Endogenous Price in a Dynamic Model for Agricultural Supply Analysis

May 2014

Wei Zhou
Graduate Assistant
Department of Economics
Iowa State University
wzhou815@iastate.edu

Bruce Babcock
Cargill Chair of Energy Economics and
Director, Biobased Industry Center
Professor of Economics
Iowa State University
babcock@iastate.edu

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

Introduction

The purpose of this paper is to present a new model of agricultural supply which combines Positive Mathematical Programming (PMP) with the rational expectations storage model. PMP is an approach widely used for calibrating mathematical models to observed multiple agricultural outputs. The PMP approach is able to generate optimal production plans that replicate observed plans. Popularized by Howitt (1995), PMP has been developed by researchers both in calibration using exogenous supply elasticities (Hechkelei and Britz (2005), Mérel and Bucaram (2010), Mérel et al. (2011)) or estimation using multiple data points (Britz and Heckelei (2000), Jansson and Heckelei (2011)). Most of these models assume exogenous output prices. Arfini et al. (2008) incorporates endogenous prices in PMP by modeling 'farm level' demand functions and cost functions while including the demand functions into profit maximization problem. Later, Arfini and Donati point out that their old approach is inappropriate because their model assumes that individual farms are not price takers so it does not fit the competitive market. Their new approach introduces endogenous prices by maximizing the difference between the total value of the output and the total cost of variable inputs subjected to the aggregating constraints and to individual farm constraints. However, existing PMP models do not consider that the acreage decisions made by forward-looking farmers are determined by expected crop prices due to the biological lags of the agricultural production. One way to include this behavior iss with a rational expectations storage model.

The rational expectations storage model has emerged as a powerful tool in crop price analysis and policy analysis (Williams and Wright (1991), Miranda and Glauber (1993), Gouel (2013a)). It endogenizes expected price by assuming futures prices formed by agents are realized expected prices given all optimal decisions about storage, acreage and consumption. Cafiero et al. (2011) validate the model empirically and raise an important issue of the quality of solutions properties of generated prices. Miranda (1997) compares different numerical methods for solving the storage model, including collocation methods, least squares, space discretization and linearization when approximating current price functional form. He finds that the collocation method with Chebyshev or spline polynomials outperforms the other methods. Gouel (2013b) compares methods for approximating various functions including the value function, the expected crop price function and the storage rule. He claims that expected price function approximation leads to the most accurate result because the expected price function is smooth and close to linear. But expected price approximation is also the most time consuming method. Gouel (2013b) recommends approximating the expected price functions especially when there are several state variables. Comparing the results given by the storage model with and without convenience yield, Gouel (2013b) finds that convenience yield smooths the approximated functions and generally gives higher accuracy for all methods.

However, the collocation method has been used to solve the storage model for no more than two crops because of the curse of dimensionality. To extend the storage model to multiple crops requires consideration of new methods for solving the storage model.

The perturbation method is widely used in solving dynamic stochastic general equilibrium (DSGE) models in macroeconomics. The perturbation method linearizes the solution at the steady state and uses the solution to infer results away from steady states. Because it only requires solving a system of linear equations, it can be easily applied to models with multiple state variables. However, the perturbation approximation performs poorly away from the steady state. Because the economy is usually not around the steady state, this method cannot be applied to the storage model (Gouel (2013b), Miranda (1997)).

Kollman et.al (2006) compares several methods for solving stochastic neoclassical growth models with multiple countries. They suggest using the Smolyak method to interpolate functional forms. Invented by the Russian mathematician Sergey Smolyak (1963), the Smolyak grid is used instead of a tensor gird to interpolate and represent multi-variate functions. First adopted by Krueger and Kubler (2004) to solve an over-lapping generation model in economics, the Smolyak collocation method is also used in many other applications. Malin et. al (2011) apply it to solve a multiple country international real business cycle model. Judd et. al (2013) extended the method by improving grid construction and by developing a non-derivative fixed point algorithm.

Besides the Smolyak method, generalized stochastic simulation algorithm (GSSA) (Judd et al. (2011)), improves the stochastic simulation algorithm (Haan and Marcet (1990)) by replacing inaccurate Monte-carlo and unstable standard least square methods to solve high dimensional dynamic models. GSSA solves the model using a relatively small number of points that are visited in equilibrium rather than by using the collocation method which requires larger domains (Judd (1992)). GSSA has been shown to be numerically stable even with a large number of state variables in multi-country neoclassical growth models (Judd et al. (2011)), an 80-period overlapping generation model (Hasanhodzic and Kotlikoff (2013)) and a search and hiring model with heterogeneous workers and hiring selectivities (Villena-Roldán (2013)).

In this paper, we present a new way of formulating a three-crop competitive storage model. In each period, crops can be consumed or stored for future use. A representative farmer maximizes expected profit using expected crop prices and a cost function. It is assumed that the only production input is land. The cost function is calibrated to conditions in a base year and the implied expected land elasticities coincide with exogenous supply elasticities (Mérel and Bucaram (2010)). The model is calibrated to three crops: corn, soybean and all other crops. The base year is set to be 2013/14 marketing year. This model is the first to combine PMP with endogeneous prices that is solved using a storage model.

We incorporate convenience yield in our model. The idea of convenience yield is first introduced by Kaldor (1939) to explain backwardation. Backwardation is a phenomenon that positive stocks exist even when spot price is less than the next to expire futures price. Convenience yield is often motivated by the option value of storage. For example, producing firms might need to meet a sudden increase in demand to keep consumer's satisfied. Thus they need to keep a certain amount of storage even in backwardation. Convenience yield is greater when stock on hand is smaller. Recent work by Joseph et al. (2011) validate the existence of convenience yield for US corn, soybean and wheat markets using 1990 to 2010 data.

Peterson and Tomek (2005) calibrated a convenience yield function and embedded it in a rational expectations storage model for U.S. corn and found that the model generates price patterns that are consistent with actual commodity prices. Roberts and Tran (2012) didn't use convenience yield and generate too little storage compared to real world storage levels. One explanation could be that they do not consider the convenience yield. Later, Roberts and Tran (2013) used a calibrated negative constant storage cost to represent the existence of convenience yield. In our model we will calibrate a convenience yield similar to Peterson and Tomek's approach.

To solve the model, the expected revenue functions are approximated to obtain the acreage decisions of farmers. In addition, either storage, expected price or current price needs to be approximated as functions of all state variables. If the storage function is approximated, the expected price can be calculated using the known storage rule. If expected price functions are approximated as functions of storage levels of all crops, then the storage levels can be solved using Euler equations.

The model is solved using three approaches: (1) Generalized stochastic simulation approach with fixed point iteration, (2) Smolyak collocation with storage rule approximation and (3) Smolyak collocation with expected price function approximation. Method (1) employs GSSA while method (2) and (3) use Smolyak method. When using Smolyak method, storage rule approximation might be the least time-consuming method and the expected price function approximation might be the most accurate one.

Storage levels for all three crops are state variables in GSSA. Expected crop price are approximated as functions of storage levels for all crops. We simulate the model with draws generated from correlated crop yields to get the state space that is visited in equilibrium. The solutions are computed using the simulated points. Regularized least-absolute deviation (LAD) linear-programming method is used as a numerically stable approximation method. Monomial integration is used instead of quadrature or Monte Carlo. Monomial rules make the integration possible for many random variables and it is applied for all three methods.

The other two methods use Smolyak collocation methods. One approximates the storage rule while the other approximates expected crop prices. The steps for Smolyak collocation method are: (1) Descretize the continuous state space. Finite points are used to approximate a continuous function. Those chosen finite points are called grid points. (2) Find basis functions and the collocation matrix. The approximated function is constructed by unknown coefficients and basis functions. (3) At each grid point, solve the true values of the approximated functions. (4) Solve for the unknown coefficients. Compared to a tensor grid, use of a Smolyak grid requires fewer supporting points to approximate a multivariate function, thus making it feasible to represent higher dimensional functions.

For Smolyak collocation with storage rule approximation, total supply for each crop is a state variable. Storage rule is approximated as a function of total supply. This is a fixed point approach associated with a sparse grid. The fixed point approach requires only direct calculations and it should require less computational time (Judd et. al (2013)).

For Smolyak collocation with expected prices being approximated, storage levels are state variables. Expected prices are represented as functions of storage. This method requires time iteration which means it approximates future prices and expected revenues and solve the current storage decisions using a numerical solver at each grid point. Thus it takes more computational time compared to a fixed point approach, but expected price approximation is shown to be the most accurate algorithm by Gouel (2013).

