Globally Flexible Modeling of County-Level Acreage Response for Primary U.S. Field Crops

Joseph Cooper and Carlos Arnade

Economic Research Service, USDA

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's 2011 AAEA & NAREA Joint Annual Meeting, Pittsburgh, Pennsylvania, July 24-26, 2011.

jcooper@ers.usda.gov and **carnade@ers.usda.gov**

The views expressed herein are those of the authors and not necessarily those of ERS or the USDA.

Overview

This study takes the standard acreage response model that stems from an expected utility framework, accounting for both price and yield variability, and nests it within a flexible semi-nonparametric (SNP) model consistent with farm-level decision models for computationally tractable results. We use county-level data to estimate the response of farmers' planting preferences to changes in revenue and other variables.

The Issue

Standard acreage response models are predicated on the assumption that observed acreage allocations result from the famer maximizing wealth, income, or profit in an expected utility (EU) framework and, as such, the models are subject to the constraints expected from this theory. To the extent that farmers depart from this theorized behavior, however, the estimates can be biased.

The Basic Parametric Model

Prices and yields are stochastic in our model and, if the farmer is not risk neutral, optimal input choices will be sensitive to price and yield distributions. Yields are denoted by Y_i, output prices (\$/bu.) by p_i, planted acres by A_j, and cost per acre by C_i, for crop *i*. Assume that the constraints on the farm household's input decisions (due to a budget constraint) are represented by $f(\boldsymbol{A}) = 0$, where $\boldsymbol{A} = (A_1,\ldots,A_n)$. The Chavas and Holt (1990) *EU* framework for modeling acreage decisions serve as the theoretical basis for our nested model. In their model, acreage maximizes a concave von Neumann Morgenstern utility function over wealth:

$$(1) \quad \max_{A} \left\{ EU\left[\frac{I}{q} + \sum_{i=1}^{n}\left(\frac{p_i}{q}Y_i - \frac{C_i}{q}\right)A_i\right]\right\} \text{ s.t. } f(\boldsymbol{A}) = 0,$$

where *I* is exogenous wealth and *q* is a price index for household goods. Our model keeps costs *C* and expected revenue *R* separate in the estimation so that costs per acre are measured at the county level. Solving for equation 1 leads to the optimal acreages expressed as exogenous variables, or $A^*(I, \boldsymbol{R}, \boldsymbol{C}, \boldsymbol{\sigma})$, where $\boldsymbol{\sigma}$ represents the covariance matrix of crop revenues. Without data to estimate wealth at the county level, we proxy for wealth via USDA crop reporting district (CRD) fixed effects dummies, where *CRD*s represent aggregations of several adjacent counties.

The parametric reduced form approximation of the equation for acreage response for crop *i* in period *t* is (leaving out the county subscript):

$$(2)\ A_{it} = \alpha_i + \sum_j \beta_{ij} R_{jt} + \sum_j \delta_{ij} C_{jt} + \sum_{k \geq j} \sum_j \gamma_{ijk} \sigma_{ikt} + \Theta_i t$$
$$+ \sum_l \eta_l Gov_{lt} + \tau_i A_{it-1} + \sum_m \lambda_m CRD_m + \varepsilon_t,$$

where R_{jt} is expected revenue per acre, C_{jt} is cost per acre, σ_{ikt} are covariances of revenue, Gov_{lt} is a Government support like the expected value of commodity payments not included in R_{jt} (i.e., R_{jt} and σ_{ikt} already account for the commodity loan rate) or supports decoupled from production. In principle, R_{jt} could contain the net expected Federal crop insurance indemnity, but the wide variety of available Federal crop insurance instruments and coverage levels suggest that is it better for practical purposes to keep separate from revenue in estimation. Other variables include lagged acreage (A_{it-1}), the time trend, and regional dummies (*CRD)*. We do not include σ_{ij}, $i \neq j$ in our empirical estimation because markets are unlikely to appreciate the cross σ_{ij} except in the most abstract sense**.**

Nesting the Parametric Model in a Globally Flexible Model

The Fourier functional form we used for the SNP is the only functional form known to have Sobolev flexibility, so the difference between the model $A^{SNP}(\boldsymbol{x},\theta)$ and the true function $f(\boldsymbol{x})$ can be made arbitrarily small for any value of $\boldsymbol{x}$ as the sample size increases (Gallant, 1982). The

