CHAPTER 6

METHODS

The objective of this thesis is to determine the extent, if any, to which private insurance companies may inflate government losses through adverse selection activities based on El Nino/Southern Oscillation. To accomplish this, reinsurance decisions over the 20-year period of 1978-1997 are simulated. Assume that in each year, insurance premiums are established by the government and are based on the conditional yield density $\hat{f}_{RMA}(y_t|t)$. Alternatively, assume that an insurance company would have at their disposal information regarding SSTs. This will allow them to base their reinsurance strategy on the conditional yield density $\hat{f}_{PIC}(y_t|t, SST)$. The econometric methods used to derive the conditional yield densities $\hat{f}_{RMA}(y_t|t)$ and $\hat{f}_{PIC}(y_t|t, SST)$ are outlined in this chapter. Section 6.1 will discuss the techniques used by the RMA to estimate the process of technology in wheat production. This is followed in section 6.2 by a discussion of the techniques used in PIC estimation of wheat production. Section 6.3 will present estimates of the technology and SST processes. Section 6.4 will explain the manner in which conditional yield densities are estimated to recover the corresponding insurance premiums. Section 6.5 compares RMA and PIC estimated conditional yield densities and their associated insurance premiums. The final section will present a chapter summary and conclusions.

6.1 RMA Estimation of the Technology Process

Wheat production is generally characterized by an increase in yield over time as technology has advanced. In calculating an insurance premium, RMA will estimate the process of technological advancement to forecast the expected yield of the crop. The residuals of the estimation are used to construct the yield distribution. To most closely approximate RMA techniques in this simulation, the technological component will be estimated in the same manner used for the county yield GRP program, specifically a restricted linear spline model. A linear spline function is estimated with one knot to accommodate a change in the rate of technological advancement. This can be expressed:

$$y_t = \alpha + \beta_1 (t \times I(0,\delta)(t) + \delta(1 - I(0,\delta)(t))) + \beta_2(1 - I(0,\delta)(t)(t-\delta)) + \varepsilon_t \qquad (10)$$

where I is the indicator function. However this is estimated not with ordinary least squares but once-iterated least squares. Ordinary least squares is first used to estimate a spline function and identify outlying yield realizations. The outlying yield realizations are then windsorized to obtain a robust estimate on the second iteration. The final estimate of the function is used by the RMA to forecast yield in time period T+1 while the residuals are used to empirically construct the conditional yield density $\hat{f}_{RMA}(y_{t+1}|t)$.

6.2 PIC Estimation of Wheat Production

PIC reinsurance strategy will be based on the conditional yield density $\hat{f}_{PIC}(y_t|t, SST)$. Thus wheat yield is modeled as a function of both the technology employed and SSTs. The relationship between these processes and wheat yield is discussed in this section. To ensure that any difference between $\hat{f}_{RMA}(y_t|t)$ and $\hat{f}_{PIC}(y_t|t, SST)$ results strictly from the incorporation of SSTs, the technology process will be estimated with the linear spline in the same manner as used by RMA. However, for reasons outlined in section 6.2.2, a two-stage procedure will be used to estimate the combined technology and SST process. The technology process is first estimated using nonparametric techniques, to extract the SST process. The SST process is then combined in a model with the linear spline for the final estimation

6.2.1 The SST Process on Wheat Yield

In analyzing the relationship between ENSO and crop yield, most studies have assumed a linear relationship. This was consistent with early theory indicating a linear relationship between ENSO and weather variables such as precipitation and temperature. However as discussed in chapter 3, more recent analysis has found that the assumption of linearity does not always hold. Studies that investigate non-linearity typically employ a phase-based analysis. A phase-based approach uses some quantitative criteria to categorize ENSO conditions as either: El Nino, La Nina, or neutral. Upon delineating the phases of ENSO, one is able to distinguish any differences or trends in the subject data that may be attributed to the different phases. Though a phase-based methodology addresses the question of linearity, it fails to recognize that the temperature and pressure anomalies associated with ENSO actually fluctuate along a continuum. Given the limitations of the linear and phase-based approaches, nonparametric techniques are used to estimate the relationship between ENSO and crop yield.

6.2.2 Recovery of the SST Process on Wheat Yield

In simulating RMA modeling techniques, a restricted linear spline was used to estimate the process of technological advancement in wheat production. However, a problem arises when a spline function is used to estimate technological advancement given SSTs have been introduced to the model. Moss and Shonkwiler (1993) have pointed out that technical innovation may be a stochastic process. Ker and Coble (1998) propose that technical innovation may be a Poisson process where there is a distribution surrounding the magnitude to which an innovation may impact crop yield, as innovations are adopted neither completely nor simultaneously by all producers. As a result, the temporal trend in yield will likely be marked by random and sporadic increase as technology advances. Under these circumstances it may be difficult or inappropriate to specify a functional form *a priori*. A linear spline function as used by the RMA may not sufficiently accommodate the sporadic nature of technological advancement. This creates a problem when estimating $\hat{f}_{PIC}(y_t|t, SST)$. Any residual effect of the technology component that the spline is unable to capture may be spuriously attributed to SSTs. Therefore nonparametric techniques will initially be used to estimate the technology component of wheat yield.

In a fully nonparametric setting we would minimize

$$E\left(y - m(t, SST)\right)^2 \tag{11}$$

where $m(t, SST)$ represents the process of yield as a multivariable function of technology and SST. However, for reasons outlined in the following section it is suitable to estimate a generalized additive model of the form

$$\text{Yield} = m_1(t) + m_2(SST) \qquad (12).$$

The estimates $\hat{m}_1(t)$ and $\hat{m}_2(SST)$ may then be obtained individually using nonparametric techniques. Nonparametric estimation of $\hat{m}_1(t)$ will accommodate the sporadic nature of technological advancement, which will allow recovery of the SST component $\hat{m}_2(SST)$ without the risk of contamination presented by the linear spline. The properly estimated SST component can then be incorporated into a model with the linear spline in the manner of:

$$y_t - \hat{m}_2(SST) = \alpha + \gamma_1 (t \times I(0,\phi)(t) + \phi\,(1 - I(0,\phi)(t))) + \gamma_2\,(1 - I\,(0,\phi)(t)(t-\phi)) + \nu_t \quad (13)$$

where I is the indicator function. The spline function is employed for the final estimation so we are assured that any difference between $\hat{f}_{RMA}(y_t|t)$ and $\hat{f}_{PIC}(y_t|t, SST)$ may be attributed strictly to the incorporation of SSTs, and not contaminated by the nonparametric estimation of the technology component, $\hat{m}_1(t)$.

6.2.3 Generalized Additive Modeling

Inherent to the additive model is the assumption that the effect of SST is independent of technological advancement in wheat production. Intuition would suggest that this is the case. In fact, the assumption of independence has typically been elemental to past studies regarding SST and crop yield. In addition, independence was supported by preliminary analysis that found an interaction term (year × SST) to be statistically insignificant. These items considered suggest the additive structure is an appropriate estimation technique for the purposes of this study.

In estimating a generalized additive model we now minimize $E(y - m(t, SST))^2$ such that $m(t,SST) = m_1(t) + m_2(SST)$. Thus we recover an estimate of the closest additive approximation to the underlying multivariable function $m(t,SST)$. A thorough review of additive modeling can be found in Buja, Hastie, and Tibshirani (1989).

In estimating the model we employ the backfitting algorithm. Backfitting essentially involves estimating each component holding all others fixed, and then iterating until convergence. That is, the current estimates of $\hat{m}_2(SST)$ are obtained by regressing the residuals $\hat{\varepsilon}_{t,} = y - \hat{m}_1(t)$ against *SSTs*. Estimation of the function in an additive form offers an advantage in that each of the individual components can be estimated using the appropriate univariate techniques, and the resulting estimators provide easily interpretable information as to how the dependent variable relates to each.

6.2.4 Choice of Nonparametric Techniques

Recall that technical innovation may be a stochastic process. In addition there may be a distribution surrounding the magnitude to which an innovation may impact crop yield, as innovations are adopted neither completely nor simultaneously by all producers. As a result, the temporal trend in yield will likely be marked by random and sporadic increase as technology advances. As such, locally weighted regression smoothing, specifically an Isotonic Robust Super Smoother (IRSS) (Ker and Coble 1998), is used to estimate the technology component $\hat{m}_1(t)$.

Locally weighted regression smoothing will estimate a weighted least-squares regression at each realization (x_t, y_t) in the set. The weights are determined by a decreasing function of the distance between x_0 and the other realizations in the local neighborhood. If the underlying function were to exhibit more curvature, then a smaller neighborhood is more desirable. Higher variance would require a larger neighborhood. Thus the size of the neighborhood may vary for each observation and is chosen using local cross-validation techniques. This is termed super smoothing.

Locally weighted least squares regression smoothing is employed under the assumption that the dependent variable is normally distributed. However crop yields in this area are generally considered non-normal. Recall the RMA uses windsorizing techniques to address this issue. Therefore the super smoother is augmented with robust techniques. Specifically, the IRSS uses the default S-Plus m-estimator. This is the Huber m-estimator until convergence followed by two iterations of the Bisquare. In using the m-estimator, outlying yield realizations are essentially down-weighted during the regression smoothing. This is used in cross validation procedures as well as in the final estimates of the regression coefficients.

Finally, technological advancement is an accumulating process. As such the impact of technological advancement on wheat yield will likely be a non-decreasing function. Therefore estimates are isonotized, or restricted to be non-decreasing. This is imposed using the pool-adjacent-violators (PAV) algorithm in Hanson, Pledger, and Wright (1973). The IRSS is fully delineated in Appendix A.

There is little in the climatology or agronomy literature which discusses the underlying functional form of the relationship between crop yield and ENSO. Most studies have assumed a linear model, while more recently a phase-based approach has been used. As stated, both of these approaches face limitations. Recall that the warm SSTs indicative of El Nino are associated with above normal precipitation in the state of Texas. Alternatively, the cold SSTs associated with La Nina often result in below normal precipitation and a tendency toward drought. Considering that moisture availability and the occurrence of drought are often limiting factors in the production of wheat in Texas, this would imply that yield is likely an increasing function with SSTs.

Given the lack of guidance from previous research, it is felt that locally weighted regression smoothing will again serve as an appropriate means of estimating the underlying function. The m-estimator is used to address the non-normality of the data. The agronomy and climatology literature suggest that yield is likely an increasing function of SSTs. As such, estimates are restricted to be non-decreasing. In short, the IRSS is well suited to estimate the unknown underlying functional form of the relationship between SSTs and wheat yield.

6.3 Analysis of Technology and SST Components

The process of technological advancement in wheat production has not been uniform throughout the state of Texas. Figure 6.1 compares technological advancement as estimated by the IRSS to that of the linear spline for Coleman county Texas over the period of 1956-1997. In some counties the spline is a good approximation of technological advancement, however in many the shortfalls are evident. For instances, note the latter part of the time series for Coleman county where yield appears to level off; the IRSS is able to approximate this while the spline estimators indicate continued technological advancement

In most cases the spline estimates by the RMA and those of the PIC are nearly identical. This is not surprising as they differ only in that insurance companies have accounted for SSTs prior to estimation of the spline. To the extent that technological advancement and SSTs are essentially independent, this should have very little effect on the spline estimate. The relative independence of the two components was also evident in back-fitting the estimates, as convergence was typically achieved in 5 to 7 iterations.