To date, rational expectations storage models are usually solved one crop with no resource constraint. Our model with multiple crops and land constraint is more suitable for policy analysis. Because of the curse of dimensionality, PMP supply models have never been combined with rational expectations storage model. With the new methods, we show that it is developed a practical to analyze agricultural supply model in a realistic way that incorporates forward-looking rational agents[1].

In the rest part of the paper, we first introduce the basic model with three crops. The computational approaches are presented and the candidate solution qualities are tested next. At the end, we show some simulation results from the model.

[1] The supply side of the model will be further developed to a regional supply model with ten crops in US. The whole US is divided into 10 regions while each region has its own supply elasticities. So the model is calibrated into 10 PMPs. There is a total demand for each crop in the model. Besides the supply side, demand and storage remains the same as the simple model illustrated here. We will simulate the model 5000 times with 5000 sequences of 10 years' crop yields, then we can get crop price distributions for 10 consecutive years. This model will be used for policy analysis similar to what the FAPRI model is used to do.

The Model

The model is a three-crops, rational expectations, competitive storage model. Each agent's optimization problem is described below:

Farmer

A representative farmer maximizes time t expected profits from planting three crops given a land constraint. Crop prices are realized in t+1. Land is the only input and production technology for each activity is Leontief . The farmer's maximization problem follows Mérel and Bucaram (2010) with some modifications:

$$\underset{x_{it+1}}{Max}\sum_{i=1}^{3}P_{it}x_{it+1}-(\lambda_{2i}-\gamma_i\overline{x}_i)x_{it+1}-\frac{1}{2}\gamma_i x_{it+1}^2 \tag{1}$$

$$s.t. \quad \sum_{i=1}^{3}x_{it+1}\leq A$$

For crop i, x_{it+1} is the planted acreage at time t, $\bar{x}_{i1}$ is the observed acreage level in the base year calibration ($t=0$). Let $y_{it+1}$and p_{it+1} be time $t+1$ yield and price for crop i. The discounted expected revenue at the time of planting for each crop i is $\delta E_t(p_{it+1}y_{it+1})$. Assume that $\delta=\frac{1}{1+r}$ where r is the interest rate. C_{it} is the per acre observed cost. $P_{it}=\delta E_t(p_{it+1}y_{it+1})-C_{it}$ is the gross margin. $\gamma=[\gamma_1,\ \gamma_2,\ \gamma_3]$ is the coefficient vector needed to be calibrated so that the model's elasticities are equal to the exogenously determined elasticities. λ_{2i} is used for exact calibration purposes with $\lambda_{2i}=P_{i0}-\bar{\lambda}$ where $\bar{\lambda}$ is the shadow price for the binding land constraint. The value of land rent is suggested to be used for $\bar{\lambda}$ (Gohin and Chantreuil (1999)). The constraint says that the total acreage for three crops is not greater than A where $A=\sum_{i=1}^{3}\overline{x}_{i1}$. In each period, it is assumed that the representative farmer faces the same land constraint.

Instead of using the land supply elasticity with respect to price (Mérel and Bucaram (2010)), we calibrate the model to land supply elasticity with respect to per-acre expected revenue. Let $\bar{\eta}_i$ for $i=1,2,3$ be the expected revenue elasticities in the base year $t=0$, we have

$$\bar{\eta}_i = \frac{dx_{i1}}{dE_0(p_{i1}y_{i1})}\frac{E_0(p_{i1}y_{i1})}{\overline{x}_{i1}} \tag{2}$$

Using (2), the acreage response to per acre gross margin is

$$\frac{dx_{i1}}{dP_{i0}} = \frac{dx_{i1}}{d(\delta E_0(p_{i1}y_{i1}) - C_{it})} = \frac{1}{\delta}\frac{\bar{x}_{i1}\bar{\eta}_i}{E(p_{i1}y_{i1})} \tag{3}$$

$\bar{\eta}_i$ and $\bar{x}_{i1}$ are assumed known. If base year revenue for each crop $E(p_{i1}y_{i1})$ is known, we can get the acreage response and calibrate the unknown parameters γ in the cost function.

Solving the farmer's constrained optimization problem[2],

$$\frac{dx_{i1}}{dP_{i0}} = \frac{1}{\gamma_i}(1 - \frac{\partial\lambda_1}{\partial P_i}) = \frac{1}{\gamma_i}(1 - (\sum_{i=1}^{3}\frac{1}{\gamma_i})^{-1}\frac{1}{\gamma_i}),\ i = 1,2,3 \tag{4}$$

Where λ_1 is the Lagrange multiplier associated with the land constraint. Let $w_i = \frac{dx_{i1}}{dP_{i0}}$, if $w_i < \sum_{j\neq i} w_j$, then we can get positive values of γ_i, $i = 1,2,3$ by solving three unknowns from three equations (4). This condition requires at least three crops in the calibration system and the response of one crop should not be greater than the sum of responses of the other crops.

Note that another assumption here for getting the γ is that the base year expected revenues are known. In the later algorithm section, we will approximate the expected revenue for each crop as a function of state variables in order to solve the farmer's problem in each period. If the base year $E_0(p_{i1}y_{i1})$ is pre-determined, it may not be the same as what is implied by the model. Thus we treat it as endogenous and approximate it in each iteration using the approximated value. More details will be provided in the algorithm section.

Storer

A representative storer maximizes his/her profit from storing crops. At time t, the revenue from storing is the expected crop price in time $t+1$. The cost of storing is storage cost plus the opportunity cost from not selling the crop in time t. Equilibrium storage satisfies the following non-arbitrage condition for crop i

[2]The detailed procedure can be found in Mérel and Bucaram (2010) page 399-402.

$$\delta E_t(p_{it+1}) - p_{it} - SC_{it} = 0 \; i = 1, 2, 3 \tag{5}$$

where $E_t(p_{it+1})$ is the crop i's expected price, per bushel storage cost of crop i is denoted by SC_{it}. The marginal storage cost includes marginal convenience yield which goes to negative infinity when stock level approaches zero. This specification of marginal storage cost will eliminate stock-out conditions.

Consumer

A representative consumer maximizes his/her utility by consuming three crops and a numeraire good m_t given a budget constraint. Assume the utility function is separable. The price for crop i is denoted by p_{it} and total income is denoted by I.

$$\underset{\{c_{it}\}_{i=1,2,3}}{Max} \sum_{t=1}^{\infty} (\sum_{i=1}^{3} U_i(c_{it}) + m_t) \tag{6}$$

$$\sum_{i=1}^{3} p_{it} c_{it} + m_t = I$$

The optimization condition gives us the inverse demand function for each crop i,

$$p_{it} = U_i^{'}(c_{it}) = D_i^{-1}(c_{it}) \tag{7}$$

Equilibrium Condition

In each period, total supply is the sum of total production in time t and carryover stocks. $TS_{it} = h_i x_{it} y_{it} + s_{it}$, where h_i is the harvest rate for crop i. Total supply is then consumed in time t or stored for future use,

$$TS_{it} = c_{it} + s_{it+1}, \; i = 1, 2, 3 \tag{8}$$

The algorithms

Monomial integration is used to descretize multi-normal yield distribution in all algorithms. Formula for the monomial rule used in the paper is described as the second formula in supplementary material to Judd et. al (2011). We used code provided by Judd (2011) to generate monomial nodes and weights. If there are three crops, the total number of nodes equal $3^3 + 1 = 19$. Monomial nodes are denoted by a $N \times 3$ matrix $[y_1, y_2, y_3]$, where y_i is an $N \times 1$ vector representing the monomial nodes for crop i. w is the weight vector where the jth element is the probability for $[y_1(j), y_2(j), y_3(j)]$.

