



**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](mailto: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.*

# Prediction of Loan Deficiency Payments

Joseph Cooper\*

Economic Research Service – USDA  
1800 M Street NW, S-4187  
Washington, DC 20036-5831  
Email: [Jcooper@ers.usda.gov](mailto:Jcooper@ers.usda.gov)  
Phone: 202-694-5482

and

Gerald Plato

Economic Research Service – USDA  
1800 M Street NW, S-4187  
Washington, DC 20036-5831  
Email: [gplato@ers.usda.gov](mailto:gplato@ers.usda.gov)  
Phone: 202-694-5604

May 24, 2007

Accepted Paper for the 2007 Annual Meetings  
of the American Agricultural Economics Association  
Portland, Oregon (July 29 - August 1)

\*The views expressed are the authors' and do not necessarily represent those of the Economic Research Service or the US Department of Agriculture.

# **Prediction of Loan Deficiency Payments**

## **Abstract**

This paper develops a stochastic model for estimating potential loan deficiency payments to U.S. corn producers in a discrete-dynamic context. We minimize the potential for misspecification bias by using nonparametric and semi-nonparametric approaches as specification checks in the model. The model permits the forecast at planting time of the resulting empirical distribution of LDP payments for that crop year. Using this model, the paper examines the sensitivity of this distribution to changes in expected price levels.

## **Key words**

Domestic support, loan deficiency payments, marketing loan benefits, corn, yield

## **Prediction of Loan Deficiency Payments**

### **Introduction**

Direct commodity support is provided in several forms to U.S. producers of the major bulk crops. Of these, marketing assistance loans are tied to current production. Producers are eligible for marketing assistance loans when they harvest the eligible commodities. If market prices are below the loan rate, farmers are allowed to repay commodity loans at a loan repayment rate (reflecting market prices) that is lower than the loan rate (except for extra-long staple cotton). The resulting “gain” is defined as the domestic support payment. To participate, farmers decide how much of their current year’s production they want to place under loan and pledge that amount as collateral. Marketing assistance loans have a 9-month maturity and accrue interest, but the interest does not need to be repaid if the loan repayment rate is less than the loan rate. These loans are “non-recourse loans,” meaning that the government must accept the collateral as full payment of the loan at loan maturity if a producer so chooses. Alternatively, the producer can choose not to take out the marketing loan and can instead take the benefit as a cash payment in the form of a loan deficiency payment (LDP) if the repayment rate is less than the loan rate. The LDP tends to be the primary vehicle by which marketing loan benefits are delivered to producers. The producer can take the LDP and then be free to sell the crop on the open market.

Loan deficiency payments are highly variable from year to year and can represent a substantial cost to the U.S. Federal government. For instance, LDP payments for corn for the 2005 crop year amounted to \$4.3 billion dollars. LDP payments to a producer

are made on the basis of a daily or weekly posted county price (PCP) at the time that the farmer takes the LDP payment or repays the marketing loan. Among the drawbacks from the government's standpoint of this frequent updating of prices is that an unusual short-term event may cause a short-term decline in market prices, triggering a large volume of LDP requests at a high LDP rate that may not reflect the longer-term or underlying market conditions (USDA, 2007). The frequent updating of PCPs may induce producers to seek out the bottom of the market when taking their LDPs. Hence, the flexibility in the current loan structure may lead to excessive costs of the marketing loan program. To reduce the sensitivity of LDPs and marketing assistance loans to short-term price fluctuations, USDA (2007) proposes that daily calculated PCPs be replaced with monthly calculated PCPs.

We take the USDA proposal as an opportunity to develop a stochastic model for estimating potential loan deficiency payments to U.S. corn producers in a discrete-dynamic context. We minimize the potential for misspecification bias by using nonparametric and semi-nonparametric approaches as specification checks in the model of LDPs based on a monthly price structure. With this model, we forecast at planting time the empirical distribution of LDP payments for that crop year and examine the sensitivity of this distribution to expected price levels. Such an approach could aid the government's budgetary forecast process.

## Background

The LDP is based on shortfalls in market price with respect to a statutory loan rate.

Formally, for producer  $i$  of a qualifying crop in time  $t$ , the marketing loan benefit, or equivalently, the loan deficiency payment, is calculated as:

$$(1) \text{LDP}_{it} = \max \{ 0, (LLR_{it} - ALR_{it}) \} \cdot Q_{it},$$

where the statutorily-set local loan rates, or  $LLR$ , is the national loan rate  $LR$  adjusted by various county-specific and quality factors. The alternative loan repayment rate, or  $ALR$ , is essentially a USDA-determined posted county price (PCP) that varies daily or weekly (depending on the crop) according to market conditions, and is adjusted to reflect product quality. Depending on the crop, the  $ALR$  may be a county (wheat, feed grains, oilseeds), national (peanuts), or world (upland cotton and rice) “posted” price. The quantity  $Q_{it}$  (measured in bushels for our corn application) is the output that the farmer takes the LDP on. Alternatively to the LDP, producers can potentially receive support benefits from nonrecourse marketing assistance loans. These marketing loan gains (MLGs) occur if a marketing loan is repaid at less than the loan principal, and follow the same formula as in equation (1). We do not distinguish between the marketing loan benefits (MLB) taken by the farmer as an LDP or an MLG as the marketing loan benefit is the same in either case. Effectively, our analysis is of marketing loan benefits, but we will use the term LDP given that most of the gains come in that form and the latter is perhaps a more commonly known term.

Producers may obtain loans or receive LDPs on all or part of their eligible production anytime during the loan availability period. For corn, this period would run from around October or November (when the crop is normally harvested) to May 31 of the next year. For output to be eligible for LDPs, the farmer must have a “beneficial

interest” in the output (FSA, 2003). Among other things, the farmer must have control over the output at the time the LDP is applied for, e.g., the farmer cannot have already sold the output.