Figure 6.2 depicts the estimation of the SST component for Coleman county. As expected, the cold SSTs associated with La Nina typically result in decreased wheat yield while the warm temperatures associated with El Nino typically resulted in increased yields. In general the SST component explained a larger share of inter-annual yield variability in the eastern and central portion of the state, then in the western.

In using nonparametric techniques to estimate the SST component there is some ambiguity as to the number of degrees of freedom involved in the estimation. This makes it difficult to conduct typical hypothesis testing. As a result, randomization procedures were used to test the statistical significance of the SST component in each county.

Define the R^2 correlation of the estimate as:

$$R^2 = 1 - \frac{\sum \left(y_t - \hat{m}_1(t) - \hat{m}_2(SST_t)\right)^2}{\sum y_t^{\,2}} \qquad (14).$$

Through randomization procedures, we are able to construct a distribution of the R^2 correlation under the null hypothesis that the SST component $\hat{m}_2(SST)$ is not significant. To accomplish this, the vector of SSTs is randomized when estimating equation 14. That is, in estimating $\hat{m}_2(SST)$, each observation is paired with a SST value selected randomly without replacement from the set of recorded SSTs. This can be expressed

$$\text{Yield} = m_1(t) + m_2(SST_j) \quad (15).$$

Because these SSTs are chosen randomly, any variation in yield that $\hat{m}_2(SST)$ may explain thus increasing the R^2 correlation, is therefore spurious. Repetitive trials will allow construction of the distribution of R^2 correlation under the null hypothesis that the SST component $\hat{m}_2(SST)$ is not significant. This will allow a test of the hypothesis that the R^2 correlation corresponding to the actual nonparametric estimate were drawn from this population.

The results indicated that the SST component was significant ($\alpha = 0.1$) in 50 out of 55 counties analyzed. All five counties where SSTs were found insignificant are located in a region of the Panhandle where irrigation is most common. Logically, the use of irrigation would likely result in mean yields that are less dependent on weather conditions and hence SSTs than are yields in other parts of the state.

Similar randomization procedures were used to test that the relationship between wheat yield and SSTs differed significantly from linearity. Consider that the relationship between wheat yield and SSTs were made up of two components: a linear component, and a second component responsible for any deviation from linearity. This can be expressed in the semi-parametric model

$$\text{Yield} = m_1(t) + \hat{\beta} \cdot SST_t + m_2(SST_t) \quad (16)$$

such that $\hat{\beta} \cdot SST$ is the least squares estimate of the linear relationship between wheat yield and SSTs, and $\hat{m}_2(SST)$ is the nonlinear component which can be recovered using the nonparametric techniques. [13] To determine whether the relationship between wheat yield and SSTs differs significantly from linearity is tantamount to determining whether the nonparametric component $\hat{m}_2(SST)$ is statistically significant. Repetitive trials, in which SSTs are randomized when estimating the nonlinear component $\hat{m}_2(SST)$, will allow construction of the distribution of R^2 correlation under the null hypothesis that the relationship was linear. This will subsequently allow a test of the hypothesis that the R^2 correlation corresponding to the actual nonparametric estimate were drawn from this population. The results indicated that 20 of the 50 counties in which the relationship between wheat yield and SSTs is significant, reject that the relationship may be linear.

6.4 Estimation of Conditional Yield Density

In calculating an insurance premium RMA will estimate the process of technological advancement to forecast the expected yield of the crop. The residuals of the estimation are used to construct the yield distribution. A number of techniques have been used to estimate conditional yield densities and derive crop insurance premiums in the past. Goodwin and Ker (1998) note that researchers have often used a beta distribution to estimate yield densities. This is appealing since it may accommodate skewness in the data. Others have used a gamma distribution. However the unknown

---

[13] When backfitting the semi-parametric process of SST's, the current estimate of $\hat{\beta}$ is updated by regressing the residuals $\hat{\varepsilon}_{t,} = y - \hat{m}_1(t)$ against *SST's*. Estimates of the nonparametric component $\hat{m}_2(SST)$ are then updated by regressing the residuals $\hat{\varepsilon}_{t,SST} = y - \hat{m}_1(t) - \hat{\beta} \cdot SST$ against SST's. Lastly, the PAV algorithm is used to impose the monotonicity constraint not on $\hat{m}_2(SST)$ but rather the sum $\hat{\beta} \cdot SST + \hat{m}_2(SST)$.

yield densities may not likely belong to the restricted parameter space of these distributions. Therefore nonparametric kernel techniques will be employed to estimate the conditional yield densities. Nonparametric kernel techniques offer an advantage in that they can estimate the unknown density without restricting it to a pre-specified parametric space. Nonparametric kernel density estimation is outlined in Appendix B, and is discussed in detail in Goodwin and Ker (1998).

The spline model of yield conditioned on technology, which is used by the RMA in deriving insurance premiums was defined in equation 10 as:

$$y_t = \alpha + \beta_1 (t \times I(0,\delta)(t) + \delta(1 - I(0,\delta)(t))) + \beta_2(1 - I(0,\delta)(t)(t-\delta)) + \varepsilon_t \tag{17}$$

Alternatively, the model conditioned on technology and SSTs was defined as:

$$y_t - \hat{m}_2(SST) = \alpha + \gamma_1(t \times I(0,\phi)(t) + \phi(1 - I(0,\phi)(t))) + \gamma_2(1 - I(0,\phi)(t)(t-\phi)) + \nu_t \tag{18}$$

Estimation of a conditional yield density requires a set of independent yield realizations that represent a sample from the unknown density in question. Therefore using the forecasts of yield recovered from the estimation of equations 17 and 18 and the estimated residuals $\{\hat{\varepsilon}_t \ldots \hat{\varepsilon}_T\}$ and $\{\hat{\nu}_t \ldots \hat{\nu}_T\}$, we adjust historic yield observations to a base of time period $T + 1$.

As technology has lead to higher yield over time, wheat may become susceptible to greater shortfall or show increased variability in the harvested yield. In estimating equation 17, the nonparametric peak test indicated the presence of heteroscedasticity in 44 of 55 counties. In estimation of equation 18 the presence of heteroscedasticity was indicated in 39 out of 55 counties. Therefore the appropriate corrections were made as follows.

In estimating $\hat{f}_{RMA}(y_t|t)$, the set of independent yield realizations is defined as:

$$\left(\frac{\hat{\varepsilon}_t}{\hat{y}_t}\right) \times \hat{y}_{T+1} + \hat{y}_{T+1},\ t = 1 \ldots T \qquad (19).$$

where $\hat{y}_{T+1}$ is the forecasted yield for time period $T+1$. The same procedure is followed when adjusting yield realizations to estimate $\hat{f}_{PIC}(y_t|t, SST)$, but with the exception that we use the forecast of yield and the residuals $\{\hat{v}_t \ldots \hat{v}_T\}$, as estimated by PICs and recovered from equation 18 rather than equation 17. Note that the yield forecast by PICs will contain the additive sum of the technology component estimated by the spline, as well as the effect of SSTs. Upon adjusting yields we then estimate the conditional yield densities and calculate the corresponding insurance premiums.

6.5 Comparison of Conditional Yield Densities

The relationship between yield and SSTs is evident by comparison of conditional yield densities. Figure 6.3 compares the alternative conditional yield densities $\hat{f}_{RMA}(y_t|t)$ and $\hat{f}_{PIC}(y_t|t, SST)$ for Coleman county in crop year 1991. The actual SST anomaly observed in July-September prior to planting the 1991 crop was very moderate, at only 0.213 °C. As a result there was little difference between the expected yield forecasted by PICs and that by the RMA (see Table 6.1). However notice the reduction in variance exhibited by $\hat{f}_{PIC}(y_t|t, SST)$ in comparison to $\hat{f}_{RMA}(y_t|t)$. Because the variation in yield attributed to SSTs has been explained, there is a smaller portion of the probability mass in the tails of $\hat{f}_{PIC}(y_t|t, SST)$ while a larger portion in the neighborhood of the expected yield. Although the difference in forecasted yield was only slight, when combined with the reduced variance there results lower premium rates calculated by the PICs than the RMA (Table 6.2). Coleman county ultimately produced a wheat yield of 16.97 bushels/acre in 1991, which resulted in a net profit at all coverage levels.

However unlike 1991, strong La Nina conditions were developing prior to planting for crop year 1989. The July-September SST anomaly of –1.38°C resulted in a substantial decrease in expected wheat yield for Coleman county (Table 6.1). Although the RMA had estimated an expected yield of 21.34 bushels/acre, the PICs included SSTs into their estimation to forecast a yield of only 11.11 bushels/acre. Figure 6.4 compares the subsequent conditional yield densities for Coleman county. Notice that $\hat{f}_{PIC}(y_t|t, SST)$ still exhibits less variance than $\hat{f}_{RMA}(y_t|t)$, however it has shifted downward to reflect the lower expected yield. As a result of the downward shift, the corresponding PIC premium rates which more accurately reflect the risk in production are substantially higher than those estimated by the RMA (Table 6.2). Dry conditions associated with the low SSTs resulted in Coleman county producing a wheat yield of only 13.40 bushels/acre in 1989 which resulted in no indemnities paid at the 60% coverage level but net losses at the 70, 80, and 90% levels.

6.6 Summary and Conclusions

To determine the extent to which PICs may recover excess profits at the expense of excess government loss, reinsurance decisions over the period of 1978-1997 are simulated. It is assumed that in each year, insurance premiums are established by the government and are based on the conditional yield density $\hat{f}_{RMA}(y_t|t)$. A linear spline model is used to estimate the process of technological advancement. This will enable a forecast of yield in time period $T+1$. The residuals of the regression are then used to empirically construct the yield density $\hat{f}_{RMA}(y_t|t)$.

Alternatively, it is assumed that a PIC would base their reinsurance strategy on the conditional yield density $\hat{f}_{PIC}(y_t|t, SST)$. The linear spline is limited in its ability to approximate the process of technological advancement (figure 6.1). Therefore the processes of technology and SSTs are estimated in two stages using a generalized additive model. First, nonparametric techniques are used to estimate the technology and SST process. This will allow extraction of the SST process without contamination resulting from the limitations of the linear spline. Secondly, the SST process is inserted into an additive model with the linear spline. This is used to forecast yield in time period $T+1$ and estimate $\hat{f}_{PIC}(y_t|t, SST)$. By using the linear spline in the final model we are assured that any difference between $\hat{f}_{RMA}(y_t|t)$ and $\hat{f}_{PIC}(y_t|t, SST)$ results strictly from the inclusion of SSTs.