Method 1. GSSA with storage rule approximation

GSSA is an algorithm developed by Judd et al. (2011). Matlab code for implementing the algorithm can be found online. The model requires solving (1), (5), (7), (8). We approximate the storage rules and expected revenues for crop i as functions of all state variables $s_{it} = f_i(TS_{1t}, TS_{2t}, TS_{3t})$ and $E_t(p_{it+1}y_{it+1}) = g_i(TS_{1t}, TS_{2t}, TS_{3t})$, respectively. Flexible functional forms $\psi_i(TS_{1t}, TS_{2t}, TS_{3t}, a_i)$, $\phi_i(TS_{1t}, TS_{2t}, TS_{3t}, b_i)$ and vectors of coefficients a_i and b_i for $i = 1, 2, 3$ are chosen such that $f_i(TS_{1t}, TS_{2t}, TS_{3t}) \approx \psi_i(TS_{1t}, TS_{2t}, TS_{3t}, a_i)$ and $g_i(TS_{1t}, TS_{2t}, TS_{3t}) \approx \phi_i(TS_{1t}, TS_{2t}, TS_{3t}, b_i)$. The detailed steps are as follows:

Initialization: Choose initial guesses $a_i^{(1)}$, $b_i^{(1)}$, $i = 1, 2, 3$. Choose the initial state $(TS_{10}, TS_{20}, TS_{30})$ for simulations. Choose a simulation length T, draw a sequence of crops yields $\{y_{it}\}_{t=1,\ldots,T}, i = 1, 2, 3$. The steps for generating correlated crop yields are as follows.

(i) Let M be variance and covariance matrix for three crops. Let L be the Cholesky decomposition of M.

(ii) Generate a $T \times 1$ vector of random normal deviates for three yields independently. Each vector is denoted by z_1, z_2 and z_3.

(iii) Impose correlation by Cholesky decomposition matrix, $[z_1, z_2, z_3] \times L$.

(iv) Impose mean yields. y_i is the ith column of $[z_1, z_2, z_3] \times L$ plus mean yield of crop i.

Step 1. At iteration p, use $\left\{a_i^{(p)}\right\}_{i=1,2,3}$, $\left\{b_i^{(p)}\right\}_{i=1,2,3}$, calibrate for γ and simulate the model T

periods forward.

(1i) Calibration for γ.

The base year total supplies for corn, soybean and all the others are TS_{10}, TS_{20}, TS_{30}, respectively.

Expected revenue at base year is $\phi_i(TS_{10}, TS_{20}, TS_{30}; b_i^{(p)})$, $i = 1, 2, 3$.

Acreage response for crop i at base year is: $\frac{dx_{i1}}{dP_{i0}} = \frac{1}{\delta}\frac{\bar{x}_{i1}\overline{\eta}_i}{E_0(p_{i1}y_{i1})} = \frac{1}{\delta}\frac{\bar{x}_{i1}\overline{\eta}_i}{\phi_i(TS_0^1, TS_0^2, TS_0^3; b_i^{(p)})}$. Then solve for γ using systems of equations (4). As $b_i^{(p)}$ converges, the expected revenue at the point of base year total supplies converges and so does γ. In this way γ is then calibrated in the whole algorithm. If expected revenues in the base year are fixed and γ is fixed at the beginning of all iterations, then expected revenue is not the same as one implied by the model.

(1ii) When expected revenues are known, solve the farmer's constrained optimization problem (1) to get x_{it+1},$i = 1, 2, 3$.

(1iii) Total supplies, stock levels and expected revenues in $t + 1$ are:

$$TS_{it+1} = \psi_i(TS_{1t}, TS_{2t}, TS_{3t}; a_i^{(p)}) + h_i x_{it+1} y_{it+1} \tag{9}$$

$$s_{it+1} = \psi_i(TS_{1t+1}, TS_{2t+1}, TS_{3t+1}; a_i^{(p)}) \tag{10}$$

$$E_{t+1}^{(p)}(p_{it+2}y_{it+2}) = \phi_i(TS_{1t+1}, TS_{2t+1}, TS_{3t+1}, b_i^{(p)}),\ i = 1, 2, 3 \tag{11}$$

The model then can be simulated T periods forward using (9), (10), (11).

Step 2. The storage can be approximated as:

$$z_{it}^s = TS_{it} - D_i(\delta E_t(P_{it+1}) - k_i) \tag{12}$$

Expected revenue for crop i in time t can be approximated as (13) using monomial nodes and weights:

$$z_{it}^{epy} = \sum_{j=1}^{N} w(j) D_i^{-1}(TS_{it+1,j} - s_{it+1,j}) y_i(j) \tag{13}$$

where the next period total supply vector for all yield nodes for crop i , the next period storage rule for each crop i at each monomial node j and expected price for crop i in time t are defined by

$TS_{it+1,j} = s_{it} + h_{it} x_{it+1} y_i(j)$

$s_{it+1,j} = \psi_i(TS_{1t+1,j}, TS_{2t+1,j}, TS_{3t+1,j}, a_i^{(p)})$

$E_t(p_{it+1}) = \sum_{j=1}^{N} w(j) D_i^{-1}(TS_{it+1,j} - s_{it+1,j})$

Step 3. Find $\{\hat{a}_i\}_{i=1,2,3}$ and $\{\hat{b}_i\}_{i=1,2,3}$ that minimize the errors ϵ_{it}, ζ_{it} in the regression equation using LAD method as described in Judd et. al (2011).

$$z_{it}^{s} = \psi_i(TS_{1t}, TS_{2t}, TS_{3t}, a_i^{(p)}) + \epsilon_{it},\ i = 1, 2, 3 \tag{14}$$

$$z_{it}^{epy} = \phi_i(TS_{1t}, TS_{2t}, TS_{3t}, b_i^{(p)}) + \zeta_{it},\ i = 1, 2, 3 \tag{15}$$

Step 4. Check the convergence and end (2) if

$$\frac{1}{T}\sum_{1}^{T}\sum_{i=1}^{3}\left(\left| \frac{s_{it}^{(p)} - s_{it}^{(p-1)}}{s_{it}^{(p-1)}} \right| + \left| \frac{E_t^{(p)}(p_{it}y_{it}) - E_t^{(p-1)}(p_{it}y_{it})}{E_t^{(p-1)}(p_{it}y_{it})} \right|\right) < \varepsilon \tag{16}$$

where $s_{it}^{(p)}$,$s_{it}^{(p-1)}$, $E_t^{(p)}(p_{it}y_{it})$, $E_t^{(p-1)}(p_{it}y_{it})$ are the storage and expected revenue series obtained on iteration p and $p-1$.

Step 5. Compute $a_i^{(p+1)}$ and $b_i^{(p+1)}$ for iteration $(p+1)$ for $i = 1, 2, 3$.

$$a_i^{(p+1)} = (1 - \xi) a_i^{(p)} + \xi \widehat{a}_i \tag{17}$$

$$b_i^{(p+1)} = (1 - \xi) b_i^{(p)} + \xi \widehat{b}_i \tag{18}$$

where $\xi \in (0, 1]$ is a damping parameter. Go to (2) with new coefficients $a_i^{(p+1)}$, $b_i^{(p+1)}$ and stop the iteration when convergence criterion is reached. After convergence is achieved, we have a model with approximated expected revenue functions and storage functions for each crop.

Smolyak collocation method

The Smolyak method was introduced by Smolyak (1963) to approximate multivariate functions. Compared to the use of standard tensor grids, Smolyak grids require fewer support nodes. For example, if we want to approximate an N dimensional function, the number of tensor nodes is 5^N if using 5 points for each dimension. If $N = 10$, the total number of collocation nodes equals 9,765,625. Smolyak nodes are constructed by the levels of approximation. Higher approximation level leads to higher accuracy. For 10 dimension with 2nd, 3rd or 4th level approximation, the Smolyak grid requires 221 points, 1581 points and 8801 points respectively. Thus, the Smolyak method makes it feasible to solve high dimensional models.

Method 2 Smolyak collocation with storage rule approximation

Step 1: Define the interval for state variables. Let the total supply of crop i between $[TS^i_{min}, TS^i_{max}]$, $i = 1, 2, 3$. The minimum and maximum total supply values should not be violated in iterations.

For each Smolyak grid point $(TS_{1k}, TS_{2k}, TS_{3k})$, $k = 1, ...K$, guess $K \times 1$ coefficient vectors $a_i = [a_{i1}, ..., a_{iK}]$ and $b_i = [b_{i1}, ..., b_{iK}]$, $i = 1, 2, 3$. Storage levels are approximated as $s_{ik} = f^{d,\mu}(TS_{1k}, TS_{2k}, TS_{3k}; a_i)$, $i = 1, 2, 3$, expected revenues are approximated as $epy_{ik} = f^{d,\mu}(TS_{1k}, TS_{2k}, TS_{3k}; b_i)$, $i = 1, 2, 3$. $f^{d,\mu}$ will be defined later by (23). d is the number of dimensions and μ is the approximation level.