As the lowest corn prices tend to occur around harvest time or soon after, that also tend to be the time of year when most LDPs are taken. For example, Farm Service Administration records show that for Hardin County, Iowa, approximately 60 percent of total electronic LDP applications for the 2005 crop year were taken by mid-November and 90 percent by the end of the first week of December.

To address the question of when producers would most likely take LDP payments under a monthly PCP system, we look to historic price data.<sup>1</sup> We use the CBOT (Chicago Board of Trade) December futures contract for corn given the availability of this data back to 1969, and given its high correlation with cash price data.<sup>2</sup> Figure 1 shows that historically over the months covering the LDP availability period, December most frequently represents the lowest monthly price. Hence, unless the producer’s cost of storing the commodity into December outweighs the expected price differential in LDPs from waiting until the end of December to take the LDP, it seems likely that most producers taking an LDP based on a monthly PCP will likely do so at the end of December.

### **Methodology for estimating payments**

We estimate the distribution of payments based on the historic relationship between national price and national average yield. Payments in crop year  $t$  are assumed to be a function of planted acres at the beginning of  $t$ , the loan rate, and the stochastic price-yield

relationship in which within-season price change is a function of within-season yield change.

### *Modeling the within-season price-yield relationship*

Our focus is on estimating the distribution of payments for a given reference crop year, in our case 2005, given the difference in realized yield and price over the levels expected at the beginning of the crop year. Hence, for the purposes of estimating the relationship between price and yield, we re-express the historic price and yield data as proportional changes between expected and realized price and expected and realized yield within each period, respectively. We can then apply this history of proportional changes in yield and price to 2005 data to develop the distribution of payments. Specifically, the realized national average yield,  $Y_t$ , is transformed to  $\Delta Y_t = (Y_t - E(Y_t)) / E(Y_t)$ . The expected value of  $Y_t$ , or  $E(Y_t)$ , is calculated from an estimated trend equation (as described in detail below). Similarly, the realized price at harvest,  $P_t$ , is transformed to  $\Delta P_t = (P_t - E(P_t)) / E(P_t)$ , where  $E(P_t)$  is derived from futures prices as discussed in the Data section below.

The historic yield data needs to be detrended before it can be used for our analysis. Namely, the upward trend in corn yields since the mid-1940s has been quite remarkable, and even mean corn yields from the 1970s are significantly lower than that which would be expected today. To generate a distribution for  $Y_{2005}$  based on historic yield shocks, the historic yields must be rescaled to reflect the proportional change in the



state of technology between that in 2005 and that in time  $t$ , i.e.,  $Y_t$  is rescaled to 2005 terms as

$$(2) Y_t^d = E(Y_{2005})(\Delta Y_t + 1), \forall t \text{ periods, } t \neq 2005.$$

A specific detrending approach used in the literature is to assume that expected yield evolves according to a time trend, or  $E(Y_t) = f(t)$ , e.g., Chavas and Holt (1990), who fit a linear model to the time trend.

To separate the stochastic component of yield from the upward trend in yields over time due to technological and managerial innovations, we detrend the yield data using a simple linear model as generally used in the literature. In addition, as a specification-check, we also detrend yields using a nonparametric LOESS (Cleveland, 1979; Cleveland and Devlin, 1988) prediction of yield trends instead of the simple parametric approaches used in the literature. As the LOESS (LOcally weighted polynomial regrESSion) procedure is available as a canned procedure in several common statistical and econometric packages, it is not described in detail here. In brief, at each point in the data set a low-degree polynomial is fit to a subset of the data. The polynomial is fit using weighted least squares, giving more weight to points near the point whose response is being estimated and less weight to points further away. A smoothing parameter denotes the degree of the local polynomial, and controls the flexibility of the model.

Our goal in fitting the trend regression was to model yield as a function of technological change and other factors that are correlated with time. Any deviations from the trend are assumed to be due to stochastic shocks. A model that is linear with respect to the time trend, e.g.,  $E(Y_t) = \alpha + \beta t + \varepsilon_t$ , may in certain cases be too restrictive in its

assumption with regards to technical change. However, note that even if the equality of a restricted model to a fully flexible model is not accepted from a statistical standpoint, this result does not imply that the fully flexible model should be used for detrending the data – doing so will be at odds with the goal of separating the yield shocks from trend effects. The LOESS yield trend provides some flexibility to  $f(t)$  over the linear model while minimizing the chasing of the stochastic yield shocks.

Figure 2a plots the yield per planted acre trend for corn for grain and silage over 1946 to 2006 and 2b over 1975 to the present. As the figures show, the nonparametric trend is roughly linear, whether over 1946 to the present or over the more narrow span of 1975 to 2005 that we use later in the analysis. Figure 2a shows the pre-1940s yields to demonstrate that most yield growth over the last century dates from the post Second World War period, with an annual postwar growth rate of 2.81 percent. The dashed lines in Figure 2a represent hypothetical 99 percent confidence intervals assuming a constant coefficient of variation over 1946 to 2005, where the estimated coefficient of variation is calculated from detrended yield over 1996 to 2005. These confidence bands suggest that the increase in yield and in its standard deviation have been roughly proportional over time, thereby providing some justification for using historical corn yields as a guide to future corn yields. Figure 3 presents the deviations in actual corn yields from expected yields over the 1950 to 2005 period. The variability in the data in the figure suggest that one will need more years of observed data to determine whether or not corn varieties are becoming significantly less sensitive to weather shocks over time.

Given the estimated trend yields as the predictions of  $E(Y_t)$ , we can construct  $\Delta Y_t^d$  and estimate the relationship between it and  $\Delta P_t$ . In particular, we assume that

$\Delta P_t$  can only be partially explained by  $\Delta Y_t^d$ , and that the uncertainty in this relationship can be incorporated into the empirical distribution. We do so by specifying  $\Delta P_t$  as

$$(3) \Delta P_t = g(\Delta Y_t^d, z_t) + \varepsilon_t,$$

where  $\varepsilon_t$  is *i.i.d.* with mean 0 and variance  $\sigma_\varepsilon^2$  given  $\{\Delta Y_t^d, z_t\}$ .<sup>3</sup> We expect that

$d\Delta P_t / d\Delta Y_t^d < 0$ , that is, the greater the realization of national average yield over expected national average yield, the more likely harvest time price will be lower than expected price.