Figures 6.3 and 6.4 presented comparisons of $\hat{f}_{RMA}(y_t|t)$ and $\hat{f}_{PIC}(y_t|t, SST)$. In years of moderate or warm SSTs the reduced variance associated with $\hat{f}_{PIC}(y_t|t, SST)$ resulted in lower premiums estimated by PICs than the RMA. However in years of strong negative SSTs, the downward shift associated with $\hat{f}_{PIC}(y_t|t, SST)$ resulted in higher premiums estimated by PICs than the RMA.

		Forecasted yield	
Year	**SST anomaly**	**RMA**	**PIC**
1991	0.213	19.62	20.81
1989	-1.38	21.34	11.11

Table 6.1: Comparison of forecasted yield for Coleman county.

	1991 Premium rates		**1989 Premium rates**	
Coverage level %	**RMA**	**PIC**	**RMA**	**PIC**
60	0.28	0.09	0.31	2.11
70	0.56	0.21	0.62	4.05
80	0.97	0.42	1.07	6.14
90	1.55	0.77	1.70	8.27

Table 6.2: Comparison of premium rates for Coleman county.

Figure 6.1: Estimated technology process, Coleman county

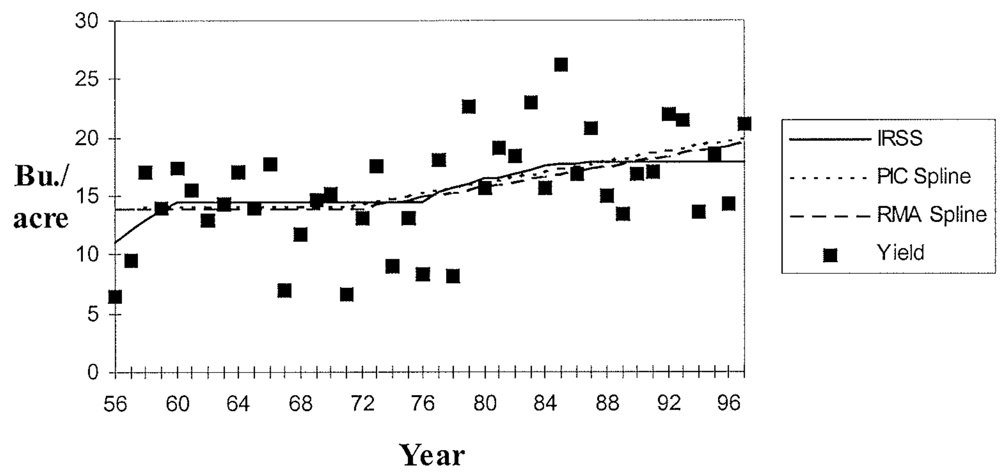
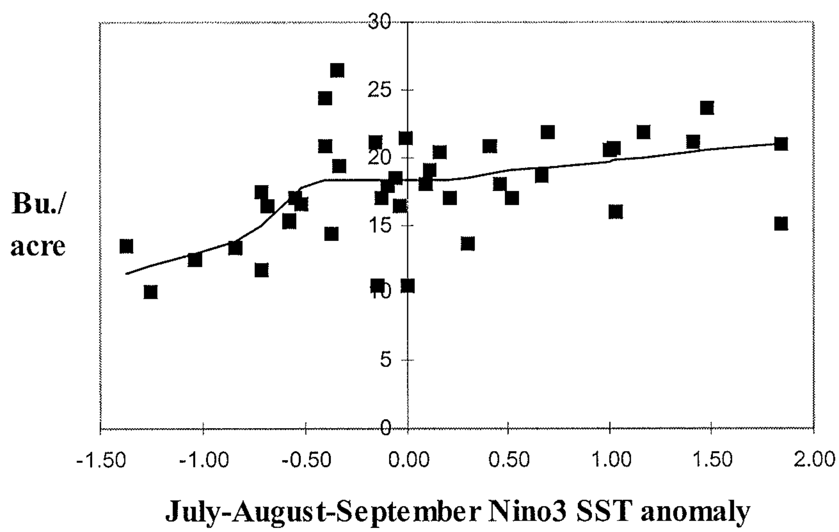


Figure 6.2: Estimated SST process, Coleman county



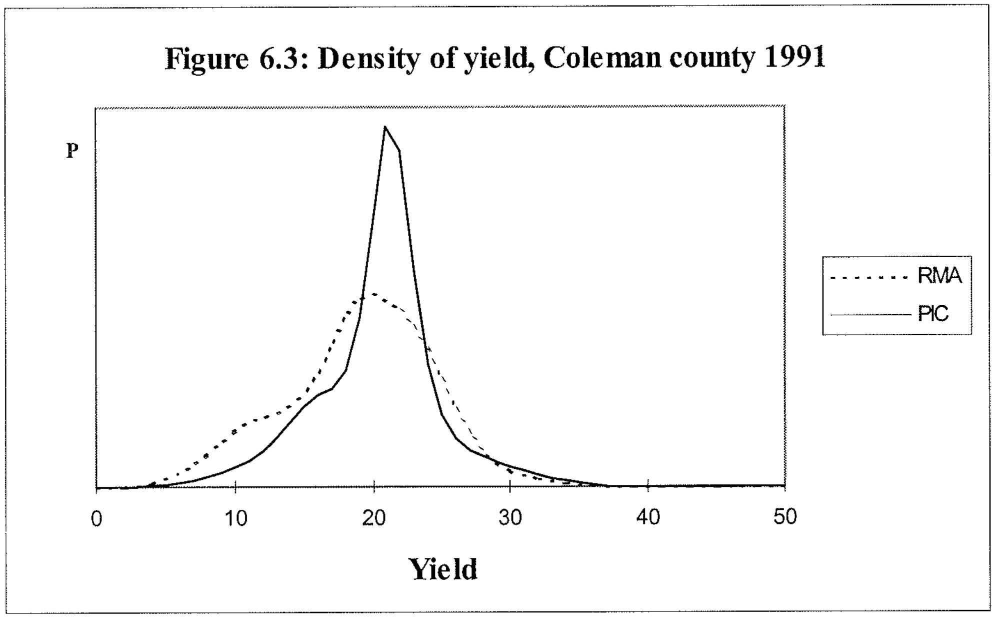


Figure 6.3: Density of yield, Coleman county 1991
P
RMA
PIC
0
10
20
30
40
50
Yield

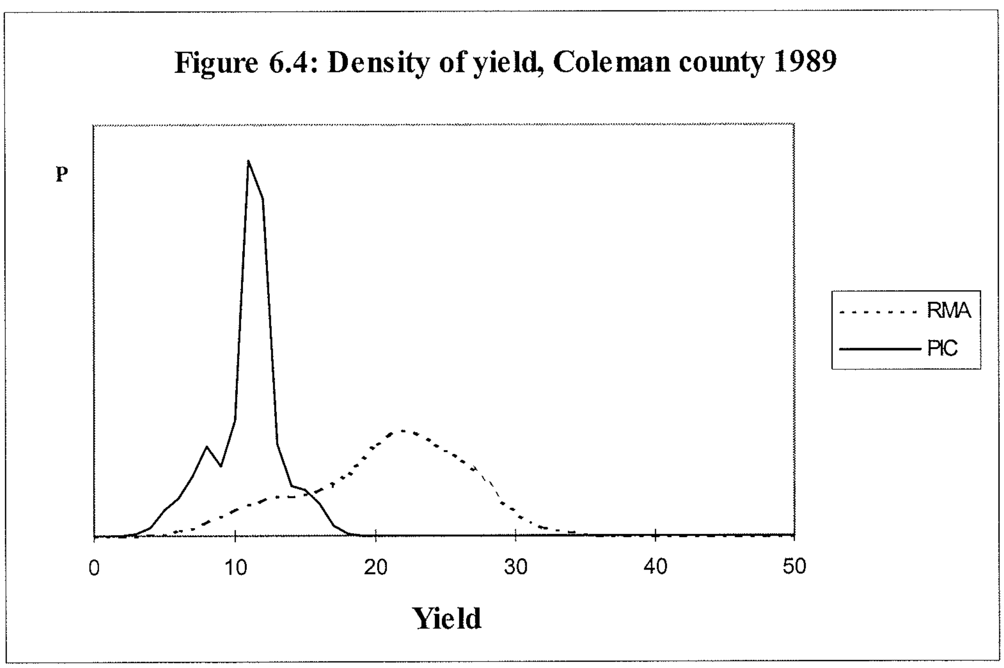


Figure 6.4: Density of yield, Coleman county 1989
P
RMA
PIC
0
10
20
30
40
50
Yield

CHAPTER 7

SIMULATION ANALYSIS AND DISCUSSION

To determine the extent to which ENSO based adverse selection may increase government losses, we simulate the reinsurance decisions of a private insurance company over the 20-year period of 1978-1997. The general guidelines are outlined in section 7.1. The reinsurance simulation is repeated under 3 different scenarios. In the first scenario, PIC decision strategy is dependent on their expectation of profit subject to the constraints of the 1998 SRA. Each contract is evaluated individually under the assumption that the loss ratio of each fund would follow the loss ratio of the contract in question. In the second scenario, PIC decision strategy is based on the expected loss of each contract. The underwriting gains and losses are subject to the constraints of the SRA. In the third scenario, PIC decision strategy is independent of the SRA, and they may either accept or cede to the government 100% of the underwriting gains or losses as they choose. These scenarios are presented and explained in more detail in section 7.2. They allow a comparison of the results both with and without the constraints of the 1998 SRA. Upon discussing the scenarios, a summary of the findings is presented in section 7.3.

7.1 Simulation Guidelines

There are 55 Texas counties included in the study, and each county-year observation is considered one insurance contract to be reinsured. Therefore 55 contracts are evaluated in each year for a total of 1100 contracts over the 20-year period of the simulation. Each contract is weighted by the average number of acres insured in that county over the period of 1995-1998. The simulation is performed at the 60,70, 80, and 90% coverage levels.

Premiums are established on a contract by contract basis and we assume that rates are set one year in advance by the RMA. For example the premium for 1978 is based on the conditional yield density as estimated from the information set $I_{RMA} = \{y_{1956} \ldots y_{1977}\}$. Alternatively, PICs will have at their disposal information regarding SSTs. Therefore they would base their decision strategy on the conditional yield density as estimated from the information set $I_{PIC} = \{y_{1956} \ldots y_{1977}, SST_{1956} \ldots SST_{1978}\}$.[14] The respective information sets are then updated in each successive year of the simulation. Thus the premium for 1979 is based on the information set $I_{RMA} = \{y_{1956} \ldots y_{1978}\}$.

Historic yield realizations are used to determine the profit or loss on each contract. This will allow calculation of the total profit or losses that are accrued by insurance companies, the government, and the overall program over the 20-year period. In addition, *pseudo* loss ratios are calculated for insurance companies, the government, and the program overall. Because profits and losses under the SRA are shared between the government and PICs, there is no delineation as to the actual premiums or indemnities accredited to each. This prohibits the calculation of actual loss ratios for each party. Thus *pseudo* loss ratios are calculated for the government, PICs, and the program overall. The *pseudo* loss ratio is defined as

[14] Note that the conditions of SST_{1978} are observed in the months of July-August-September 1977, just prior to the deadline date for 1978 reinsurance decisions.

$$\frac{\textit{total losses}}{\textit{total profits}} \tag{20}$$

where *total losses* = the sum of the losses for those years in which PICs [or government] realized a net loss, and *total profits* = the sum of the profits for those years in which PICs [or government] realized a net profit.