Equidistant grid points performs worse than Chebyshev-based nodes for interpolation, therefore the Chebyshev-Gauss-Lobatto grid is used as suggested in Klimke (2006). Chebyshev-Gauss-Lobatto grid is a kind of sparse grid using extrema of Chebyshev polynomials. The details of forming uni-dimension Chebyshev-Gauss-Lobatto grid can be found in Judd et al. (2013) appendix A. We use the sparse grid interpolation toolbox developed by Andreas Klimke (2007) to obtain the Chebyshev-Gauss-Lobatto points.

To see how the Smolyak grids are constructed. We first show the nodes constructed in uni-dimention. Then we show a special case used in our algorithm, three dimensions with three approximation level.

The part of forming grid points and basis functions follows Malin et. al (2007) and Judd et al. (2013). The set of grid points X_i is defined as the set of the extrema of the Chebyshev polynomials with a number of $m(i)$ points in each set. $m(i) = 2^{i-1} + 1$ when $i \geqslant 2$ and $m(1) = 1$.

The formula for extrema of the Chebychev polynomials is $X_i = -cos(\frac{\pi(i-1)}{m(i)-1})$ $i = 1, 2, ..., m(i)$.

The set for extrema of the Chebyshev polynomials are:

$i = 1$, $X_1 = \{0\}$;

$i = 2$, $X_2 = \{-1, 0, 1\}$;

$i = 3$, $X_3 = \{-1, -\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}, 1\}$;

when $i = 5$, there are 17 points in the set, the set of grid points $X_5 = -cos(\frac{\pi(i-1)}{17-1})$, $i = 1, 2, ..., 17$.

From the construction, we can see that X_i is a subset of X_j when $j > i$.

For higher dimensions, the Chebyshev-Gauss-Lobatto grid is formed as follows:

In the three dimension case, we must select tensor products of points selected from unidimention according to

$d \leqslant i_1 + i_2 + i_3 \leqslant d + \mu$.

For example, in the three dimensional case, $d = 3$.

If $\mu = 1$, $3 \leqslant i_1+i_2+i_3 \leqslant 4$. Thus the sets for $\{(i_1, i_2, i_3)\} = \{(1, 1, 1), (1, 1, 2), (1, 2, 1), (2, 1, 1)\}$.

Given a value of i_1, i_2, i_3, find the set of extrema of the Chebychev polynomials (X_i) in each dimension with $i = i_1$, $i = i_2$, $i = i_3$.

When $i_1 = 1$, $i_2 = 1$, $i_3 = 1$, then the tensor product of the multidimensional nodes of each dimension is $\{(0, 0, 0)\}$.

When $i_1 = 1, i_2 = 1, i_3 = 2$, we have X_2 for the third dimension and X_1 for the other dimensions. The tensor product of the points are $\{(0, 0, 0), (0, 0, 1), (0, 0, -1)\}$.

Doing this for all $\{i_1, i_2, i_3\}$, we have seven points in the first level approximation are

$\{(0, 0, 0), (0, 0, 1), (0, 0, -1), (0, 1, 0), (0, -1, 0), (1, 0, 0), (-1, 0, 0)\}$.

If $\mu = 2$, $3 \leqslant i_1 + i_2 + i_3 \leqslant 5$, There are several combinations of i_1, i_2, i_3 that satisfy this restriction: $\{(1, 1, 1), (1, 1, 2), (1, 2, 1), (2, 1, 1), (1, 1, 3), (1, 3, 1), (3, 1, 1), (1, 2, 2), (2, 1, 2), (2, 2, 1)\}$.

Thus we have 25 Smolyak grid points,

$$\{(0,0,0),(0,0,1),(0,0,-1),(0,1,0),(0,-1,0),(1,0,0),(-1,0,0),(0,0,-\frac{1}{\sqrt{2}}),(0,0,\frac{1}{\sqrt{2}}),$$
$$(0,-\frac{1}{\sqrt{2}},0),(0,\frac{1}{\sqrt{2}},0),(-\frac{1}{\sqrt{2}},0,0),(\frac{1}{\sqrt{2}},0,0),(0,-1,-1),(0,-1,1),(0,1,-1),(0,1,1),$$
$$(-1,0,-1),(-1,0,0),(1,0,-1),(1,0,1),(-1,-1,0),(-1,1,0),(1,-1,0),(1,1,0)\}$$

If $\mu = 3$, $3 \leqslant i_1 + i_2 + i_3 \leqslant 6$. The i_1, i_2, i_3 satisfy for the restriction are

$$\{(i_1,i_2,i_3)\} = \{(1,1,1),(1,1,2),(1,2,1),(2,1,1),(1,1,3),(1,3,1),(3,1,1),(1,2,2),$$
$$(2,1,2),(2,2,1),(1,1,4),(1,4,1),(4,1,1),(1,2,3),(1,3,2),(2,1,3),(2,3,1),(3,2,1),(3,1,2)\}.$$

and there are 69 Smolyak grid points.

Step 2: Construct Smolyak Chebyshev basis functions. Here we only provide the details for 3 dimensions (d) with 3rd approximation level (μ) . The Smolyak polynomial function is given by a general form:

$$f^{d,\mu}(TS_{1k},...,TS_{nk};a) = \sum_{max(d,\mu+1)\leq|i|\leq d+\mu} (-1)^{d+\mu-|i|}\left(\begin{array}{c} d-1 \\ d+\mu-|i| \end{array}\right)p^{|i|}(TS_{1k},...,TS_{nk}) \tag{19}$$

$(-1)^{d+\mu-|i|}\left(\begin{array}{c} d-1 \\ d+\mu-|i| \end{array}\right) = (-1)^{d+\mu-|i|}\frac{(d+\mu-|i|)!}{(d-1)!(\mu-|i|+1)!}$ is a counting coefficient to insure that there are no repeated basis functions. n is the number of state variables. A tensor product operator $p^{|i|}(TS_{1k},...,TS_{nk})$ is defined as

$$p^{|i|}(TS_{1k},...,TS_{nk}) = \sum_{i_1+...+i_d=|i|} p^{i_1,...,i_d}(TS_{1k},...,TS_{nk}) \ ,$$

where $p^{i_1,...,i_d}$ is defined as

$$p^{i_1,...,i_d}(TS_{1k},...,TS_{nk}) = \sum_{l_1=1}^{m(i_1)} ... \sum_{l_d=1}^{m(i_d)} a_{l_1...l_d}\psi_{l_1}(TS_{1k})...\psi_{l_d}(TS_{nk}) \ ,$$

where $m(i_j) = 2^{i_j-1}+1$, $i_j \geq 2$, $m(1) = 1$. In our example we have $d = 3$, $\mu = 3$, $n = 3$.

Counting factor is $(-1)^{3+3-|i|}\left(\begin{array}{c} 3-1 \\ 3+3-|i| \end{array}\right)$, $|i| = 4, 5, 6$, $a_{l_1 \ldots l_d}$ are the coefficients. We have

$$p^4 = p^{1,1,2} + p^{1,2,1} + p^{2,1,1} \tag{20}$$

$$p^5 = p^{1,1,3} + p^{1,3,1} + p^{3,1,1} + p^{1,2,2} + p^{2,2,1} + p^{2,1,2} \tag{21}$$

$$p^6 = p^{1,1,4} + p^{1,4,1} + p^{4,1,1} + p^{2,1,3} + p^{2,2,2} + p^{3,1,2} \tag{22}$$

let $c_{j_1 j_2 j_3}$ represents $a_{l_1 l_2 l_3}\psi_{l_1}(TS_{1k})\psi_{l_2}(TS_{2k})\psi_{l_3}(TS_{3k})$. $\psi_{l_1}(TS_{1k})\psi_{l_2}(TS_{2k})\psi_{l_3}(TS_{3k})$ is a basis function where $\psi(\cdot)$ is Chebyshev polynomial basis function.