Based on the econometric estimate of the function for  $\Delta P_t$ , we can then generate a distribution of estimates of  $\Delta P_t$ , or  $\Delta \hat{P}_{gt}(g[\Delta Y_t^d, z_t])$ , for each  $\Delta Y_t^d$ . As will be explained in the next section, to reduce the potential for bias due to the misspecification of equation (3), we utilize a semi-nonparametric (SNP), or flexible, econometric approach as a specification check on a parametric estimate of  $g(\cdot)$ .

#### *Semi-nonparametric estimation of the price-yield relationship*

A potential drawback of a parametric model for the estimation of equation (3) is that it could potentially be subject to biases associated with incorrect specifications of functional form of  $g(\cdot)$ . As a specification check on a parametric model, we use the Fourier flexible functional form (e.g., Fenton and Gallant, 1996) to model equation (3). The Fourier functional form is one of the few functional forms known to have Sobolev flexibility, which means that the difference between a function  $g(\mathbf{x}, \theta)$  and the true function  $f(\mathbf{x})$  can be made arbitrarily small for any value of  $\mathbf{x}$  as the sample size becomes large (Gallant, 1987). Letting  $\mathbf{x}_t$  represent the

vector of explanatory variables in equation (3) with 3 or more unique values each, our SNP specification of  $g(x, \theta)$  is:

$$(4) Y_i = g(x_i, \theta_i) = x_i' \beta + \sum_{m=1}^M \sum_{l=1}^L (v_{lmj} \cos[r_{mj}' s(x_i)] - w_{lmj} \sin[r_{mj}' s(x_i)]),$$

where the  $p \times 1$  vector  $x_i$  contains all arguments of the utility difference model,  $k$  is the number of coefficients in  $\theta$ , which consists of the  $\beta$ ,  $v_{lm}$ , and  $w_{lm}$  coefficients to be estimated,  $M$  and  $L$  are positive integers, and  $r_m$  is a  $p \times 1$  vector of positive and negative integers that forms indices in the conditioning variables and that determines which combinations of variables in  $x_i$  from each of the transformed variables.<sup>4</sup> The integer  $m$  is the sum of the absolute value of the elements in the multi-indexes in vector  $r_m$  and  $L$  is the order of the transformation, and is basically the number of inner-loop transformations of  $x_i$ . For example, if  $x_i$  contains 3 variables and  $M = L = 1$ , then the  $r_m$  vectors are (1,0,0), (0,1,0), and (0,0,1), resulting in  $k = 9$  (not counting the constant). The  $p \times 1$  function  $s(x_i)$  prevents periodicity in the model by rescaling  $x_i$  so that it falls in the range  $[0, 2\pi - 0.000001]$  (Gallant, 1987). This rescaling of each element in  $x_i$  is achieved by subtracting from each element in  $x_i$  its minimum value (from across the sample), then dividing this difference by the maximum value (from across the sample), and then multiplying the resulting value by  $[2\pi - 0.000001]$ . For example, if bid is the only explanatory variable, then  $r_m$  is a (1x1) unit vector and  $\max(M)$  equals 1. If a variable has only three unique values, then only the v or w transformation may be performed. A dummy variable is not transformed. In practice, the level of transformation embodied in  $M = L = 1$  generally adds sufficient flexibility to the model, and the parametric model is nested in the SNP model.

A formal criterion for choosing  $M$  and  $L$  is not well established. Chalfant and Gallant (1985) suggest a rule of thumb that the dimension of  $\theta = N^{2/3}$ . Asymptotic theory calls for  $\theta = N^{1/4}$  (Andrews, 1991), but Fenton and Gallant note that  $\theta = N^{1/2}$  is likely to be more representative of actual practice. Hong and Pagan (1988) found that the Fourier approximation had low bias in the estimators even for sample sizes as low as  $n = 30$ .

Taken individually, Fourier coefficients do not have an economic interpretation. To give those regression coefficients an economic interpretation, they must be re-expressed in terms of the base variables. One way to do this is to evaluate  $\partial g(\mathbf{x}, \theta) / \partial \mathbf{x}$ , noting that

$$(5) \quad \partial g(\mathbf{x}, \theta) / \partial \mathbf{x} = \mathbf{b} + 2 \sum_{\alpha=1}^A \left\{ \sum_{j=1}^J j (v_{j\alpha} \cos[j \mathbf{r}'_{\alpha} s(\mathbf{x})] + w_{j\alpha} \sin[j \mathbf{r}'_{\alpha} s(\mathbf{x})]) \mathbf{r}_{\alpha} \right\}.$$

### *Generating the empirical distribution of payments*

To generalize our empirical distribution of payments, we use a general bootstrap method that can allow for flexible right-hand-side regression modeling and allow for modeling interactions between variables. In particular, we use a bootstrap approach in a joint resampling methodology that involves drawing *i.i.d.* observations with replacement from the original data set (Efron, 1979; Yatchew, 1998). The bootstrap data-generating mechanism is to create replications by treating the existing data set of size  $T$  as a population from which samples of size  $T$  are extracted. Equation (3) is re-estimated for each of these bootstrapped data sets. Variation in estimates results from the fact that upon selection, each data point is replaced within the population. We use this standard bootstrap to generate a distribution of  $\Delta P_t$  given  $\Delta Y_t^d$ .

## Data

Data on planted yields and acres for corn are supplied by the National Agricultural Statistics Service (NASS) of the U.S. Department of Agriculture. As payments can be collected for corn for silage as well as corn for grain, and because silage can be a significant portion of corn production in some regions outside the Heartland, we merge together data on corn for grain and corn for silage. We convert tons of silage to bushels using a conversion rate of 7.94 bushels per ton, as per FSA (2006).