7.2 Simulation Analysis and Results

Recall from chapter 5 that the 1998 SRA designates three funds to which a contract may be assigned: the commercial, developmental, and assigned risk. The government and PICs share any underwriting gains or losses from each fund in a given year by a two-tiered risk-sharing structure. At the first tier, an insurance company will retain a predetermined share of the profit or loss (see table 5.1). At the second-tier, the shares are determined by the loss ratio of the fund (see table 5.2).

The reinsurance simulation is repeated under three different scenarios. In the first, assume that the loss ratio of each fund would be entirely dependent on the loss ratio of the contract in question. Thus the fund loss ratios would equal the contract loss ratio. PIC reinsurance decisions are then based on their expectation of profit under the constraints of the 1998 SRA. This is termed the *dependent* scenario, and is presented in section 7.2.1. In the second scenario, assume that the loss ratio of each contract is independent of the loss ratios of the different funds. Reinsurance decisions are then based on the expected loss of each contract. However, the underwriting gains and losses are subject to the constraints of the SRA. This is termed the *independent* scenario, and is presented in section 7.2.2. In the third scenario PICs are fully unconstrained by the SRA. Their decision strategy is independent of the SRA, and they may either accept or cede to the government 100% of the underwriting gains or losses as they choose. The *unconstrained* scenario is presented in section 7.2.3.

In the *dependent* scenario, beginning in 1978 we estimate $\hat{f}_{PIC}(y_t|t, SST)$ for every contract. As demonstrated in section 5.2 this will allow calculation of the profit a PIC would expected from each of the three funds. Each of the 55 contracts is then placed in the fund that offers the highest expected profit. Using historic yield realizations for 1978, the loss ratios for each fund are then calculated and the gains or losses divided between the government and PICs accordingly. This process is repeated for each year of the simulation to recover the total profit or losses that are accrued by PICs, the government, and the overall program over the 20-year period. In addition, *pseudo* loss ratios are calculated for each party.

Randomization procedures were used to test the statistical significance of ENSO-based reinsurance strategy. From the original simulation, it is determined how many contracts are placed in each of the three funds. The simulation is then repeated for 1000 trials in which contracts are assigned randomly to one of the three funds. However, the random trials are restricted so that the total number of contracts assigned to each fund is equal to the quantity so assigned during the original simulation. For example, suppose that during the original simulation 150 contracts were placed in the assigned risk fund, 100 in the developmental fund and 850 in the commercial fund. For each random trial, 150 contracts are placed in the assigned risk fund, 100 in the developmental fund and 850 in the commercial fund. By comparing the results from the original simulation, to those from the random trials, the statistical significance of this reinsurance strategy can be evaluated. Probability values are calculated to indicate the likelihood that ENSO based adverse selection did not result in higher [lower] PIC [government] profits and lower [higher] PIC [government] loss ratios.

In the *independent* scenario of the simulation, PIC reinsurance decisions are made under the assumption that the loss ratio of each contract is independent of the loss ratios for the three reinsurance funds. However, the underwriting gains and losses are subject to the constraints of the SRA. Using $\hat{f}_{PIC}(y_t|t, SST)$ PICs will calculate the expected loss on a contract, which will differ from the premium established by the RMA using $\hat{f}_{RMA}(y_t|t)$. If they estimate the expected loss to be more than reflected in the premium, they will assign the contract to the commercial fund. Alternatively, if they estimate the expected loss to be higher than reflected in the premium, they will place it in the developmental or assign risk fund. As with the first scenario, this is repeated for each contract to recover the total profit or losses that are accrued by PICs, the government, and the overall program for the 20-year period. In addition, the loss ratios and pseudo loss ratios are calculated for each of these parties. Randomization procedures similar to those in the first scenario are used to test the statistical significance of ENSO based adverse selection.

In the *unconstrained* scenario of the simulation, insurance companies are unrestricted in their reinsurance decisions. Using $\hat{f}_{PIC}(y_t|t, SST)$ they will calculate the expected loss on a contract, which will differ from the premium established by the RMA

$\hat{f}_{RMA}(y_t|t)$. If they estimate the expected loss to be more than reflected in the premium, they will accept 100% of the contract liability and profits or losses. Alternatively, if they estimate the expected loss to be higher than reflected in the premium, they will cede 100% of the liability and profits or losses to the government. This is repeated to calculate the total profits and losses, pseudo loss ratios, and loss ratios for the government, PICs, and the program over the twenty-year period. This scenario will provide a base for comparison, to determine the effect of the SRA on PIC adverse selection.

Randomization procedures are again used to test the statistical significance of this strategy. The procedure is repeated for 1000 trials in which 100% the liability of each contract is randomly assigned to either the government or PICs. However, the random trials are restricted so that contracts are assigned to the government and PICs with the same likelihood to which they were assigned in the original simulation. The profits, losses, and loss ratios from the random trials are then compared to those of the original simulation to evaluate the statistical significance of ENSO-based adverse selection.

7.2.1 Simulation Results, *Dependent* Scenario

Results of the simulation indicate that overall program loss ratios for the period of consideration range from 0.90 at the 60% coverage level to 1.39 at the 90% level (Table 7.1). These values are consistent with the historical performance of the program, as was outlined in chapter 2. Section A of table 7.2 presents a comparison of PIC and government profits and loss ratios when reinsurance decisions were made under the assumption that contract loss ratios were equal to fund loss ratios. Thus, contracts were placed in whichever of the three reinsurance funds offered the highest expected profit.

At the 3 lowest coverage levels PICs retained a profit while the government suffered losses, and at the 90% level PICs suffered a substantially smaller loss than the government. Likewise, PIC loss ratios were substantially lower at all levels. Excess PIC profit due to ENSO-based adverse selection was statistically significant at all coverage levels (see Pvalues in table 7.2).

Figure 7.1 compares the cumulative distributions of PIC profit for the commercial and assigned risk funds for crop year 1991. If PICs were to assign this contract to the commercial fund in a year of moderate SSTs such as 1991, there is a risk that they may suffer a loss of over 3 bushels/acre. However they are much more likely to retain a fair profit in the commercial fund, so the expected value of the contract was 0.48 bushels/acre. In comparison, the negligible share for which they are liable in the assigned risk fund is evident. They were not likely to realize a significant loss or profit, resulting in an expected value of 0.08 bushels/acre. In a case such as this a PIC would assign this contract to the commercial fund.

At all coverage levels a large majority of the 1100 contracts are assigned to the commercial fund. Intuition may suggest that negative SST anomalies are associated with below average yields and high risk, while positive SST anomalies are associated with above average yields and low risk, resulting in approximately 50% of the contracts placed in the commercial fund and the other 50% in the assigned risk fund. The large percentage placed in the commercial fund may then appear inordinately high, but it results of several factors. Consider figure 7.2, which compares the cumulative distributions of government and PIC profit for the commercial fund, on a Colby county contract at the 90% coverage level in crop year 1991. If premiums were actuarial fair and risk sharing between the government and PICs equal, then there would be 0 expected profit for each party. However recall from section 6.3 that the reduced variance associated with $\hat{f}_{PIC}(y_t|t, SST)$ indicates that in a moderate year such as 1991 (SST anomaly = 0.213) the risk in production is less than reflected in the premium. As a result there is an expected profit which is shared between the two parties and equal to the difference between the RMA premium and the PIC premium:

$$RMA\ premium - PIC\ premium \text{ (see table 6.2)} = 1.55 - 0.77 = 0.78 \qquad (21).$$

Thus the government and PIC expected profits as shown on figure 7.2 were 0.30 and 0.48 bushels/acre respectively. Similarly, even when SST anomalies are negative there may still be less likelihood of a shortfall and a higher expected profit than reflected in the premium calculated by the RMA. This will increase the percentage of contracts that PICs wish to place in the commercial fund.

A second reason for the large number of contracts placed in the commercial fund results from the risk sharing structure of the SRA. In particular, recall the 2nd tier of the commercial fund as outlines in table 5.2. Note the asymmetry by which profits and losses are divided between the government and PICs. When yields are favorable and the loss ratio is below 1, PICs tend to share a larger portion of the profits than they do the losses when there is a shortfall and the loss ratio is above 1. This too can be seen in figure 7.2. In the event of no shortfall they share the profits almost evenly, with the government retaining 0.772 and PICs 0.778 bushels/acre. Thus the maximum profit each may retain as indicated on the figure is almost equal. However in the event of a shortfall, particularly a severe one, the government may absorb a loss of up to 13 bushels/acre while PICs less than 4. Because of this asymmetry, potential PIC loss due to shortfall is often offset by the larger share they will retain in the event of no shortfall, making the commercial fund again the most desirable in which to place a contract. Only in extreme cases would the assigned risk fund be the option of choice.

Figure 7.3 compares the cumulative distribution of PIC profit for the commercial and assign risk funds for crop year 1989. The associated strong negative SST anomalies and low expected yield indicated by $\hat{f}_{PIC}(y_t|t, SST)$ (figure 7.4) have resulted in distributions of PIC profit which reflect the high likelihood of a loss in each fund. As a result the expected profit of the commercial fund was -1.82 bushels/acre while in the assigned risk a negligible –0.04 bushels/acre. With an expected net loss in each, a PIC would choose to put this contract in the assigned risk fund. Figure 7.4 compares the distributions of government profit for the commercial and assign risk funds in 1989. Notice the government is more likely to sustain larger losses in the assigned risk fund, where the PIC share is negligible, then in the commercial where it is more considerable.

Finally, notice in table 7.2 that a larger number of contracts were placed in the assigned risk fund as the coverage level increased. Consider a case where $\hat{f}_{PIC}(y_t|t, SST)$ lies on the lower tail of $\hat{f}_{RMA}(y_t|t)$. Since $\hat{f}_{PIC}(y_t|t, SST)$ exhibits less variance then $\hat{f}_{RMA}(y_t|t)$, it may still contain less probability mass in its tails. Therefore at lower coverage levels, $\hat{f}_{PIC}(y_t|t, SST)$ would still indicate less likelihood of a shortfall then $\hat{f}_{RMA}(y_t|t)$, making the commercial fund more favorable. However, as coverage level increased the high probability of shortfall associated with the main body of $\hat{f}_{PIC}(y_t|t, SST)$ would indicate the contract was not favorable, so that it the assigned risk fund was more favorable.