The Smolyak polynomial function is:

$$\begin{aligned} f^{3,3}(TS_{1k}, TS_{2k}, TS_{3k}; a) = c_{111} + c_{112} + c_{113} + c_{114} + c_{115} + c_{116} + c_{117} + b_{118} + c_{119} + c_{121} \\ +c_{131} + c_{141} + c_{151} + c_{161} + c_{171} + c_{181} + c_{191} + c_{211} + c_{311} + b_{411} \\ +c_{511} + c_{611} + c_{711} + c_{811} + c_{911} + c_{122} + c_{123} + c_{124} + c_{125} + b_{132} \\ +c_{133} + c_{134} + c_{135} + c_{142} + c_{143} + c_{152} + c_{153} + c_{212} + c_{213} + b_{214} \\ +c_{215} + c_{312} + c_{313} + c_{314} + c_{315} + c_{412} + c_{413} + c_{521} + c_{531} + b_{221} \\ +c_{231} + c_{241} + c_{251} + c_{321} + c_{331} + c_{341} + c_{351} + b_{421} + c_{431} + b_{521} \end{aligned} \tag{23}$$

$$+c_{531} + c_{222} + c_{223} + c_{232} + c_{322} + c_{333} + c_{332} + c_{333} + c_{323} + c_{323} + c_{233} \tag{24}$$

The set of Chebyshev polynomial basis functions are defined recursively as follows: $\psi_1(x) = 1$, $\psi_2(x) = x$, $\psi_n(x) = 2x\psi_{n-1}(x) - \psi_{n-2}(x)$. The Chebyshev matrix is denoted by Φ. In this case Φ is a 69×69 matrix.

Approximate storage and expected revenues for each crop i as follows:

$s_i = f^{d,\mu}(TS_{1k}, ..., TS_{nk}; a_i)$.

$epy_i = f^{d,\mu}(TS_{1k}, ..., TS_{nk}; b_i)$, where $f^{d,\mu}$ is defined as equation (19).

Step 3. At iteration p, use $\left\{a_i^{(p)}\right\}_{i=1,2,3}$, $\left\{b_i^{(p)}\right\}_{i=1,2,3}$. Calibrate for γ.

(3i) The base year total supplies for corn, soybean and all the others are TS_{10}, TS_{20}, TS_{30}. All Chebyshev-Gauss-Lobatto grid points are normalized to $[-1, 1]$, so we need to normalize total supplies in the base year before constructing the base year expected revenues.

$TS_0^1 = \frac{2(TS_{10} - TS^1_{min})}{(TS^1_{max} - TS^1_{min})} - 1$, $TS_0^2 = \frac{2(TS_{20} - TS^2_{min})}{(TS^2_{max} - TS^2_{min})} - 1$, $TS_0^3 = \frac{2(TS_{30} - TS^3_{min})}{(TS^3_{max} - TS^3_{min})} - 1$.

(3ii) Expected revenues at base year are $f^{3,3}(TS_0^1, TS_0^2, TS_0^3; b_i^{(0)})$, $i = 1, 2, 3$.

Acreage responses at base year are: $\frac{dx_{i1}}{dP_{i0}} = \frac{1}{\delta}\frac{\bar{x}_{i1}\bar{\eta}_i}{E_0(p_{i1}y_{i1})} = \frac{1}{\delta}\frac{\bar{x}_{i1}\bar{\eta}_i}{f^{3,3}(TS_0^1, TS_0^2, TS_0^3; b_i^{(0)})}$, $i = 1, 2, 3$. Then solve γ using (4).

Step 4. At each grid point k, solve the optimal acreage decisions x_{ik}, $i = 1, 2, 3$, solve storage and expected revenue at each grid point using monomial integration of expectations. The expected revenues are

$$z_{ik}^{epy} = \sum_{j=1}^{N} w(j) D_i^{-1}(TS_{ikj} - f^{3,\mu}(TS_{1kj}, TS_{2kj}, TS_{3kj}; a_i^{(p)}))y_i(j),\ i = 1, 2, 3,\ k = 1, 2, ...K. \quad (25)$$

The storage levels are

$$z_{ik}^{s} = TS_{ik} - D_i(\delta ep_{ik} - SC_{ik},\ i = 1, 2, 3,\ k = 1, 2, ..., K. \quad (26)$$

where the expected prices are

$$ep_{ik} = \sum_{j=1}^{N} w(j) D_i^{-1}(TS_{ikj} - f^{3,\mu}(TS_{1kj}, TS_{2kj}, TS_{3kj}; a_i^{(p)})),\ l = 1, 2, 3,\ i = 1, 2, 3. \quad (27)$$

The next period total supply for crop i with yield $y_i(j)$ is $TS^n_{ikj} = f^{3,\mu}(TS_{1k}, TS_{2k}, TS_{3k}; a_i^{(p)}) + h_i x_{ik} y_i(j)$. Because we need to transform our nodes which normalized in the interval $[-1, 1]$ to the interval $[TS^i_{min}, TS^i_{max}]$, each crop i will be $TS_{ikj} = \frac{(TS^n_{ikj}+1)}{2}(TS^i_{max} - TS^i_{min}) + TS^i_{min}$

Step 5. See if the approximated storage and expected revenue functions converge,

$$\frac{1}{K}\sum_{1}^{K}\sum_{i=1}^{3}(\mid \frac{s_{ik}^{(p)} - s_{ik}^{(p-1)}}{s_{ik}^{(p)}} \mid + \mid \frac{epy_{ik}^{(p)} - epy_{ik}^{(p-1)}}{epy_{ik}^{(p-1)}} \mid) < \varepsilon \quad (28)$$

Step 6. Update the coefficients if the convergence criterion is not satisfied.

z_i^s is a $K \times 1$ vector with kth element equals to z_{ik}^s. z_i^{epy} is a $K \times 1$ vector with kth element equals to z_{ik}^{epy}.

$$a_i^{(p+1)} = (1-\xi)a_i^{(p)} + \xi\Phi^{-1}z_i^s,\ i = 1,2,3. \tag{29}$$

$$b_i^{(p+1)} = (1-\xi)b_i^{(p)} + \xi\Phi^{-1}z_i^{epy},\ i = 1,2,3. \tag{30}$$

Step 7. go to step 3 until the condition in step 5 is satisfied.

Method 3. Smolyak collocation with expected price approximated

Initialization: storage levels of three crops are state variables. Pick grid points and construct basis functions as described in method 2. The Chebyshev matrix is denoted by Φ.

Step 1. Define the intervals for state variables. The storage levels of three crops is contained in $[0, s_{max}^i]$ for $i = 1,2,3$.

At each collocation node (s_{1k}, s_{2k}, s_{3k}), expected prices are approximated as $ep_{ik} = f^{3,\mu}(s_{1k}, s_{2k}, s_{3k}; a_i^{(0)})$, expected revenues are approximated as $epy_{ik} = f^{3,\mu}(s_{1k}, s_{2k}, s_{3k}; b_i^{(0)})$ where $\left\{a_i^{(0)}\right\}_{i=1,2,3}$, $\left\{b_i^{(0)}\right\}_{i=1,2,3}$ are coefficient vectors with initial guesses.

Step 2. At iteration p, use $a_i^{(p)}$, $b_i^{(p)}$. Calibrate γ. The base year observed storage levels for corn, soybean and all the others are s_{10}, s_{20}, s_{30}. Normalize storage levels to $[-1, 1]$.

$s_0^1 = \frac{2s_{10}}{s_{max}^1} - 1$, $s_0^2 = \frac{2s_{20}}{s_{max}^2} - 1$, $s_0^3 = \frac{2s_{30}}{s_{max}^3} - 1$.

Expected revenues at base year are $f^{3,3}(s_0^1, s_0^2, s_0^3; b_i^{(0)})$, $i = 1,2,3$.

Acreage responses at base year are: $\frac{dx_{i1}}{dP_{i0}} = \frac{1}{\delta}\frac{\bar{x}_{i1}\overline{\eta}_i}{E_0(p_{i1}y_{i1})} = \frac{1}{\delta}\frac{\bar{x}_{i1}\overline{\eta}_i}{f^{3,3}(s_0^1, s_0^2, s_0^3; b_i^{(0)})}$, $i = 1,2,3$. Then solve γ using systems of equations (4).

Step 3. For each collocation node, solve the optimal acreage decision x_{ik}, $i = 1,2,3$.

Solve the storage decision at each grid point and each monomial node, where s_{ikj} is the storage decision solved from the non-arbitrage condition:

$$\delta f^{3,\mu}(s_{1kj}, s_{2kj}, s_{3kj}; a_i) - D_i^{-1}(TS_{ikj} - s_{ikj}) - SC_{ikj} = 0,\ i = 1,2,3 \tag{31}$$

where total supply for crop i is $TS_{ikj} = s_{ik} + h_i x_{ik} y_i(j)$.