Since we use national level yield figures for our analysis, we simplify Equation (1) by not adjusting the national loan rate for county-specific and quality factors. In addition, since we use the national loan rate, we also use the actual national market price as the alternative loan repayment rate rather than the institutionally-determined “posted county prices.”

In particular, for  $P_t$ , we use the average of the daily December prices of the December CBOT corn future in period  $t$ . Note that perhaps a better consensus for a harvest time price would perhaps be the average of the November prices, but we are interested in modeling the prediction of the price when LDPs are most likely to be taken, and not the harvest time price.<sup>5</sup> For the expected value of price  $P_t$ , or  $E(P_t)$ , we utilize a non-naive expectation, namely the average of the daily February prices of the December Chicago Board of Trade corn future (CBOT abbreviation CZ) in period  $t$ ,  $t = 1975, \dots, 2005$ . While we have prices back to 1969, the data before the mid-1970s does not reflect China and Russia as regular participants in global grain markets, and is unlikely to be representative of contemporary global markets. The immediate post-harvest time price  $P_t$  is the average of the daily December prices of the December CBOT

corn future in period  $t$ . The choice of expected corn price is consistent with that of the USDA's Risk Management Agency in pricing certain crop revenue insurance products for corn. To best match the price deviations, the yield deviations are estimated using the nonparametric specification of expected yield estimated over the same time period of 1975 to 2005. The data in Figure 4 suggests that this time span strikes a balance between covering a representative range of yield shocks while at the same time modeling the price-yield relationship over a span more likely to represent the future.

### **Analytical results**

Table 1 provides the econometric results for the parametric and the SNP models over 1975 to 2005. The dummy variable *FarmAct* takes the value of "1" for years 1996 and above (and 0 otherwise), reflecting the Federal government being out of the commodity storage business under recent Farm Acts. We would expect the market to be more efficient in predicting harvest time price without the government build-up of stocks, suggesting a negative sign for *FarmAct*. Regression results show its coefficient to be negative and significant at the 10 percent level in the parametric regression and 12 percent in the SNP regression. Of course, being a dummy variable, *FarmAct* is treated as fully parametric in the SNP regression.

In addition to the  $t$ -statistics for the actual data, the table presents confidence intervals for the regression results that were produced with the bootstrap approach using 1,000 simulated data sets. The confidence intervals presented in Table 1 are constructed from the regression results on each simulated data set and are of the bias corrected accelerated (BCa) type (Efron, 1987), which gives the bootstrap results an interpretation

analogous to  $t$ -statistics by making the estimated confidence interval symmetric around the mean. Using this statistic, the coefficient on *FarmAct* on the SNP regression is significantly different from 0 at the 10 percent level.

The coefficient on  $\Delta Y_t$  is significant at the 1 percent level in both regressions. The higher order transformation terms in the SNP regression are not significant, and the value for  $d\Delta P_t/d\Delta Y_t$  is nearly the same for both regressions.<sup>6</sup> In fact, a likelihood ratio test cannot reject the hypothesis of the equivalence of the parametric and SNP results.

While the expectation is that the within season yield change would be the most significant explanation of within season change in corn prices – and the  $R^2$  values show that over 50 percent of the variation in the within season change in corn prices is explained by the within season change in corn yields – other potential nonendogenous explanatory variables were examined as well. These included  $\Delta Y_t$  for corn produced in the rest of the world, which was not significant in explaining  $\Delta P_t$  for U.S. corn. The  $\Delta Y_t$  for soybeans was not included in the regression as the correlation coefficient between  $\Delta Y_t$  for soybeans and  $\Delta Y_t$  for corn is quite high, as one would expect, at 0.74. A proxy for the  $\Delta Y_t$  for corn in the regression is the  $\Delta Y_t$  for all U.S. feed grains. However, substituting this latter value for the former produced almost identical regression results, suggesting that corn is the driving force in producing within season changes in corn prices.<sup>7</sup> Also examined was the change in the corn trade-weighted U.S. exchange rate (Economic Research Service, 2007) between the beginning of the crop year and the end of the crop year, and the change in GDP between the first quarter and the fourth quarter of the calendar year. Neither variable was significant in the regression.



Another factor that could explain within season change in corn prices is the stocks-to-use ratio, but variables such as this were not included in the analysis as our goal was to model corn price change in a reduced form purely as a function of exogenous shocks, and in particular, yield uncertainty. We treat all other variables as exogenous shocks via the error term in Equation (3).

To generate the forecast of the probability distribution of LDP payments, we use the regression results of the bootstrap analysis discussed above to generate the distribution of price shocks associated with each yield shock. Specifically, the simulated  $(1 \times G)$  vector  $\Delta \hat{\mathbf{P}}_t = \{\Delta \hat{P}_{1t}, \Delta \hat{P}_{2t}, \dots, \Delta \hat{P}_{Gt}\}$  corresponding to each deviation in planted yield  $\Delta Y_t^d$  is generated from coefficients sets corresponding to the  $G = 1,000$  bootstrapped data sets, with *FarmAct* set equal to 1 in the equation to adjust the predictions of  $\Delta P_t$  to reflect the post-1996 Farm Act regime:<sup>8</sup>

$$(6) \Delta \hat{P}_{gt} = \hat{\beta}_{1g} + \hat{\beta}_{2g} \Delta Y_t^{d*} + \hat{\delta}_g \text{FarmAct}_t \quad t = 1, \dots, T \text{ and } g = 1, \dots, G$$

To increase the smoothness in the empirical distribution of the forecast,  $\Delta Y^d$  is divided into  $T = 1,000$  increments in the ascending sequence  $\Delta Y_t^{d*} = \{\min(\Delta Y_t^d) \dots \max(\Delta Y_t^d)\}$ . The end result is therefore a  $1,000,000 \times 1$  vector of estimated price shocks.