To better distinguish PIC profit gained directly through ENSO-based adverse selection activities, from those resulting of asymmetry in the SRA, the scenario was repeated for which PIC reinsurance decisions were made without regard to SSTs. That is, for a given contract the expected PIC profit from each of the three reinsurance funds was calculated using the conditional yield distribution $\hat{f}_{RMA}(y_t|t)$ rather than $\hat{f}_{PIC}(y_t|t, SST)$. Not surprisingly given the asymmetric risk sharing structure of the SRA, when SSTs were excluded from reinsurance strategy a PIC would maximize their expected profit by assigning all contracts to the commercial fund. The resulting loss ratios and profits are presented in section C of table 7.2. The government still bore a majority of the losses and suffered higher loss ratios than PICs. Table 7.3 compares government and PIC profits and losses with and without ENSO-based adverse selection activities. By engaging in ENSO-based adverse selection, insurance companies were able to lower their loss ratios while increasing profit substantially. This came at the direct expense of the government however, which saw higher loss ratios and net loss increases ranging from 12% to almost 40% at the various coverage levels.

Recall from chapter 5 that in the state of Texas a PIC may allot no more than 75% of their reinsured contracts to the Assigned risk fund. Therefore, a situation may exist where the assigned risk fund is an unavailable option for reinsurance. To address this possibility, the *dependent* scenario was repeated a third time in which PIC reinsurance options were restricted to the commercial and developmental funds. In this setting the developmental fund is viewed as the preferred option in instances where PICs wish to cede a contract to the federal government. As such, it is assumed that they would opt to retain the minimum 35% share of first tier liability (see table 5.1).

Even without availability of the assigned risk fund, ENSO-based adverse selection resulted in reduced loss or increased profit for PICs at the expense of increased government loss (table 7.2, section B). Likewise PIC loss ratios were lower and government's higher than when SSTs were excluded from analysis and contracts were assigned to the commercial fund. The higher liability shares of the developmental fund in comparison to the assigned risk resulted in PIC gains and government losses that were not as large as when afforded the use of the latter fund. However the results were again statistically significant (see table 7.2). This demonstrates that the developmental fund, like the assigned risk fund, is also a viable outlet for PICs to shift a significant share of unwanted contracts to the federal government.

7.2.2 Simulation Results, *Independent* Scenario

To this point reinsurance decisions have been determined by projecting the terms of the SRA onto the conditional yield distribution associated with each contract.

However the shares stipulated in the SRA are determined not by the loss ratio associated with each individual contract, but by the loss ratios of each fund overall. Therefore an informed reinsurance decision would require not only the distribution of yield for a contract in question, but the joint distribution for all that will be assigned to each fund. This information is not attainable since it would require an immense number of possible yield combinations and knowledge as to what contracts will be placed in each fund. Rather than assume that the yield distribution for each contract is an appropriate substitute for the joint distribution of an overall fund, an alternative is to assume that the distribution of each contract is independent of the fund. A PIC could then evaluate the profitability of each contract independent of the SRA and then assign it to the various funds accordingly. For instance, if a PIC were to estimate the expected loss (the PIC premium as in table 6.2) to be lower than calculated by the RMA (the RMA premium) the contract would be placed in the commercial fund. If they were to estimate the expected loss to be higher than calculated by the RMA it, would be put in the assigned risk or developmental fund.

Table 7.2 sections D and E present the results when undesirable contracts were placed in the assigned risk and developmental funds respectively. When the advantageous structure by which profits and losses are divided in the SRA, were removed from PIC decision strategy, they opt to place fewer contracts in the commercial fund. Subsequently the number allotted to the assigned risk or developmental funds increased dramatically. Although carrying less liability, PIC profits were similar to those under previous scenarios. Randomization procedures again indicated that PIC profit attributed to ENSO-based adverse selection was statistically significant regardless of whether the assigned risk or developmental fund were used to shift unwanted contracts to the federal government. Finally, notice that the tendency to place a higher number of contracts in the assigned risk fund as the coverage level increases, is even more clearly evident when the asymmetry of the SRA is removed from PIC decision strategy.

7.2.3 Simulation Results, Unconstrained by the 1998 SRA

For comparison purposes the simulation was lastly repeated such that PICs were not constrained by the terms of the 1998 SRA, and were free to accept themselves or cede to the government 100% of the liability as they saw fit. Under this scenario the decision to accept or cede a contract is again based on the expected loss of a contract. If a PIC were to estimate the expected loss (the PIC premium) to be less than as calculated by the RMA (the RMA premium) the contract liability would be retained. Conversely if a PIC were to estimate the expected loss to be higher than calculated by the RMA, the contract liability would be ceded to the government. The results are presented in table 7.4. Though PICs would retain most of the contracts (72-88%) they would successfully avert a majority of the losses. As with previous scenarios, excess PIC profit recovered by adverse selection was statistically significant. However the overall discrepancy between PIC and government profit was less than under the constraints of the SRA since the asymmetry of the risk sharing is no longer a contributing factor.

7.3 Summary of Findings

These simulations over the years of 1978 to 1997 have considered PIC reinsurance strategy under three scenarios. The loss ratio of each fund, which under the 1998 SRA determines the share of underwriting gains or losses retained by PICs, depends jointly on every contract placed within it. Since this information is unobtainable, assumptions were made to circumvent the problem. In the *dependent* scenario, the reinsurance decision on each contract was made under the assumption that the loss ratio of each fund would be entirely dependent on the contract in question. Contracts were then placed in the fund that offered the highest expected profit. In the *independent* scenario, fund loss ratios were assumed to be entirely independent of each contract. If PICs estimated the expected loss on a contract to be less than reflected in the premium, it was placed in the commercial fund. If they estimated the expected loss to be greater than reflected in the premium, it was placed in the assigned risk or developmental fund. In the final scenario, PICs were unconstrained in their reinsurance decisions and either accepted or ceded to the government 100% of the underwriting gain or loss on each contract.

Having considered the results of the analysis, there are some generalizations that can be made about the findings. Under all scenarios, insurance companies were able to recover excess profits thus reducing their loss ratios by engaging in ENSO based adverse selection activities. These excess profits come at the direct expense of the federal government, which subsequently saw losses inflated by 12% to almost 40% at various coverage levels. Furthermore, excess PIC profits gained through ENSO based adverse selection, were found to be statistically significant under every scenario.

When looking more specifically at the effect of the 1998 SRA, several interesting points emerge. The asymmetric structure by which shares are divided under the 1998 SRA found the government bearing substantially more losses over time than PICs even regardless of ENSO based adverse selection. When PICs consider the asymmetry of the SRA into their reinsurance strategy, there result minimal instances where they would not opt to assign a contract to the commercial fund.

Also notable, when an insurance company would place a contract in the assigned risk fund, they would carry a nearly negligible portion of the liability. Thus when SST conditions would dictate, it was a most profitable option to cede undesirable contracts to the government. However in the event that the assigned risk fund were unavailable, the developmental fund serves as a viable outlet as well. There was surprisingly little difference in the results when using the assigned risk or developmental fund to cede unwanted contracts.

% coverage	Loss ratio
60	0.90
70	1.14
80	1.35
90	1.39

Table 7.1: Program loss ratio by coverage level.

	%	#Contracts by fund			*Pseudo* Loss Ratio				Profit			
	Level	Com	Dev	AR	RMA	Pvalue	PIC	Pvalue	RMA	Pvalue	PIC	Pvalue
	60	1070	1	29	1.29	0.022	0.45	0.028	-1.02	0.026	2.14	0.026
A	70	1064	4	32	2.22	0.001	0.57	0.004	-6.66	0.004	3.04	0.004
	80	1047	10	43	3.13	0.000	0.97	0.001	-18.62	0.001	0.28	0.001
	90	991	49	60	3.84	0.000	1.11	0.000	-38.16	0.000	-2.01	0.000
	60	1071	29	-	1.27	0.028	0.46	0.03	-0.97	0.031	2.08	0.031
B	70	1068	32	-	2.19	0.003	0.59	0.008	-6.48	0.007	2.85	0.007
	80	1053	47	-	3.10	0.000	1.01	0.000	-18.26	0.000	-0.08	0.000
	90	1025	75	-	3.78	0.000	1.18	0.000	-36.71	0.000	-3.46	0.000
	60	1100	-	-	1.21	-	0.52	-	-0.74	-	1.85	-
C	70	1100	-	-	2.05	-	0.70	-	-5.74	-	2.11	-
	80	1100	-	-	2.92	-	1.16	-	-16.60	-	-1.73	-
	90	1100	-	-	3.56	-	1.36	-	-33.25	-	-6.92	-
	60	971	-	129	1.35	0.000	0.33	0.002	-1.32	0.000	2.43	0.000
D	70	951	-	149	2.03	0.016	0.55	0.02	-6.44	0.029	2.81	0.029
	80	876	-	224	2.77	0.017	0.99	0.055	-18.47	0.054	0.14	0.054
	90	797	-	303	3.23	0.009	1.06	0.000	-39.21	0.000	-0.96	0.000
	60	971	129	-	1.33	0.006	0.36	0.008	-1.24	0.009	2.35	0.009
E	70	951	149	-	2.03	0.026	0.57	0.024	-6.31	0.039	2.69	0.039
	80	876	224	-	2.83	0.018	1.01	0.058	-18.28	0.057	-0.06	0.057
	90	797	303	-	3.32	0.01	1.10	0.002	-38.58	0.002	-1.59	0.002

Table 7.2: Simulation results under the constraints of the 1998 SRA. Profits reported in millions of bushels.

A. Reinsurance options open to all three funds.

B. Reinsurance options restricted to commercial and developmental fund.

C. SST's excluded from reinsurance strategy and all contracts assigned to commercial fund.

D. Reinsurance decision based on expected loss; assigned risk fund available.

E. Reinsurance decision based on expected loss: assigned risk fund unavailable.

	PIC			RMA		
%	**Without EBAS**	**With EBAS**	**Change**	**Without EBAS**	**With EBAS**	**Change**
60	1.85	2.14	15%	-0.74	-1.02	39%
70	2.11	3.04	44%	-5.74	-6.66	16%
80	-1.73	0.28	116%	-16.60	-18.62	12%
90	-6.92	-2.01	71%	-33.25	-38.16	15%

Table 7.3: Change in profits upon the introduction of ENSO-based adverse selection (EBAS).
Reported in 1,000,000's of bushels.

Level	Contracts retained by PIC's	*Pseudo* Loss Ratio				Loss Ratio				Profit			
		RMA	pval	PIC	pval	RMA	pval	PIC	pval	RMA	pval	PIC	pval
60	971	6.08	0.000	0.45	0.000	3.36	0.000	0.63	0.000	-2.72	0.026	3.83	0.026
70	951	4.13	0.000	0.84	0.000	2.59	0.000	0.92	0.000	-5.38	0.004	1.76	0.004
80	876	3.30	0.015	1.51	0.006	1.92	0.003	1.20	0.004	-9.74	0.001	-8.59	0.001
90	797	3.92	0.004	1.60	0.000	1.82	0.001	1.22	0.000	-24.59	0.000	-15.58	0.000

Table 7.4: Simulation results unconstrained by the 1998 SRA.

Profits and losses are reported in 1,000,000's of bushels.