Thus the expected revenues are

$$z_{ik}^{epy} = \sum_{j=1}^{N} w(j) D_i^{-1} (TS_{ikj} - s_{ikj}) y_i(j), \; i = 1,2,3, k = 1,2,...,K \tag{32}$$

The expected prices are

$$z_{ik}^{ep} = \sum_{j=1}^{N} w(j) D_i^{-1} (TS_{ikj} - s_{ikj}), \; i = 1,2,3, \; k = 1,2,...,K \tag{33}$$

Step 5. See if the approximated expected prices and expected revenues converges

$$\frac{1}{K}\sum_{1}^{K}\sum_{i=1}^{3}(\mid \frac{ep_{ik}^{(p)} - ep_{ik}^{(p-1)}}{ep_{ik}^{(p)}} \mid + \mid \frac{epy_{ik}^{(p)} - epy_{ik}^{(p-1)}}{epy_{ik}^{(p-1)}} \mid) < \varepsilon \tag{34}$$

Step 6. If not, update the coefficients:

$$a_i^{(p+1)} = (1-\xi) a_i^{(p)} + \xi \Phi^{-1} z_i^{epy}, \; i = 1,2,3. \tag{35}$$

$$b_i^{(p+1)} = (1-\xi) b_i^{(p)} + \xi \Phi^{-1} z_{ik}^{ep}, \; i = 1,2,3. \tag{36}$$

z_i^{epy} is a $K \times 1$ vector with $z_i^{epy}(k) = z_{ik}^{epy}$, z_i^{ep} is a $K \times 1$ vector with $z_i^{ep}(k) = z_{ik}^{ep}$.

step 6. go to step 2 until the condition in step 5 is satisfied.

Calibration for the Model

The model is calibrated to corn, soybeans and “other” which includes wheat and cotton for simplicity in this example.

Land Allocation:

We calibrate the land allocation problem to exogenous elasticities and endogenous expected revenue so that the optimal allocation is the same as what was projected to happen in 2013/14 according to WASDE 2013/14 January report. The acreages allocated to corn, soybean and other in the base year are 95.4 million acres, 74.5 million acres and 66.61 million acres, respectively. Thus, total acreage is 238.51 million acres for all time periods. The exogenous supply elasticities for corn, soybean and all others are assumed to be 0.25, 0.2, 0.2 as illustration purposes. Land rent is assumed to be \$200 per acre, so $\bar{\lambda} = 200$.

Yield Distributions

National acreage crop yields from 1970 to 2013 given by USDA NASS are employed to get the yield distributions. All crop yield distributions are assumed to be normal for illustration of the approach. Detrended data for each crop is used separately to fit a normal distribution with mean μ and standard deviation σ, $N(\mu, \sigma^2)$, we have corn yield in bushel per acre, $y_1 \sim N(157, 15.4^2)$, soybean yield in bushel per acre, $y_2 \sim N(43.6, 3.18^2)$. The yield data for other in each year is the average yield of wheat and cotton weighted by output. Thus yield distribution for others is $y_3 \sim N(1.1801, 0.0818^2)$ with tons per acre as the unit. The correlation between crop yield variables are also calculated using the same detrended crop yield data. The covariance between crop i and crop j is estimated as

$$COV_{ij} = \frac{1}{N}\sum_{t=1}^{N}(y_{it} - \overline{y}_i)(y_{jt} - \overline{y}_j) \tag{37}$$

where N is the total number of observations, y_{it} is the detrended yield for crop i at year t and $\bar{y}_i$ is the average yield for crop i for all observations. The covariance between crops are $COV_{12} = 21.34$, $COV_{13} = 0.20704$, $COV_{23} = -0.02897$. With the above information, we can construct the variance covariance matrix for the multivariate normal distribution for all three crops. Using Cholesky decomposition, the variance covariance structure is imposed on the simulated yields.

Demand Functions

Demand functions are assumed to be constant elasticity:

$$D^{-1}(\cdot) = \alpha_{1i} c_i^{-\alpha_{2i}},\ i = 1, 2, 3 \tag{38}$$

The demand parameters are calibrated to the total use and average prices received by farmers for 2013/14. Total use of corn including food, feed, export and ethanol is 13.150 billion bushels. Total use of soybean is 3.3040 billion bushels. Average prices received by farmers for corn and soybean are \$4.40 per bushel and \$12.50 per bu. Total use and price of wheat and cotton, defined as average of wheat and cotton weighted by output, are 68.36 million ton and \$312.40 per ton. In our program, we set the units for quantities of corn and soybean to be 10 billion bushels. The units for all the other to be 100 million tonnes. The price units are dollar per bushel for corn and soybean and dollar per ton for all the others. Demand elasticities for corn and soybean are -0.44 (Adjemian and Smith. (2012)) and -0.236 (Roberts and Schlenker (2013)), respectively. The other crop demand elasticity is assumed to be -0.1 as an illustration purpose. Thus we have $\alpha_{11} = 8.192456$, $\alpha_{21} = 2.27$, $\alpha_{12} = 0.114537$, $\alpha_{22} = 4.237288$, $\alpha_{13} = 0.069614$, $\alpha_{23} = 10$.

Storage Cost and Marginal Convenience Yield

The storage cost per unit includes per unit observed cost and per unit unobserved cost. The observed part of the storage cost (OSC) is a constant physical storage cost paid by the storer. We assume the observed per bushel storage cost is 3 cents per bushel per month (Peterson and Tomek (2010)). The yearly observed storage cost is thus \$0.36 per bushel, $OSC = 0.36$. One component of unobserved storage cost is the opportunity cost that increases with stock level when stock levels are large. This is because holding more stock of one crop decreases the opportunity of holding other more profitable crops (Paul (1970)). The other unobserved storage cost is the marginal convenience yield. We can get unobserved storage cost from the storage non-arbitrage condition: $USC_{it} = \delta E_t(p_{it+1}) - p_{it} - OSC_{it}$. We collect our data including current price p_t, expected price p_{t+1} from 2001/2002 to 2011/2012. Use average price received by the farmer in each marketing year from USDA NASS for the current year price. For corn, the December corn futures price from September 1st to August 31st is used as the average of yearly expected price. The discount factor, δ, is defined as $\frac{1}{1+r}$ where r is the interest rate. We use the return for 1-year treasury constant maturities as risk free interest rate. The unobserved storage cost should be increasing with stock level and it is negative when stock level s_t is small and positive when stock level is high. Rui and Miranda (1995) uses a logarithmic function to achieve it. In Tomek and Peterson (2005), they incorporate expected total supply in marginal convenience yield function and successfully avoid stock-out condition using that specification. Because current price indicates relative shortage of the crop (Na Jin(2013)), we assume the unobserved cost is also increasing in price when it is positive, $UOC_{it} = p_{it}(a_i + b_i log(s_{it+1}))$ where a_i and b_i are parameters needed to be calibrated. For each crop i, when stock level is low, $a_i + b_i log(s_{it+1})$ is negative, current price is positive, unobserved storage cost is negative. When stock level is large, the unobserved storage cost is positive.

For soybean and the other crops, the soybean expected price is the average of November Soybean futures from September 1st to August 31st. We use wheat unobserved cost to represent the expected price of other crops. Wheat expected price is the average of July wheat futures from June 1st to May 31st.

We use two points to calibrate the parameters a_i and b_i in the unobserved marginal cost function. One point is the unobserved cost/current price and end year stock in 2012/13. The other point is the (average unobserved cost)/price and average ending stock from 2001/02 to 2012/13. Thus we find that $a_i = 0.5229$, $b_i = 0.2772$. for soybean $a_i = 0.5488$, $b_i = 0.1593$, for all others $a_i = 0.1799$, $b_i = 0.1286$.

Beginning stock:

The base year is set to be 2013/14, beginning stocks for corn and soybean are 0.0821 10 billion bushels and 0.0141 10 billion bushels. The other stock is the sum of wheat stock and cotton stock. Wheat stock is 718 million bushels and cotton stock is 3.9×480 pounds. Then the total stock is 0.203897 100 million tonnes. The total supply for corn, soybean and all the others are 1.4781 10 billion bushels, 0.3454 10 billion bushels and 0.8550 100 million ton.

Harvest rate:

Harvested acres are less than acres planted. The harvest rate is defined as the ratio of harvested acreage to the planted acreage. The harvest rates for corn, soybean and all other crops are assumed to be 92%, 99% and 85% respectively.