Fourier flexible functional form, which attaches linear and quadratic terms to the Fourier terms to decrease the terms needed to model nonperiodic functions, is specified as

$$\textbf{(3)} \quad A_{it}^{SNP}(\boldsymbol{x}, \theta_k) = \quad U_0 + \boldsymbol{b}'\boldsymbol{x} + 0.5\boldsymbol{x}'D\boldsymbol{x} + 2\sum_{\alpha=1}^{A}\left\{ \sum_{j=1}^{J}\left(v_{j\alpha}\cos[j\boldsymbol{k}'_{\alpha}s(\boldsymbol{x})] - w_{j\alpha}\sin[j\boldsymbol{k}'_{\alpha}s(\boldsymbol{x})]\right)\right\}$$

where (k-A-J) x 1 vectors **b** and **x** are the set of coefficients and variables, respectively, in equation 2, $U_0 = \quad u_0 \quad + \sum_{\alpha=1}^{A}\{ u_{0\alpha} \}$, and $D = \sum_{\alpha=1}^{A} \quad u_{0\alpha}k_{\alpha}' \; k_{\alpha}$, k is the dimension of θ, A (the length) and J (the order) are positive integers, and $\boldsymbol{k}_{\alpha}$ are vectors of positive and negative integers that form indices in the conditioning variables, after shifting and scaling of x by s(x). The function s(x) prevents periodicity in the model.

As parametric equation 2 is nested in equation 3, validity of the parametric specification can simply be assessed by statistically testing whether or not the parameters $u_{0\alpha} = v_{j\alpha} = w_{j\alpha} = 0$, $\forall\, j, \alpha$.

We estimated the parametric and SNP models as acreage shares in a system of seemingly unrelated regression (SUR), where hay is the omitted activity in the system. Using restricted least squares, we imposed symmetric restrictions on the cross revenue and cross cost effects for the untransformed (or parametric) variables, that is $\beta_{ij} = \beta_{ji}$ and is $\delta_{ij} = \delta_{ji}$ for each crop (not imposed on hay). We also imposed homogeneity of degree zero on revenues and costs in each equation. All restrictions apply only to the untransformed variables. We found that the parametric models of acreage response were rejected at any measurable level of significance based on F-tests comparing them with the SNP models.

Data

U.S. county-level data from 1975 to 2007 were used for corn, soybean, winter wheat, and hay. Crop prices were drawn from futures markets. Price and yield densities were converted into within-season deviates (Cooper, 2010). The price deviate was derived from (harvest price-planting price)/planting price. The yield deviate was derived from (actual yield-expected yield)/expected yield. Generated prices for each commodity were truncated by their respective loan rates.

The data are grouped according to their growing history for crops included the regression systems:

Group 1: 573 counties that produce corn for grain, soybeans, and winter wheat (but not spring wheat or cotton).

Group 2: 367 counties that produce corn for grain and soybeans (but not spring wheat or cotton).

Given that the previous 10 years of data were used to generate the means and variances of revenue for each year, the time span for the econometric analysis covers 1985-2007. Sample for the SUR regression for group 1 (group 2) includes 39,537 (16,882) observations. The SNP group 1 (group 2) regression has 516 (274) variables.

Parametric and semi-nonparametric acreage elasticities in counties that grow corn and soybeans only

	Corn				Soybeans			
Variable	Para		SNP		Para		SNP	
R_C ($/acre)	0.310	***	0.291	***	-0.308	***	-0.234	***
R_S ($/acre)	-0.093	***	-0.111	***	0.104	***	0.122	***
R_H ($/acre)	0.018		0.031	**	0.033	***	0.053	***
σ_C	-0.011	*	-0.011	**	0.011		0.015	**
σ_s	-0.007		-0.002		-0.004		-0.011	
C_C ($/acre)	-0.025		-0.017		0.201	***	0.207	***
C_s ($/acre)	0.0452		0.064	**	-0.099	***	-0.151	***
CCP ($/acre)	-0.013	***	-0.018	***	0.004		0.017	***
NI ($/acre)	0.007	***	0.007	***	0.000		0.004	
Trend	0.168	***	0.419	***	-0.109	***	-0.711	***
DP ($/acre)	0.016		0.013		-0.040		-0.036	
A_{t-1}	0.708	***	0.699	***	0.886	***	0.880	***