For our example, we forecast the 2005 LDP distribution at the beginning of the 2005 crop year. The estimated price deviation  $\Delta \hat{P}_{gt}$  used to derive the harvest time price in 2005,  $\hat{P}_{g2005}$ , is determined as  $\hat{P}_{g2005} = E(P_{2005}) \cdot (\Delta \hat{P}_{gt} + 1)$ . Since LDP payments are based on cash prices, and since the deviation in cash and futures prices are quite close, we set  $E(P_{2005})$  to the February cash price for that year (\$1.86/bu., Illinois No. 2 yellow Corn), and thus, remove the need to add the basis to the price prediction to convert from

futures to cash prices. We assume that the producers take out the LDP on total farm production of corn for the crop year, where  $\hat{Q}_{2005} = A_{2005} \cdot Y_t^{d*}$ , where  $A_{2005}$  is the planted corn acreage in 2005.

### **Discussion of Results**

Given the February 2005 cash price of \$1.86 per bushel, planted acres of 81.76 million, and the 2005 Farm Act loan rate of \$1.95 per bushel for the 2005 crop year, the mean and standard error of the forecasted LDPs for December 2005 are \$3.01 and \$2.78 billion, respectively. The 90 percent confidence interval is \$0.00 to \$8.00 billion, and the coefficient of variation = 0.924. This band may seem wide, but actual LDPs do vary greatly from year to year, and include \$0 payments some years. Note that the actual total corn LDP payment of \$4.3 billion for the 2005 crop year is within this 90 percent confidence interval.

Figure 4 summarizes the bootstrap results at each yield deviation increment using the 2005 expected price and planted acreage. Read vertically, it gives the mean and 90 percent confidence interval in LDPs associated with each yield shock relative to the 2005 expected yield. The vertical dotted line represents the actual 2005 baseline yield level (i.e., yield deviation is zero). As the figure suggests, the estimated price in February 2005 was low enough that it would have taken substantial yield shortfalls to reduce LDP payment to zero that year.

Our approach can be used to predict the distribution of LDPs as a function of other market conditions at planting time. Figure 5 presents the mean LDP payments for a range of expected prices listed in the figure, where again, expected price is taken as that

in February.<sup>9</sup> As expected, payments fall substantially as the expected price increases. We do not examine payments under expected prices lower than the \$1.86/bu value of February 2005, as in real terms, prices have not been much lower than that in recent years.

Mean LDP payments effectively fall to zero when expected prices reach \$2.90, although at that price and at the maximum observed percent yield increase, the 90 percent upper bound is \$0.54 billion (not shown in figure). With the biofuel-induced February 2007 price of \$3.90, the forecasted 90 percent upper bound for LDP payments in December 2007 is \$0 for any feasible yield increase.

### **Future Research Directions**

In the options literature, an American put option is one that can be exercised before its expiration. For example, the LDP based on daily posted county prices is equivalent to an American put with payment based on the end of May price (the end of the LDP availability period). However, with an American option, payment can be taken early by exercising the option, i.e., taking payment on the current price. Option pricing theory implies that this type of option will be exercised and the underlying asset (corn in this paper) sold when the option time value falls to zero (Stoll and Whaley, 1993, pp. 187-190). Exercising an option in our context here is taking the LDP payment. Gaming with this type of option refers to waiting for the option's time value to fall to zero. Time value sometimes falls to zero when there is a short-term aberration in the county posted price.

The USDA (2007) proposed LDP payment based on the average posted county price for a month is equivalent to an average price European option. The option expires at

the end of month for which the average price is calculated. There is no gaming with this option because it cannot be exercised if posted county price takes an unexpected dip. Future research could examine the producer's timing of taking LDP using Stoll and Whaley's American option pricing model (Stoll and Whaley, 1993). This approach could aid in understanding producer behavior by determining if LDP payment timing decisions are influenced by option time value.

## Endnotes

---

<sup>1</sup> USDA (2007) recommends that the monthly PCP be an average of five daily PCPs on pre-set days during the previous month, excluding the high and low daily PCP. For the sake of generality in our application, we assume that the monthly average PCP would be calculated as a simple average over the whole calendar month.

<sup>2</sup> The level of correlation between cash and futures prices for corn depends to some extent on the time span examined, but Plato and Linwood (2007) find it to be at least 0.95.

<sup>3</sup> For the purposes of the regression itself, either  $\Delta Y_t^d$  or  $\Delta Y_t$  can be used as the difference between the two is simply that the former is expressed with respect to a particular base year.

<sup>4</sup> In addition to appending  $x\beta$  to the Fourier series in Equation (4), Gallant suggests appending quadratic terms when modeling nonperiodic functions. Our experiments suggest that inclusion of the quadratic terms as well in the regressions had little impact on the slope estimates. Hence, we leave them out for the sake of efficiency.

<sup>5</sup> These same products use the November price of the December contract as the harvest time price for corn. We depart a bit from RMA practice by calculating average price over a whole month instead of a portion of the month.

<sup>6</sup> These results also hold for the analysis of the data over 1969 to 2005, which is available upon request.

<sup>7</sup> The correlation coefficient between  $\Delta Y_t$  for U.S. feed grains and  $\Delta Y_t$  for U.S. corn is 0.96.

---

<sup>8</sup> Note that as the *FarmAct* coefficient is not strongly significant in the SNP regression, it would not be expected to have much of an impact on the predicted price deviations.

Nonetheless, it still needs to be included given that *FarmAct* was one of the regressors included in the regression upon which the prediction is based.