Figure 7.1: PIC profit by fund

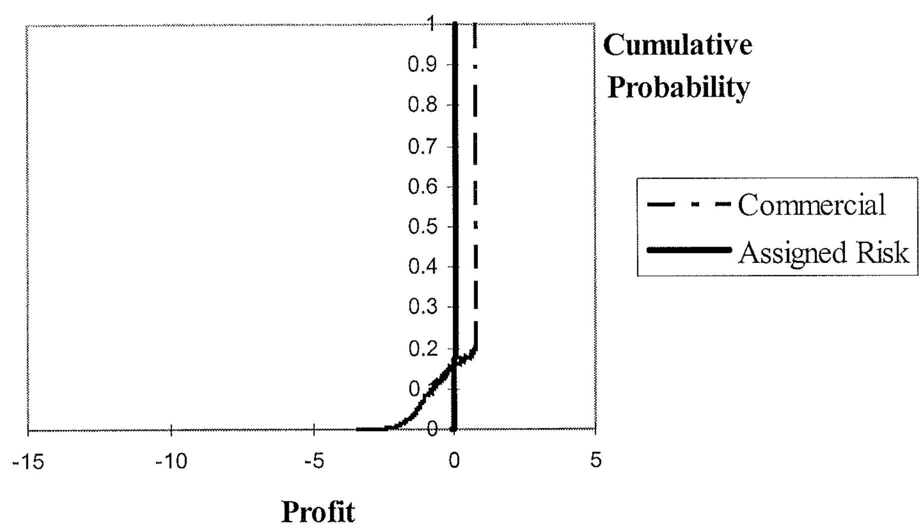


Figure 7.2: Commercial Fund

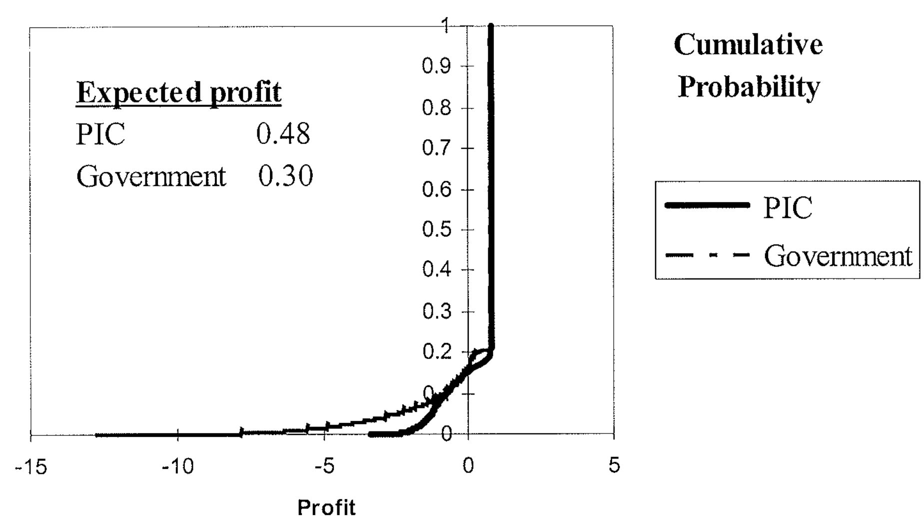


Figure 7.3: PIC profit by fund

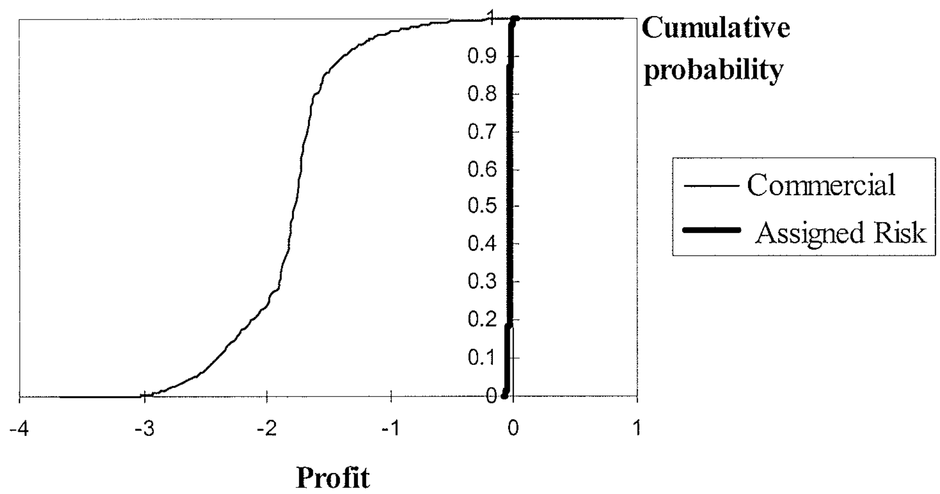
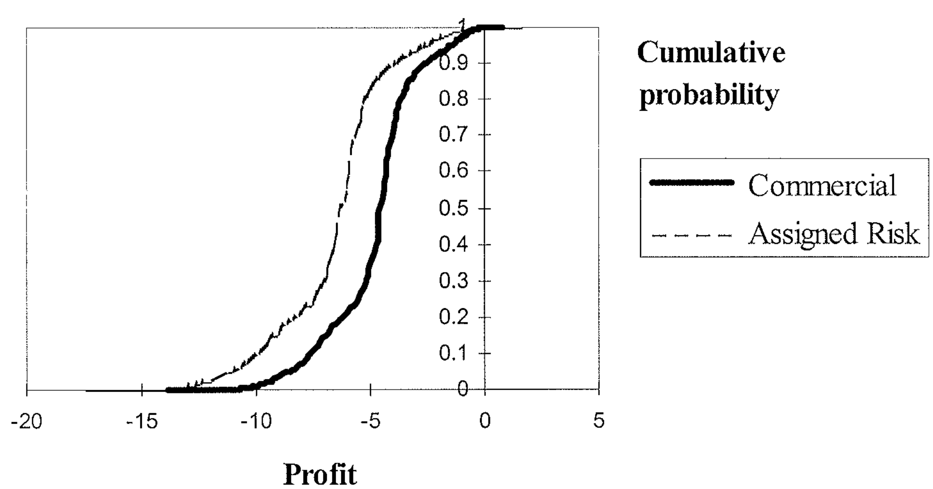


Figure 7.4: Government profit by fund



CHAPTER EIGHT

SUMMARY AND CONCLUSIONS

Since their inception over 60 years ago, government supported crop insurance programs have been persistently plagued by low participation and poor actuarial performance. Though initially available only to wheat farmers in limited geographic regions and guaranteeing only against yield shortfall, federal crop insurance has now broadened in scale and scope. Following early expansions to cover cotton, larger land areas and other major grains, it eventually became available for such diverse agricultural products as greenhouse plants and winery grapes. Most recently introduced have been revenue guarantees. If a farmer does not receive the revenue guaranteed by contract, whether it is due to yield shortfall, low commodity price or both, he/she will collect an indemnity payment. Thus through market forces insurance policies are impacted by production decisions, weather, and other factors occurring worldwide.

Due to expanded programs, premium subsidies and attempted cutbacks in federal disaster relief, participation in U.S. crop insurance has increased dramatically in recent years. The added numbers have done little to improve fiscal performance however, as 1998 saw a total cost to the government of about $2.1 billion (USDA-OIG 1999). In addition Congress in 1998 passed over $1 billion in supplemental appropriations for crop losses resulting from widespread disaster. This same year however, insurance companies collected an estimated $759 million in revenue consisting of underwriting gains and administrative fees paid for by the government. The period 1995 through 1998 saw $8.8

billion in government expenditures on crop insurance while private insurance companies received about $2.8 billion in revenue. In an attempt to gain control of program losses, Congress has set a future loss ratio target of 1.075. The government will also attempt to reduce the disparity between their losses and PIC profits.

Historically the largest factor inflicting crop loss in the U.S. has been drought, or more broadly weather variability in general. However recent years have seen great strides in advanced forecasting of seasonal weather, which offers to lessen the vulnerability of agriculture to such vagaries through improved planning and risk management. Crop insurance programs will surely play an important part of this strategy.

This study has focused specifically on the relationship between Nino3 SSTs and Texas wheat yields, and the implications that this relationship may have on the federal crop insurance program. Nino3 SSTs are a commonly used indicator of ENSO, which exhibits a strong influence on Texas climate. As a result, Texas wheat yields were found to be correlated with Nino3 SSTs observed prior to the deadline date for reinsurance decisions on wheat contracts. By incorporating this additional information into the yield distribution of the crop, a private insurance company would be able to more accurately than the RMA assess the forthcoming risk in production, and could make reinsurance decisions accordingly.

Insurance premiums are established by the RMA over a year in advance and are based on the yield distribution conditioned on technology. At this time, information regarding ENSO conditions during the growing season is not yet available. However insurance companies can base their decision strategy on the yield distribution conditioned on technology and SSTs. Figures 7.1 and 7.2 demonstrated examples of this distribution which will offer an alternative expected yield (table 6.1), and because the variation resulting from changes in SST has been explained, it will have less variance. Subsequently insurance premiums conditioned on SST differed from those of the RMA as well (table 6.2). Wheat yield in Texas is generally an increasing function of SST, so the lowest SSTs were associated with higher production risk than estimated by the RMA while moderate and warm SSTs were generally associated with lower risk.

The 1998 SRA, by which reinsurance terms are stipulated, is such that any profits or losses on crop contracts are not shared evenly or symmetrical between insurance companies and the government. There are three funds to which a PIC may assign a crop contract: the commercial, developmental, or assigned risk. They are responsible for varying portions of liability in each, although PICs generally retain a larger share of any profit than they do of a loss. The asymmetry of the risk sharing structure essentially becomes an implicit subsidy to the private insurance industry, which regardless of adverse selection activities would result in the federal government bearing responsibility for a majority of crop losses over time. Furthermore, via the assigned risk fund and to a lesser degree the developmental fund, the SRA allows insurance companies to cede all but a negligible portion of selected liability to the government. In doing so companies are afforded the opportunity to evaluate the potential profitability of contracts based on SSTs, and strategically designate them between the available reinsurance funds so as to maximize their expected profit.

To perform the analysis the study simulated the actions of an adverse selecting private insurance company over the 20-year period of 1978-1997. Under various scenarios it was found that insurance companies would opt to place a majority (from 70% to almost 100%) of the contracts in the commercial fund. The large percent assigned to the commercial fund resulted of two factors: the advantageous risk sharing structure of the SRA, and the minimal production risk associated with years of all but the lowest SSTs. As the coverage level would increase however, so would the tendency to shift liability on to the federal government.

The simulation results indicated that under all scenarios, insurance companies were able to recover excess profits thus reducing their loss ratios by engaging in ENSO based adverse selection activities. This came at the direct expense of the federal government, which saw its losses inflated by 12 to 39 percent. In addition to being economically significant, randomization procedures found these results to be statistically significant as well.