Accuracy Test for Algorithms 1, 2 &3

The purpose of this section is to subject the candidate solutions to an independent and stringent test to compare the quality of the solutions among the three algorithms. GSSA and Smolyak collocation method both approximate the functions using a finite set of points. In the accuracy check, we want to see how the candidate solutions perform for other points in the state space.

Euler equation (EE) error developed in Judd (1992) is used to evaluate the accuracy. The accuracy tests check how far the Euler equations for both storage and acreage decisions deviate zero when using the approximated solution functions.

How to conduct an accuracy test for GSSA. Using simulation to generate points in the state space for the test. Total supplies for three crops in 2013/14 are used as the starting point. Construct another set of crop yields $\{y_{i\tau}\}_{\tau=1,\ldots,T^{test}}, i = 1, 2, 3$ from the joint distribution of three crop yields, with the length of period $T^{test} = 10,000$. Using the solved rules for storage and expected prices to simulate a time series of total supply of all crops for 10,000 periods.

EE error is developed from the Euler equation for storage

$$TS_{i\tau} - s_{i\tau}(TS_{1\tau}, TS_{2\tau}, TS_{3\tau}) = D_i(\delta E_t(p_{i\tau+1}) - SC_{i\tau}) \tag{39}$$

The left hand side of (39) is today's consumption given today's storage decision. The right hand side is what today's consumption would be if the representative storer using storage rule in the next period which determines $E_t(p_{i\tau+1})$. The EE error shows how much the storer deviates from the optimization rule. The Euler Equation error is then defined in a unit free way as shown in (40).

$$EE_{si\tau} = 1 - \frac{D_i(\delta E_t(p_{i\tau+1}) - SC_{i\tau})}{TS_{i\tau} - s_{i\tau}(TS_{1\tau}, TS_{2\tau}, TS_{3\tau})} \tag{40}$$

The subscript s means EE for storage Euler equation, i denotes a specific crop. log10 of EE is used to show the error. To interpret the error, $EE = -1$ means the consumer makes a 1 dollar mistake in consumption when spending 10 dollars. $EE = -4$ means the consumer makes a 1 dollar mistake when spending 10000 dollars.

We can get EE errors for acreage decision Euler equations in the same way. From the first order condition of the farmer's maximization problem, we have

$$x_{i\tau} = \frac{1}{\gamma_i}(\delta E_t(p_{i\tau+1}y_{i\tau+1}) - E_0(p_{i1}y_{i1}) + \overline{\lambda}) + \bar{x}_i - \frac{1}{\gamma_i}\lambda_\tau \tag{41}$$

Where λ_τ is the Lagrangian multiplier with land constraint in time τ. Because the expected revenues are approximated, for any given total supply, the Euler equation holds. Define the unit free Euler equation error for acreage planted as:

$$EE_{xi\tau} = 1 - \frac{1}{x_{i\tau}}(\frac{1}{\gamma_1}(\delta E_t(p_{i\tau+1}y_{i\tau+1}) - E_0(p_{i1}y_{i1}) + \overline{\lambda}) + \bar{x}_i - \frac{1}{\gamma_i}\lambda_\tau) \tag{42}$$

The subscript x means EE for acreage Euler equation. The $log_{10}|EE_H|$ shows the mistake made by making the acreage decision. -1 means the farmer makes 1 acre mistake by planting 10 acres. -4 means the representative profit optimizing farmer makes 1 acre mistake when planting 10,000 acres.

Results

For each period τ, compute $EE_{si\tau}$, $EE_{xi\tau}$, $i = 1, 2, 3$. We evaluate the quality of a candidate solution by computing the maximum and mean of the $EE_{si\tau}$ and $EE_{xi\tau}$ for $\tau \in [1,\ T^{test}]$. The computational time reported below is for an AMD 2.2 GHz server PC running Windows 2008R2 server and Matlab R2014a. Parallel computing with 12 workers are used.

The GSSA with 700 time series and 3rd order polynomial basis functions takes 944 seconds, however the accuracy is the worst among the three for all EE errors. The maximum EE errors for

for corn and soybean are above -2. Because of the low approximation quality, we don't show the EE errors for each of the other Euler equation in details. By changing the basis functions to Chebyshev polynomials or by increasing the basis functions to 4th order polynomials, the maximum Euler equation error across all equations doesn't fall below -2. Thus GSSA is not suitable for solving multi-crop storage model.

Table 1-2 below show both the maximum errors and mean errors across 10,000 periods for sparse grid with storage approximation and sparse grid with expected price approximation. To read the data from the tables below, -2 means the maximum or mean error is 10^{-2}.

Sparse grid with storage rule approximation and Smolyak grid to 3rd order takes 236 seconds. Except for soybean storage, all maximum Euler equation error lies below -3 which mean a maximum 1 dollar error for a 1000 dollars consumption. The maximum error in soybean is $10^{-2.95} \approx 0.0011$ which means a maximum 1.1 dollars mistake in 1000 dollars consumption. The worst approximation in soybean rule is because the storage rule is more nonlinear around low soybean supplies. This can be seen from Figure 1.1, Figure 2.1 and Figure 2.2.

The same approach with 4th level approximation takes 912 seconds and the absolute maximum EE across drops by 0.5 from the 3rd level approximation. The maximum mean value of EE errors across all Euler equations is -4.57.

The computational time for sparse grid with expected price approximation is 3747 seconds. The EE errors for storage Euler equation are smaller than those using method 2. However, the EE errors for acreage decisions are greater than that using method 2. Maximum absolute EE error across all Euler equations is -3.28.

Among all three approaches, Smolyak method with storage rule approximation is the most efficient way to solve multi-crop storage model with convenience yield. Expected price function approximation takes more time but performs better than storage rule approximation with the same number of grid points. This findings are the same as Gouel (2013).

Table 1: Euler equation errors for sparse grid with storage rule approximation

	EE	EE_{sc}	EE_{ss}	EE_{so}	EE_{xc}	EE_{xs}	EE_{xo}
3rd Level	max	-2.95	-2.88	-3.21	-3.77	-3.70	-3.01
	mean	-4.26	-3.82	-4.20	-4.99	-4.72	-4.57
4th Level	max	-3.70	-3.46	-3.64	-4.66	-4.54	-4.07
	mean	-4.99	-4.85	-4.95	-5.60	-5.69	-5.01

Table 2: Euler equation errors for sparse grid with expected price approximation

EE	EE_{sc}	EE_{ss}	EE_{ss}	EE_{xc}	EE_{xs}	EE_{xo}
max	-4.02	-4.04	-3.93	-3.84	-3.82	-3.28
mean	-5.21	-5.22	-4.81	-5.21	-5.20	-4.43

Simulation Results

We use the solution functions obtained from algorithm 2 with 4th level approximation to simulate the model. Various functions of total supply of corn are shown by Figure 1.1-1.4. In each graph, the dotted lines, solid lines and dashed lines represent functions given low supplies, medium supplies and high supplies of soybean and other crops respectively. Low supplies are defined as $TS_s = 2.7$ billion bushels, $TS_o = 70$ million tons. Medium supplies are defined as $TS_s = 3.454$ billion bushels, $TS_o = 85.59$ million tons. High supplies are defined as $TS_s = 4$ billion bushels, $TS_o = 110$ million tons. The medium supplies for all three crops are set to be the real total supplies in 2013/14 marketing year.

Figure 1.1 shows that stock level rises with higher supply level. For higher supplies of both soybean and other crops, the whole storage curve shifts to the right. To see why, when total supplies for soybean and other crops rise in this period, the acreage decisions for other two crops except corn decrease. As land constraint binds, corn acreage must increase and the corn storage decreases at each total supply level because there is less incentive to store as much when expected supply increases.

Both the expected price and the acreage level of corn fall with higher corn supply given fixed soybean and other crops levels as described by Figure 1.2 and 1.3 respectively. The reason is that for a certain total corn supply, higher supplies of soybean and other crops result in lower acreage level for these two. Corn acreage rises in the binding constraint in Figure 1.3 and thus expected corn price decreases as shown by figure 1.2.

The current corn price decreases with an increasing rate as corn supply rises for fixed values of soybean and other crop total supplies as shown in Figure 1.4. Different total supplies of the other two crops shift little of the current prices.