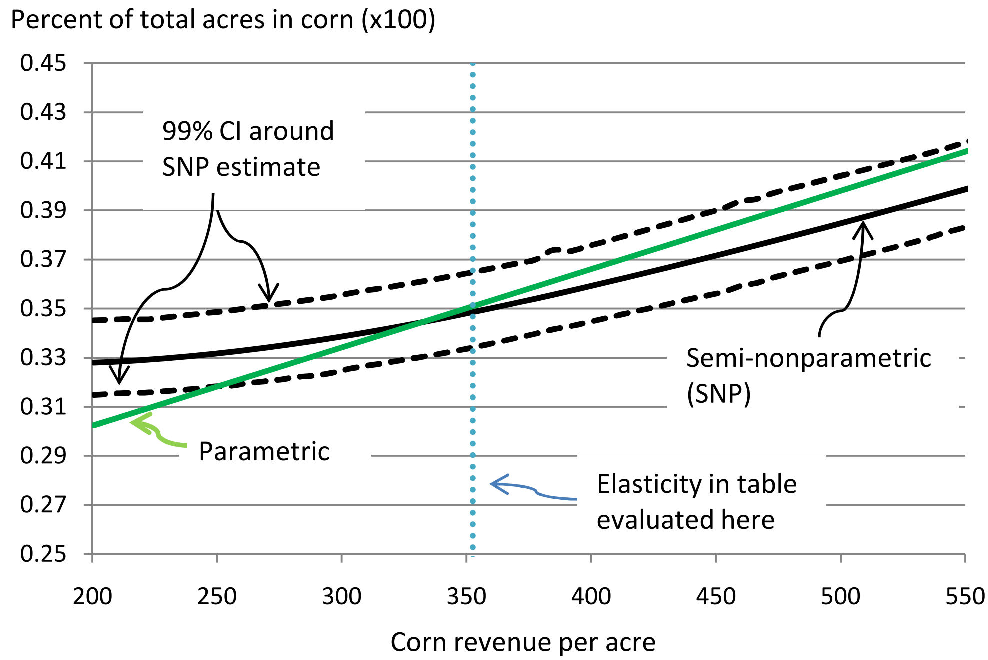
SNP=Semi-nonperametric.
***=Significance at the 1 percent level.
**=Significance at the 5 percent level.
*=Significance at the 10 percent level.
Notes: System adjusted R^2 is 0.95 and 0.94 for the parametric and SNP models, respectively. Elasticities are evaluated at the average of 2006/2007 revenues per acre.
Source: USDA, Economic Research Service calculations based on National Agricultural Statistics Service, Farm Service Agency, and other data.

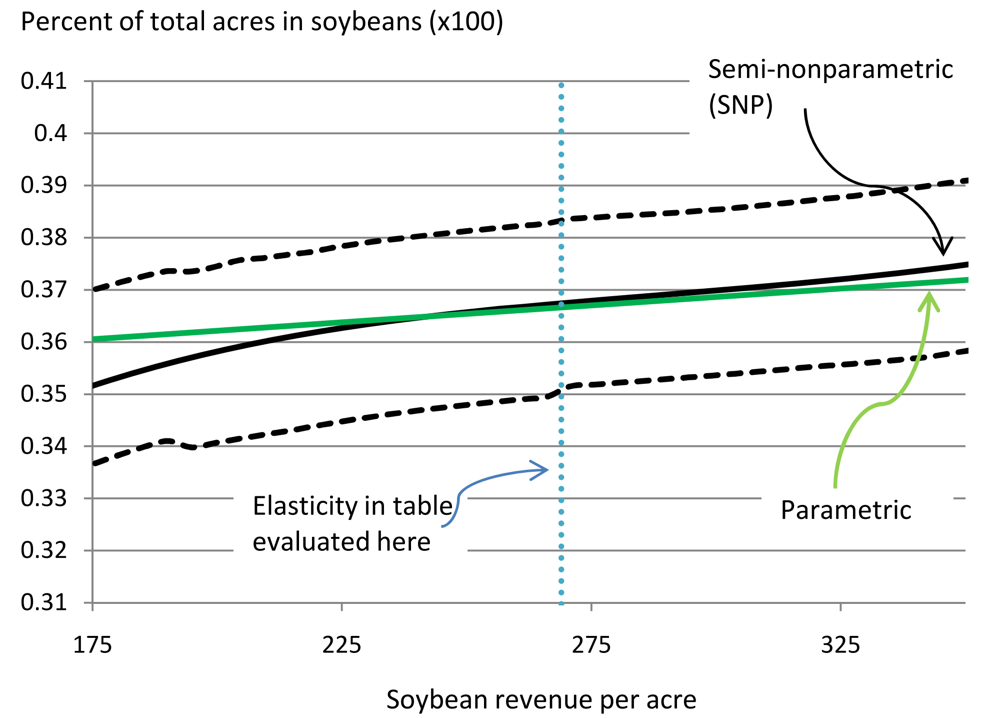
Parametric and semi-nonparametric acreage elasticities in counties that grow corn, soybeans, and winter wheat only