<sup>9</sup> Farmers would be expected to change their planted acres in response to changes in expected price. For the purpose of this simulation, we use a simple a simple single crop acreage response model that assumes a Cobb-Douglas functional form, i.e.,  $Acres_t = \alpha \cdot E(P_t)^\beta$ . We assume a corn acreage supply elasticity of 0.303, a value which is used for certain USDA market analysis procedures (e.g., the Partial Equilibrium Agricultural Trade Simulator [PeatSim] model). We calibrate the model for 2005 expected price and planted acreage data to derive  $\alpha$ . In calibrating the acreage response model, as the loan rate is effectively the farmer's price floor (e.g., Chavas and Holt, 1990), the farmer's expected price can take one of two

values,  $E(P_t) = \begin{cases} P_t^{Feb} & \text{if } P_t^{Feb} \geq LR \\ LR & \text{if } LR > P_t^{Feb} \end{cases}$ . Given this calibration of the acreage response

model, we can estimate the planted acreage associated with each alternative expected price (*ceteris paribus*) scenario, and forecast the LDP payments.

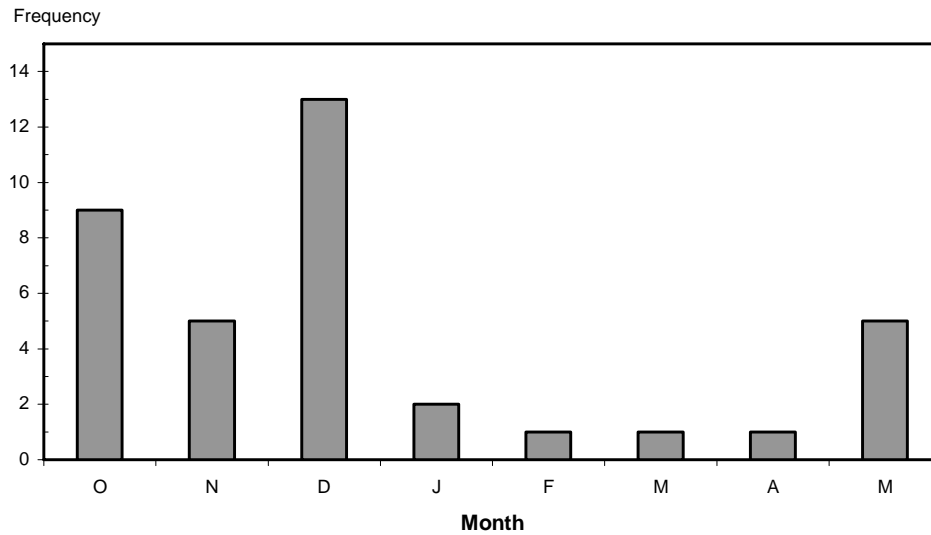
## References

- Andrews, D. “Asymptotic Normality of Series Estimators for Nonparametric and Semiparametric Regression Model.” *Econometrica* 59(March 1991): 307-345.
- Chalfant, J. and A. Gallant, “Estimating Substitution Elasticities with the Fourier Cost Function: Some Monte Carlo Results,” *J. Econometrics* 28(May 1985):205-222.
- Chavas, J. and M. Holt. “Acreage Decisions Under Risk: the Case of Corn and Soybeans,” *American Journal of Agricultural Economics* 72 (1990): 529–538.
- Cleveland, W and S. Devlin, “Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting”, *J. Amer. Statist. Assoc.*, 83(Sep., 1988):596-610.
- Cleveland, W. “Robust Locally Weighted Regression and Smoothing Scatterplots,” *J. Amer. Statist. Assoc.*, Vol. 74 (Dec., 1979), pp. 829-836.
- Economic Research Service. “Agricultural Exchange Rate Data Set,” U.S. Department of Agriculture, Washington, DC, March 2007 update.
- Efron, B. “Better Bootstrap Confidence Intervals.” *J. Amer. Statist. Assoc.* 82(1987):171-185.
- Efron, B. “Bootstrap Methods: Another Look at the Jackknife,” *Annals Statist.* 7(1979: 1–26.
- Farm Services Agency, 2006. “Direct and Counter-Cyclical Program 1-DCP Amendment 38”, U.S. Department of Agriculture, August 4, 2006.
- Farm Services Agency 2003. “Nonrecourse Marketing Assistance Loan and Loan Deficiency Payment Program,” Fact Sheet, U.S. Department of Agriculture, Washington, DC, June 2003.

- Fenton, V., and A. Gallant. "Qualitative and Asymptotic Performance of SNP Density Estimators." *J. Econometrics* 74(1996):77-118.
- Fenton, V. and Gallant, A. "Qualitative and Asymptotic Performance of SNP Density Estimators," *J. Econometrics* 74 (1996):77-118.
- Gallant, A. "Identification and Consistency in Semi-Nonparametric Regression." *Advances in Econometrics*, T. F. Bewley, ed., pp. 145-169, New York: Cambridge University Press, 1987.
- Gallant, A. "Unbiased Determination of Production Technologies. *J. Econometrics* 20(1982):285-323.
- Greene, W. *Econometric Analysis*, 4<sup>th</sup> Edition, Prentice Hall, 2004.
- Hong, Y and Pagan, Adrian, 1988. "Some Simulation Studies of Nonparametric Estimators," *Empirical Economics*, Springer, vol. 13(3/4), pages 251-66.
- Plato, G. and L. Hoffman. "Measuring the Influence of Commodity Fund Trading on Soybean Price Discovery," paper presented at the NCCC-134 Meeting on Applied Commodity Price Analysis, Forecasting, and Market Risk Management, Chicago, Illinois, April 16-17, 2007
- Stoll, H. and R. Whaley. *Futures and Options: Theory and Applications*, South Western Publishing Co., Cincinnati, Ohio, 1993.
- USDA. "USDA 2007 Farm Bill Proposals," USDA, Washington DC, February, 2007.
- Yatchew, A. "Nonparametric Regression Techniques in Economics," *J. Economic Lit.*, Vol. XXXVI (June 1998): 669-721.