Given that the potential for ENSO-based adverse selection is present in the insurance market for Texas wheat, there lies the question as to whether private insurance companies would take advantage this opportunity. The USDA Office of the Inspector General (1999) recently reported to the Secretary of Agriculture that "Reinsured companies have become very proficient at assigning policies to the various pools to maximize underwriting gains on low-risk policies and to minimize underwriting losses on high-risk policies." (p.12) As the Inspector General's office states private insurance companies are very adept at taking advantage of such activities. Thus ENSO-based adverse selection is an opportunity which they are clearly not likely to forego. The aforementioned report additionally recommends that RMA revise the SRA to assign more risk to private insurance companies. As PIC liability increases, it may become even more advantageous to engage in such activities.

Although the focus of this study was very specific in considering Texas wheat and Nino3 SSTs, the findings are more widely applicable. The vast improvements in seasonal weather forecasting will likely continue for the foreseeable future. This will create similar opportunities throughout the crop insurance industry. This study is just one example of the impact that such forthcoming science can have on the agricultural sector, and it is imperative that government policies are able to evolve with these advances. Given the significant opportunity for ENSO based adverse selection that has been demonstrated by this study, it seems that the RMA should consider this in future negotiations of the SRA.

Unfortunately there are few policy options available to eliminate or alleviate the informational asymmetries that enable PICs to engage in ENSO based adverse selection. The RMA must establish premiums well in advance to accommodate farmers, however reinsurance decisions cannot be made until after contracts have been purchased. Under these circumstances the informational asymmetries are unavoidable. An option that the RMA could take would be to restructure the SRA so that PIC and government profits were equitable after any information asymmetries had been exploited. Essentially the RMA would recognize that adverse selection will take place, and then structure the SRA to reflect this. For instance the asymmetry of the risk sharing structure could be reduced to offset the excess PIC profits gained through adverse selection.

APPENDIX A

ISOTONIC ROBUST SUPER SMOOTHER[15]

The isotonic robust super smoother performs a locally weighted regression smoothing, at each observation (x_t, y_t) in the set of interest. Define an observation x_0; and its k nearest neighbors as the neighborhood $N_k(x_0)$. In addition we will define the span as the number of observations in neighborhood $N_k(x_0)$. The weights are then determined by a decreasing function of the distance between x_0 and all the other observations in neighborhood $N_k(x_0)$. Defining the farthest point from x_0 in neighborhood $N_k(x_0)$ as $d\max(x_0) = \max_{N_k(x_0)}(|x_0 - x|)$, we assign weights according to the function

$$W\left(\frac{|x_0 - x_i|}{d\max(x_0)}\right) \tag{22}$$

where

$$W(u) = \begin{cases} (1 - u^3)^3 & \text{for } 0 \le u \le 1 \\ 0 & \text{otherwise.} \end{cases} \tag{23}$$

Note that for any x_i outside of neighborhood $N_k(x_0)$, $|x_0 - x_i| > d\max(x_0)$ and thus it receives a weight of 0. Using these weights we calculate the least squares regression coefficients and recover the vector of residuals for observation x_0 and neighborhood $N_k(x_0)$.

The optimum span for observation x_0 is then choosing through local cross validation procedures. Locally weighted regression smoothing with the neighborhood chosen by cross validation procedures is termed super smoothing. Cross validation using the leave one out approach chooses the span which minimizes the sum

$$\sum_{i=1}^{n}\left[y_i - \hat{y}_{(i)}^{k}\right]^2 \tag{24}$$

where $\hat{y}_{(i)}^{k}$ is the weighted least squares estimate of y_i, using span k, after excluding observation (y_i, x_i) from the calculation. A constant span over the entire domain may not be desirable. An increase in the curvature of the underlying function would require a smaller span, while an increase in the variance would require a larger one. Thus we employ *local* cross validation and choose a span for each x_0 based only on the neighborhood $N_{\Upsilon}(x_0)$. That is, for each predictor value x_0 the span is chosen on the sum

$$\sum_{N_{\Upsilon}(x_0)}\left[y_i - \hat{y}_{(i)}^{k}\right]^2 \qquad (25).$$

Note that the sum is only over $N_{\Upsilon}(x_0)$ and not the entire sample. This is calculated separately for each realization leading to an individual span (k_i) for each. An overall span, γ, must be specified beforehand to define the neighborhoods $N_{\Upsilon}(X_0)$ to undertake the local cross validation.

Because crop yields in the area of interest are not considered to be Gaussian, the super smoother is augment with robust techniques. The IRSS employs the default S-Plus m-estimator for robust regression, which is the Huber (1979) m-estimator until convergence followed by two iterations of the Bisquare. The robust techniques are used

[15] This outline of the IRSS draws heavily from Ker and Coble (1998).

in both the local cross validation procedure as well as in the final estimates of the coefficients.

Given observation x_0 , k, $N_k(x_0)$, W, and the vector of residuals $\{\varepsilon_1 \ldots \varepsilon_n\}$ recovered from the corresponding estimated coefficients; we calculate the Mean Absolute Deviation (MAD):

$$\frac{\sum|\varepsilon_i|}{k} \quad \text{s.t. } x_i \in N_k(x_0) \tag{26}$$

Thus we calculate the MAD of only those residuals whose corresponding $x_i \in N_k(x_0)$. We then define u_i as the absolute value of residual i divided by the MAD, and recover the Huber weights, Ω, where Ω is defined as

$$\Omega(u) = \begin{cases} 1, & \text{for } u < 1.345 \\ 1.345/u & \text{otherwise.} \end{cases} \tag{27}$$

Using the new weights ΩW, we re-estimate the coefficients and recover a new set of residuals for neighborhood $N_k(x_0)$. Using the new residuals, we re-calculate the Huber weights, and repeat the process until convergence is achieved.

Upon convergence using the Huber weights, the estimated residuals are again used to define the bisquare weights Ψ, where Ψ is defined as

$$\Psi(u) = \begin{cases} (1-(u/4.685)^2)^2, & \text{for } u < 4.685 \\ 0 & \text{otherwise.} \end{cases} \tag{28}$$

We then perform two iterations using weights ΨW, to recover our final estimates of the regression coefficients.

The robust weighted least squares estimates are calculated for all observations and all possible spans using the leave one out approach, to determine the optimum span for each x_0. Once the optimum span has been determined for each observation, the final estimates are calculated with the inclusion of the (x_0, y_0) in question.

Finally, we wish to isotonize, or restrict our estimates to belong to the class of non-decreasing functions. This is done using the pool-adjacent-violators (PAV) algorithm in Hanson, Pledger, and Wright (1973).

To begin, the estimates must be ordered so that they are increasing in x. Starting with the estimate of $\hat{y}_1$, we progress through the series until the monotonicity constraint is violated ($\hat{y}_{i+1} < \hat{y}_i$). If the monotonicity constraint is violated, then pool ($\hat{y}_i, \hat{y}_{i+1}$) and replace the estimates with their average, $\hat{y}_i^* = \hat{y}_{i+1}^* = \dfrac{\hat{y}_i + \hat{y}_{i+1}}{2}$. Having now altered $\hat{y}_i$, we then must check that the preceding estimate $\hat{y}_{i-1} \le \hat{y}_i^*$, and if not then pool ($\hat{y}_{i-1,} \hat{y}_i, \hat{y}_{i+1}$) and average. This process is continued with subsequent preceding estimates until the monotonicity constraint is satisfied, at which point we again progress through the series until completion.

APPENDIX B
NONPARAMETRIC KERNEL DENSITY ESTIMATION[16]

Nonparametric kernel techniques are an effective way of estimating an unknown probability density function without restricting the estimate to a known parametric space. A required input of the kernel density estimator is a set of independent observations from the unknown density of interest. Intuitively, kernel density estimation centers an individual kernel on each observation in the set. The estimate of the unknown probability density function at any point in the domain is then the sum of the individual kernels at that point.

The estimate of the unknown density at any given point (say y_0) is defined as

$$\hat{f}_Y(y_0) = \sum_{i=1}^{T} \frac{K\left(\frac{y_0 - y_i}{h}\right)}{Th} \tag{29}$$

where $K(\cdot)$ is the kernel function, h is a smoothing parameter, and T is the number of realizations and hence the number kernels. A decision must then be made as to the choice of the kernel function $K(\cdot)$, and the choice of the smoothing parameter h.

Epanechnikov (1969) derived the optimum non-negative kernel function that would minimize Mean Integrated Squared Error (MISE). Rosenblatt (1971) however, showed that choice of a suboptimal kernel would result in only a moderate loss in the asymptotic MISE. Therefore a standard Gaussian kernel is generally used in practice, and will be employed for this study as well.

---

[16] This appendix draws heavily from Goodwin and Ker (1998).

In choosing the smoothing parameter, *h*, a decision must be made as to whether one be used globally or locally. If the smoothing parameter is global, then it will smooth all of the observations equally. This may be problematic however, as minimal smoothing to obtain greater detail in the main body of the distribution may, lead to spurious detail in the tails of the distribution where observations are sparse. Thus we use an adaptive kernel method, which allows the smoothing parameter to vary for each observation depending on the density of its surrounding neighborhood. We smooth relatively little in the main body of the distribution where observations are heavily clustered, while smoothing much more at the tails of the distribution where observations are sparse. We therefore run a pilot estimate of the distribution, and weight the smoothing parameter for each observation based on the relative density of its surrounding neighborhood.

Silverman (1986) noted that the adaptive estimate is relatively insensitive to the pilot. Thus we estimate the pilot density with the smoothing parameter chosen by Silverman's rule of thumb:

$$\hat{h} = 0.9 \times \min\left[\text{standard deviation}, \frac{\text{interquartile range}}{1.34}\right] \times T^{-\frac{1}{5}}. \qquad (30).$$

Denoting the pilot estimate $\breve{f}$, we then define the local scale as:

$$\lambda_i = \left(\frac{\breve{f}(y_i)}{g}\right)^{-\alpha} \qquad (31)$$

where $\log(g) = \frac{1}{T}\sum \log \breve{f}(y_i)$ and $\alpha \in [0, 1]$ is the sensitivity parameter. For the theoretical reasons outlined by Abramson (1982) we set $\alpha = 1/2$. The adaptive kernel estimate of f_y at a given point, say y_0 is then defined as

$$f_Y(y_0) = \sum_{i=1}^{T} \frac{K\left(\frac{y_0 - y_i}{\lambda_i h}\right)}{Th\lambda_i} \tag{32}$$

where $h\lambda_i$ is essentially now the smoothing parameter for realization i.

An unfortunate problem with using kernel estimators is that the estimated density does not necessarily have its moments equal to the sample moments. The consistency of these estimators indicate that it is a finite sample problem, but with the limited observations that we are working with, it can be disconcerting. In estimating the density, each kernel and hence each observation is weighted equally. Thus the density mean will obviously equal the sample mean. However by smoothing or in a sense spreading each observation, the density will almost surely have a variance equal to or greater than the sample variance. This is not desirable, as the sample provides an unbiased estimate of the population variance. The variance of the estimate is given by:

$$\frac{h^2 \sum_{i=1}^{T} \chi_i^2}{T} + \frac{T-1}{T} s_E^2 \tag{33}$$

Thus we transform the variance of the estimate by the scalar

$$\sqrt{\frac{s_E^2}{\frac{h^2 \sum_{i=1}^{T} \chi_i^2}{T} + \frac{T-1}{T} s_E^2}} \quad (34).$$

REFERENCES

Abramson, I. S. "On Bandwidth Variation in Kernel Estimates-a Square Root Law." *Annals of Statistics* 10(1982):1217-1223.