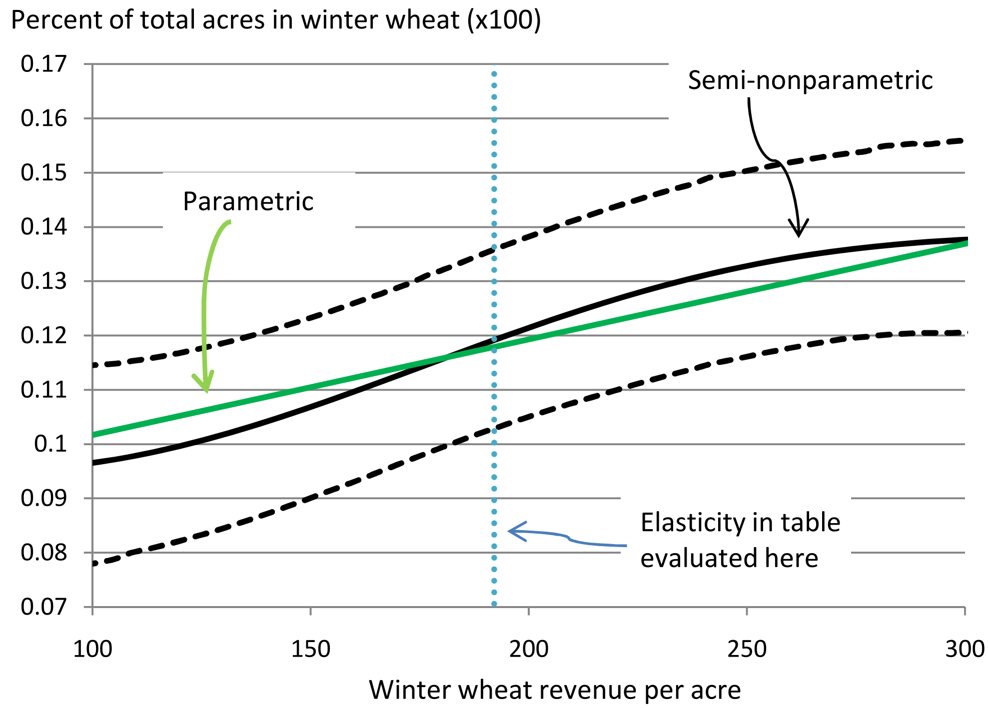
	Corn				Soybeans				Winter wheat			
Variable	Para		SNP		Para		SNP		Para		SNP	
R_C ($/acre)	0.321	***	0.212	***	-0.210	***	-0.137	***	-0.021		-0.076	***
R_S ($/acre)	-0.164	***	-0.064	***	0.048	***	0.061	***	-0.003		0.091	***
R_W ($/acre)	-0.004		-0.081	***	-0.001		-0.067	***	0.272	***	0.447	***
R_H ($/acre)	-0.010	***	0.030	***	-0.009	***	0.011		-0.013	**	-0.006	
σ_c	-0.000		-0.013	**	0.000		0.004		0.0001	**	0.010	
σ_s	0.000	***	0.002		0.000		0.002		-0.0002	***	-0.057	***
σ_w	-0.000	***	-0.012		0.000		-0.006		0.0002	***	0.045	***
C_c ($/acre)	-0.165	***	-0.089	***	0.176	***	0.169	***	-0.235	***	-0.114	***
C_s ($/acre)	0.154	***	0.101	***	-0.077	***	-0.231	***	0.130	***	-0.191	***
C_w ($/acre)	-0.054	***	0.122	***	0.034	***	0.150	***	-0.195	***	-0.093	**
CCP ($/acre)	-0.002	***	-0.042	***	0.000		0.027	***	0.002		-0.079	***
NI ($/acre)	0.002	***	0.005	***	-0.003	***	-0.002		0.003	***	-0.001	
Trend	-0.055	***	0.561	***	0.018	***	-0.816	***	0.174	***	-0.926	***
DP ($/acre)	-0.014	***	0.037	***	0.018	***	-0.041	***	0.015	***	0.020	
A_{t-1}	0.838	***	0.808	***	0.945	***	0.916	***	0.781	***	0.740	***

SNP=Semi-nonparametric.
***=Significance at the 1 percent level.
**=Significance at the 5 percent level.
*=Significance at the 10 percent level.
Notes: System adjusted R^2 is 0.97 for the parametric and SNP models. Elasticities are evaluated at the average of 2006/2007 revenues per acre.
Source: USDA, Economic Research Service calculations based on National Agricutlural Statistics Service, Farm Service Agency, and other data.

Percent of total acres (simulation) as a function of the crop's revenue per acre from corn, soybean, and winter wheat regressions (normalized to 2008 prices)







CI=Confidence interval.
SNP=Semi-nonparametric.
Notes: Functions are graphed over roughly ± 1.5 times the standard deviation of expected crop revenue. Total acres includes hay.
Source: USDA, Economic Research Service calculations based on National Agricultural Statistics Service, Farm Service Agency, and other data.

Findings

Estimated SNP expected revenue elasticities are within the range of those in the existing literature. The nonlinear relationships, however, between acreage and explanatory variables demonstrated by the SNP models show that elasticities vary depending at which points the derivatives of the response functions are evaluated.

Acreage response, with respect to expected revenue and costs, are consistent with expected utility theory for both the parametric and SNP models. Acreage response, with respect to the volatility of revenue, is generally not consistent with standard *a priori* expectations for either model. Acreage response, with respect to this volatility, however, is also quite inelastic.