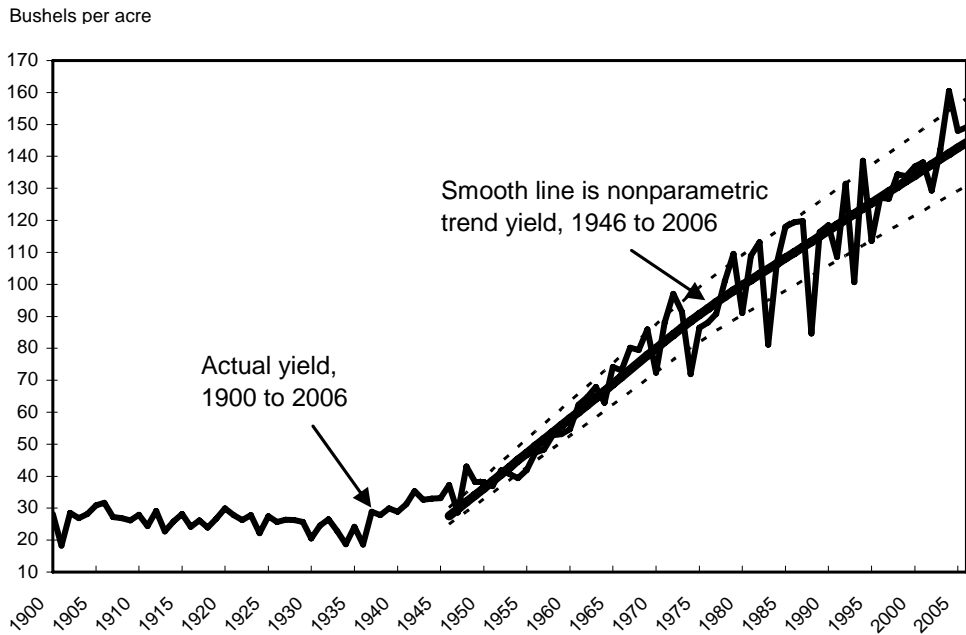


**Figure 1. Month in the LDP availability period with the lowest monthly average price**  
Frequencies are calculated over October 1969 through May 2006 corn price data.

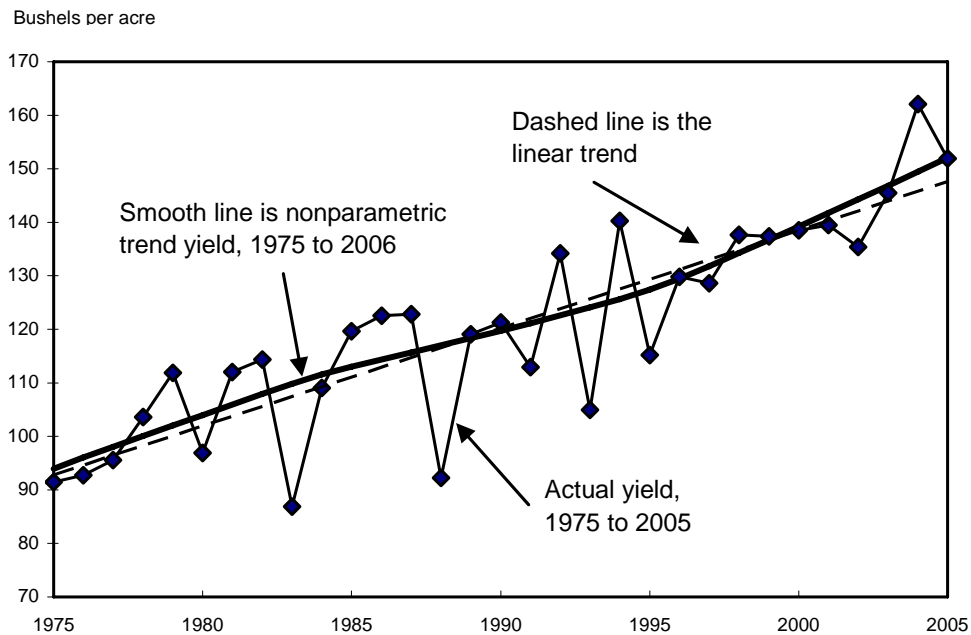


Note: Prices are the CBOT prices for the December corn futures contract.

**Figure 2a. National Average Yield per Planted Acre of Corn for Grain and Silage**  
Annual growth rate in trend yield is 2.81 percent, or 1.95 bushels, over 1946 - 2006.

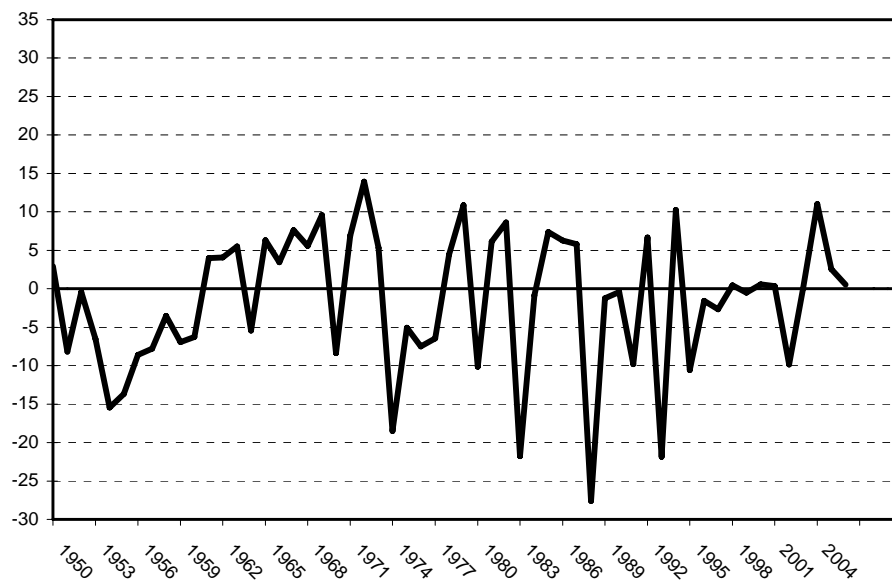


**Figure 2b. National Average Yield over 1975 - 2005**



**Figure 3. Deviation in Actual Yield from Expected Yield for Corn**  
Deviations derived from the nonparametric yield trend.