Ash, Mark S., and William Lin. *Regional Crop Yield Response for U.S. Grains.* Washington DC: U.S. Department of Agriculture, Commodity Economics Division, Economic Research Service, Agricultural Economic Report 577, September 1987.

Bjerknes, J. "Atmospheric Teleconnections from the Equatorial Pacific." *Monthly Weather Review* 97(1969):163-172.

Buja, Andreas, Trevor Hastie, and Robert Tibshirani. "Linear Smoothers and Additive Models." *The Annals of Statistics* 17(1989):453-555.

Coble, K. H., T. O. Knight, R. D. Pope, and J. R. Williams. "Modeling Farm-Level Crop Insurance Demand with Panel Data." *American Journal of Agricultural Economics* 78(May 1996):439-447.

Epanechnikov, V. A. "Nonparametric Estimation of a Multidimensional Probability Density." *Theory of probability* 14(1969):153-158.

Glantz, Michael H. *Currents of Change.* Cambridge: Cambridge University Press, 1996.

Goodwin, B. K. "Premium Rate Determination in the Federal Crop Insurance Program: What Do Averages Have to Say About Risk?" *Journal of Agricultural and Resource Economics* 19(December 1994):382-395.

Goodwin, Barry K., and A. P. Ker. "Nonparametric Estimation of Crop Yield Distributions: Implications for Rating Group-Risk (GRP) Crop Insurance Contracts." *American Journal of Agricultural Economics* 80(February 1998):139-153.

Goodwin, Barry K., and Vincent H. Smith. *The Economics of Crop Insurance and Disaster Aid.* Washington D.C.: AEI Press, 1995.

Gershunov, Alexander. "ENSO Influence on Intraseasonal Extreme Rainfall and Temperature Frequencies in the Contiguous United States: Implications for Long-Range Predictability." *Journal of Climate* 11(July 1998):3192–3203.

Gershunov, Alexander, and Tim P. Barnett. "ENSO Influence on Intraseasonal Extreme Rainfall and Temperature Frequencies in the Contiguous United States: Observations and Model Results." *Journal of Climate* 11(July 1998):1575–1586.

Handler, Paul. "USA Corn Yields, The El Nino And Agricultural Drought:1867-1988." *International Journal Of Climatology* 10(1990):819-828.

Handler, Paul, and Ellen Handler. "Climate Anomalies in the Tropical Pacific Ocean and Corn Yields in the United States." *Science* 220(June 10, 1983):1155-56.

Hanson, D.L., G. Pledger, and F.T. Wright. "On Consistency in Monotonic Regression." *Annals of Statistics* 1(1973):401-421.

Hoerling, Martin P., and Arun Kumar. "Why do North American Climate Anomalies Differ from One El Nino Event to Another?" *Geophysical Research Letters* 24(9)(1997):1059-62.

Hoerling, Martin P., Arun Kumar, and Min Zhong. "El Niño, La Niña, and the Nonlinearity of Their Teleconnections." *Journal of Climate* 10(August 1997):1769-1786.

Huber, P.J. "Robust Smoothing." *Robustness in Statistics*. E. Launer and G. Wilkinson, eds. New York: Wiley, 1979.

Ker, Alan P.,and Keith H. Coble. "On Choosing a Base Coverage Level for Multiple Peril Crop Insurance Contracts." *Journal of Agricultural and Resource Economics* 23(December 1998):427-444.

Kiladis, G. N., and H. F. Diaz. "Global Climatic Anomolies Associated with the Extremes of the Southern Oscillation." *Journal of Climate* 2(1989):1069-1090.

Kramer, Randall A. "Federal Crop Insurance:1938-1982. "*Agricultural History* 57(1983a):181-200.

Kramer, Randall A. "Federal Crop Insurance:1938-1982. "*Agricultural History* 57(1983b):186. Quoting "Crop Insurance." *The Christian Science Monitor* (23 September 1936):18.

Lagos, P., and J. Butler. *Natural and Technological Disasters: Causes, Effects and Preventive Measures*. S. K. Majumdar, G. S. Forbes, E. W. Miller, and R. F. Schmalz, eds. pp.223-238, Easton: Pennsylvania Academy of Science, 1992.

Luo, H., J. R. Skees, and M. A. Marchant. "Weather Information and the Potential for Inter-temporal Adverse Selection." *Review of Agricultural Economics* 16(1994):441-51.

Mjelde, James W., and Keith Keplinger. "Using the Southern Oscillation to Forecast Texas Winter Wheat and Sorghum Crop Yields." *Journal of Climate* 11(January 1998): 54–60.

Mjelde, J. W., T. N. Thompson, F. M. Hons, J. T. Cothren, and C. G. Coffman. "Using Southern Oscillation Information for Determining Corn and Sorghum Profit-Maximizing Input Levels in East-Central Texas." *Journal of Production Agriculture* 10(1)(1997):168-75.

Montroy, David L. "Linear Relation of Central and Eastern North American Precipitation to Tropical Pacific Sea Surface Temperature Anomalies." *Journal of Climate* 10(April 1997):541-558.

Montroy, David L., Michael B. Richman, and Peter J. Lamb. "Observed Nonlinearities of Monthly Teleconnections between Tropical Pacific Sea Surface Temperature Anomalies and Central and Eastern North American Precipitation." *Journal of Climate* 11(July 1998):1812–1835.

Moss, C. B., and J. S. Shonkwiler. "Estimating Yield Distributions with a Stochastic Trend and Nonnormal Errors." *American Journal of Agricultural Economics* 75(November 1993):1056-72.

Nicholls, Neville. "Impact of the Southern Oscillation on Australian Crops." *Journal of Climatology* 5(1985):553-560.

Nicholls, N. "Use of the Southern Oscillation to Predict Australian Sorghum Yields." *Agricultural and Forest Meteorology* 38(1986):9-15.

Patridge, I J., ed. *Will it Rain? The Effects of the Southern Oscillation and El Nino on Australia*. 2nd ed. Toowoomba, Queensland, Australia: QDPI/CSIRO, 1994.

Philander, S. G. H. *El Nino, La Nina, and the Southern Oscillation*. San Diego: Academic Press, 1990.

Piechota, Thomas C., and John A. Dracup. "Drought and Regional Hydrologic Variation in the United States: Associations with the El Niño-Southern Oscillation." *Water Resources Research* 32(5)(1996):1359-73.

Quinn, W.H., D. O. Zopf, K. S. Short, and R. T. Kuo Yang. "Historical Trends and Statistics of the Southern Oscillation, El Nino, and Indonesian Droughts." *Fish Bulletin* 109(1978):663- 678.

Quiggin, J., G. Karagiannis, and J. Stanton. "Crop Insurance and Crop Production: An Empirical Study of Moral Hazard and Adverse Selection." *Australian Journal of Agricultural Economics* 37(August 1993):95-113.

Rasmusson, E. M., and T. H. Carpenter. "The Relationship between Eastern Equatorial Pacific Sea Surface Temperature and Rainfall over India and Sri Lanka." *Monthly Weather Review* 111(1983):517-528.

Rasmusson, Eugene M., and John M. Wallace. "Meteorological Aspects of El Nino/Southern Oscillation." *Science* 222(December 16, 1983):1195-1202.

Reynolds, R. W., and T. M. Smith. "A High Resolution Global Sea Surface Temperature Climatology." *Journal of Climate* 8(1995)1571-83.

______. "Improved Global Sea Surface Temperature Analysis Using Optimum Interpolation." *Journal of Climate* 7(1994)929-948.

Rimmington, Glyn M., and Neville Nicholls. "Forcasting Wheat Yield in Australia with the Southern Oscillation Index." *Australian Journal of Agricultural Research* 44(1993):625-632.

Ropelewski, C. F., and M. S. Halpert. "North American Temperature and Precipitation Patterns Associated with the El Nino/Southern Oscillation." *Monthly Weather Review* 114(1986):2352-62.

______. "Precipitation Patterns Associated with the High Index Phase of the Southern Oscillation." *Journal of Climate* 2(1989)268-284.

Rosenblatt, M. "Curve Estimation." *Annals of Mathematical Statistics* 42(1971):1815-42.

Rosenman, Samuel I., ed. *The Public Papers and Addresses of Franklin D. Roosevelt*, vol.9, *War and Aid to Democracies*. New York: Macmillan, 1940. Quoted in Kramer, Randall A. "Federal Crop Insurance:1938-1982. "*Agricultural History* 57(1983):190.

Silverman, B. W. *Density Estimation for Statistics and Data Analysis*. New York: Chapman and Hall, 1986.

Smith, T. M., R. W. Reynolds, R. E. Livezey, and D. C. Stokes. "Reconstruction of Historical Sea Surface Temperatures Using Empirical Orthogonal Functions." *Journal of Climate* 9(1996)1403-20.

Soule, P. T. "Spatial Pattern of Drought Frequency and Duration in the Contiguous USA Based on Multiple Drought Event Definitions." *International Journal of Climatology* 12(1992):11-24.

Trenberth, K. E. "El Nino Southern Oscillation." *Climate Change: Developing Southern Hemisphere Perspectives*. T. Giambelluca and A. Anderson-Sellers, eds., pp. 145-173. Chichester: John Wiley and Sons Ltd., 1996.

______. "Short Term Climate Variations: Recent Accomplishments and Issues for Future Progress." *Bulletin of the American Meteorology Society* 78(June 1997a):1081-96.

______. "The Definition of El Nino." *Bulletin of the American Meteorological Society* 78(December 1997b):2771-2777.

Trenberth, Kevin E., and Grant W. Branstrator. "Issues in Establishing Causes of the 1988 Drought over North America." *Journal of Climate* 5(1992):159-172.

United States Department of Agriculture. "Radio Address of Agriculture Secretary Dan Glickman [Online]." Available HTTP: http://www.usda.gov/news/releases/1998/12/0524 (December 1998).

United States Department of Agriculture Office of Inspector General. *Report To The Secretary On Federal Crop Insurance Reform.* Washington D.C., March 1999.

United States General Accounting Office. *Crop Revenue Insurance: Problems with New Plans Need to Be Addressed.* Washington D.C., April 1998.

Walker, G. T. "Correlation in Seasonal Variations of Weather IX: A Further Study of World Weather." *Memoirs of the Indian Meteorological Department* 24(1924):275-332.

White, Charles. Personal Communication. United States Department of Agriculture Risk Management Agency (May 1998).

White House. "President William Jefferson Clinton State of the Union Address [Online]." Available HTTP: http://www.pub.whitehouse.gov/uri-res/I2R?urn:pdi://oma.eop.gov.us/1999/1/20/1.text.1 (January 1999).