Percent change



Note: Yield is defined as yield per planted acre of corn for grain and silage

**Table 1. Parametric and Semi-Nonparametric (SNP) Regression Results for the Function Explaining  $\Delta P_t$**

Variable	Parametric	SNP
Constant	-0.054902 (-1.878){-0.103, -0.008}	-0.053 (-2.000){-0.101, -0.006}
$\Delta Y_t$	-1.334 (-5.154){-1.794, -0.889}	-1.284 (-4.547){-1.757, -0.827}
$\sin s(\Delta Y_t)$	—	0.008 (0.495){-0.020, 0.036}
$\cos s(\Delta Y_t)$	—	0.004 (0.248){-0.026, 0.035}
FarmAct	-0.081 (-1.838){-0.144, -0.020}	-0.079 (-1.602){-0.155, -0.006}
$Ln-L$	24.738	24.918
$R^2$	0.529	0.534
$d\Delta P_t/d\Delta Y_t$	-1.334 {-1.796, -0.891}	-1.329 {-1.803, -0.887}

Notes:  $T$ -values are shown in parentheses.

The BCa 90% confidence intervals apply the bias corrected accelerated approach (Efron) to 1000 bootstrap runs, and are shown in brackets.

For the parametric case, the parameter value for  $d\Delta P_t/d\Delta Y_t$  is the same the coefficient on  $\Delta Y_t$ .

**Figure 4. LDP Payments as a Function of Yield Shock – Corn**

The results are derived from the bootstrap analysis of the parametric model.

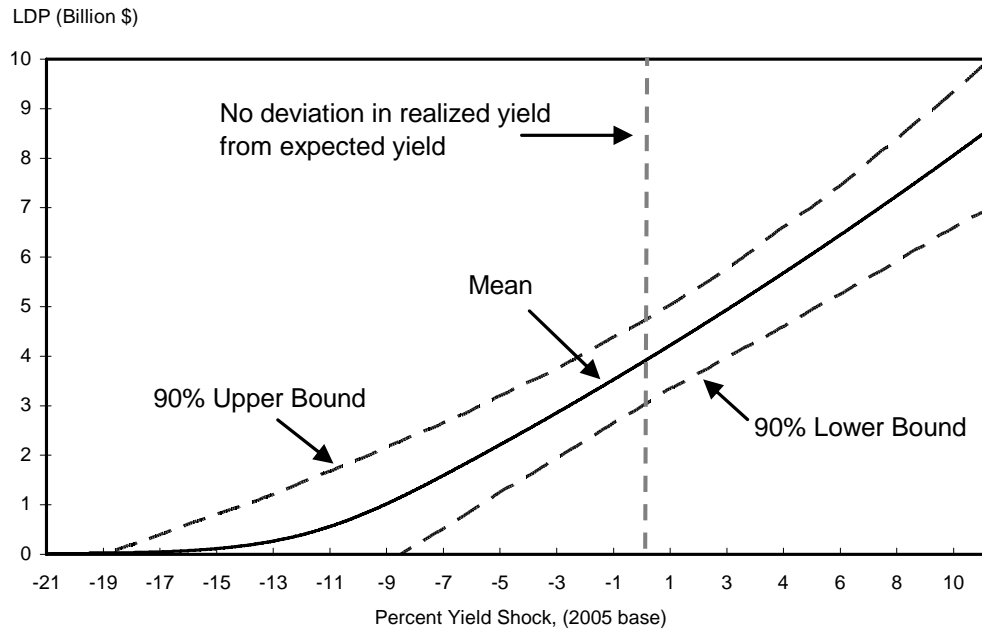
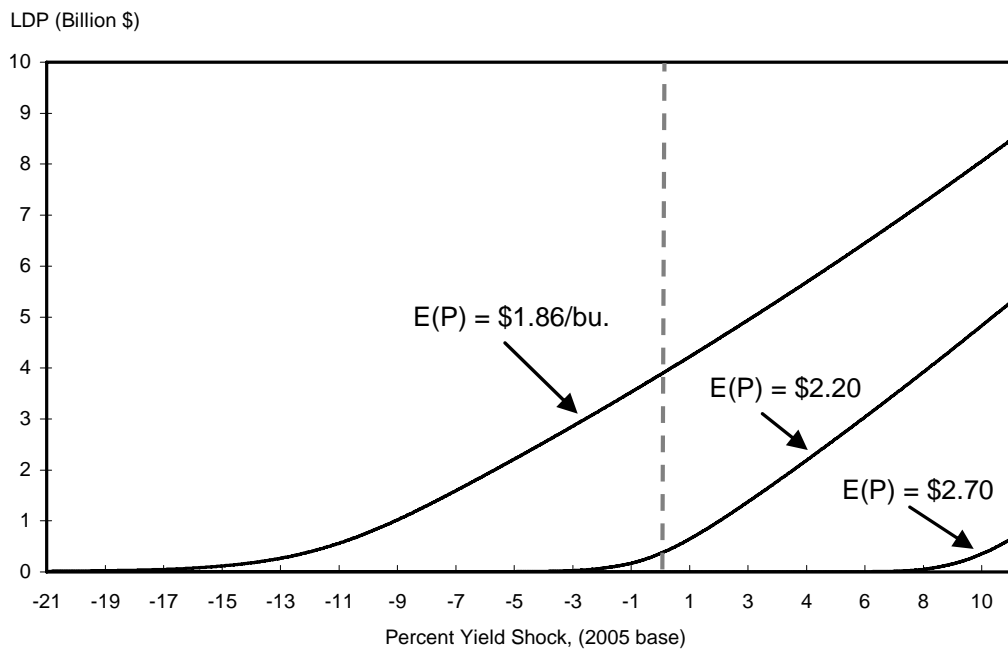


Figure 5. Forecasted Mean LDP Payments for a Range of Expected Prices – Corn



Note: At  $E(P) = \$1.86$ , the producer's effective price is the \$.195 loan rate.