We only show the storage curve for soybean and all the others as their own total supply rises given low medium and high corn and other supplies in Figures 2.1 and Figure 2.2 using dotted lines, solid lines and dashed lines. Low corn supply is defined as $TS_c = 10$ billion bushels, medium corn supply is defined as $TS_c = 14.781$ billion bushels, high corn supply is defined as $TS_o = 18$ billion bushels. Soybean storage curve looks similar to corn's. Considering the scales of stock levels for soybean and corn in Figure 2.1 and 1.1, the soybean storage function is more nonlinear. The storage curve for all other crops is almost linear in the graph, and the curve does not respond much to different supply levels of the other two crops. The more linearity in the other stock curve leads to higher accuracy in storage rule approximation.

Shocks of corn yield.

In this part, we show how decisions and prices respond to a high yield, mean yield and low yield in the second period. High corn yield, medium corn yield and low corn yield are defined as 170 bu/acre, 157 bu/acre and 120 bu/acre. As shown in Table 3, the yield shock only happens in the second period in all three scenarios. Corn yield stays at 157 bu/acre for the other two periods. Yields of soybean and all others are at the mean levels for all periods. Mean yield values of soybean all other crops are 43.8 bushels per acre and 1.18 tons per acre.

Figures 3.1-3.4 describe how storage decisions, expected prices, harvested acres and current prices of corn change with yield shock respectively. Total supply in 2nd period rises with an increase in crop yield. Thus the stock level is highest among the three cases as shown in Figure 3.1. Figure 3.2 says that the expected prices rise and fall in the opposite direction with total supply. With bumper crop, expected corn prices fall. As shown by Figure 3.3 and 3.4, the acreage decisions and current prices perform in the same pattern as expected prices. Changes of corn acreage are driven by expected revenue. When corn yield is low, the expected corn price will be

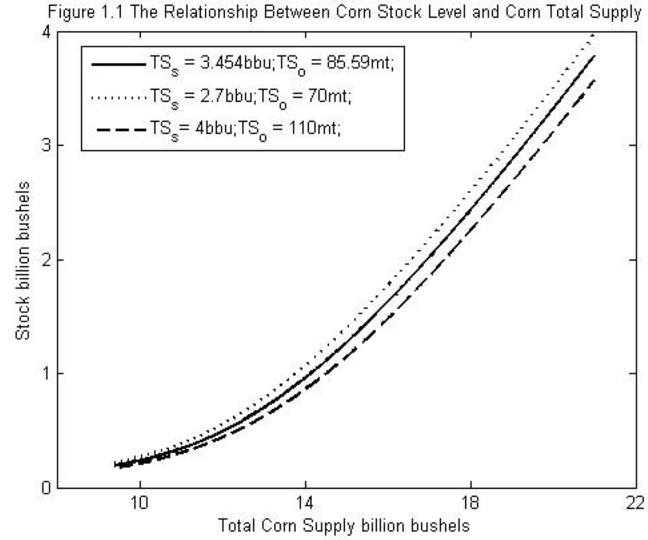


Figure 1.1 The Relationship Between Corn Stock Level and Corn Total Supply

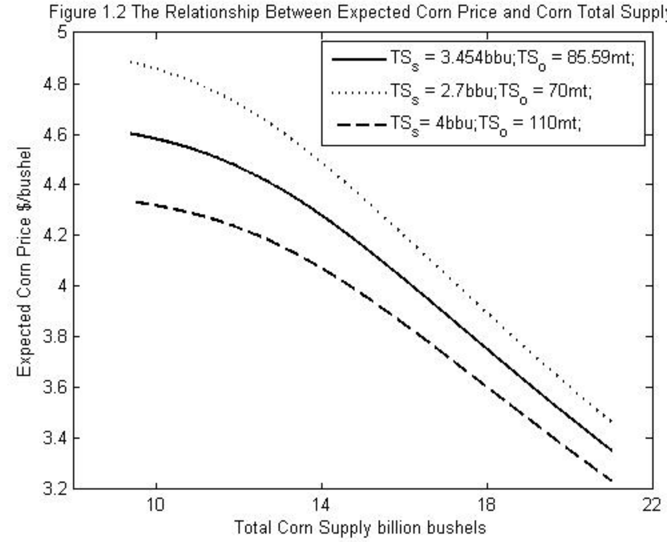


Figure 1.2 The Relationship Between Expected Corn Price and Corn Total Supply

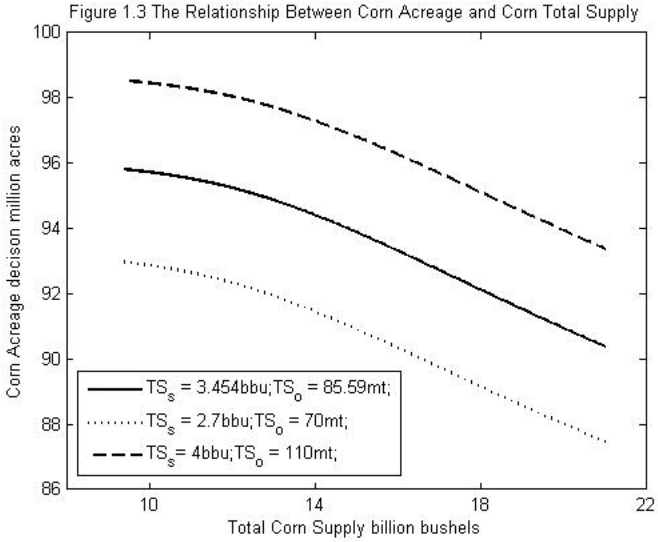


Figure 1.3 The Relationship Between Corn Acreage and Corn Total Supply

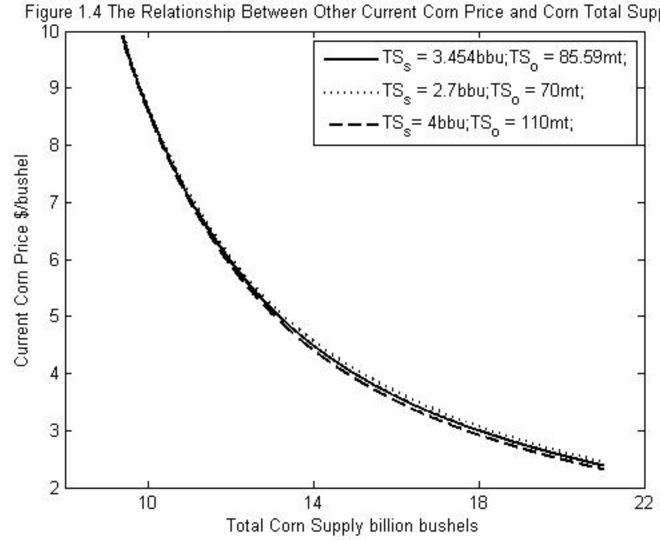


Figure 1.4 The Relationship Between Other Current Corn Price and Corn Total Supply

*bbu: billion bushel, mt: million ton.

Table 3: Corn Yield Shock in 2nd Period

Period	1	2	3
mean	157	157	157
low	157	120	157
high	157	170	157

higher, and the acreage harvested will be higher.

To see the impacts of yield shock for decisions and prices of crops other than corn, we use soybean as an example and the results are shown by Figure 4.1-4.4. Figure 4.1 depicts that all else being equal, if the corn supply is relatively higher, soybean stock levels will be lower. To explain it, considering an increase in corn supply due to the yield shock, the expected revenue decreases and hence the corn acreage for next year decreases. The decrease in the corn acreage will lead to an increase in soybean acreage (Figure 4.3), implying a fall in expected prices of soybean (Figure 4.2) and the level of stocks for both coen and soybean (Figure 4.1). The total supplies of soybean are the same in 2nd period with the same soybean yields and acreage levels. Thus current prices of soybean are only determined by the storage decisions. As we can see from Figure 4.4, higher carryover stock leads to higher current soybean prices while lower carryover stock result in lower current soybean price.

Conclusions

Smolyak collocation methods perform better than GSSA considering computational time and accuracy in solving multi-crop storage model.

The most promising approach for solving an extended model with more than three crops is Smolyak method with storage rule approximation. At the same time, there are also several approaches to improve the current method as described by Judd et. al (2013). For example, using more grid points to those dimensions that are most important for overall quality of approximation. In our case, this means we may put more grid points to descretize soybean total supply space in method 2.

Figure 2.1 The Relationship Between Soybean Stock Level and Soybean Total Supply

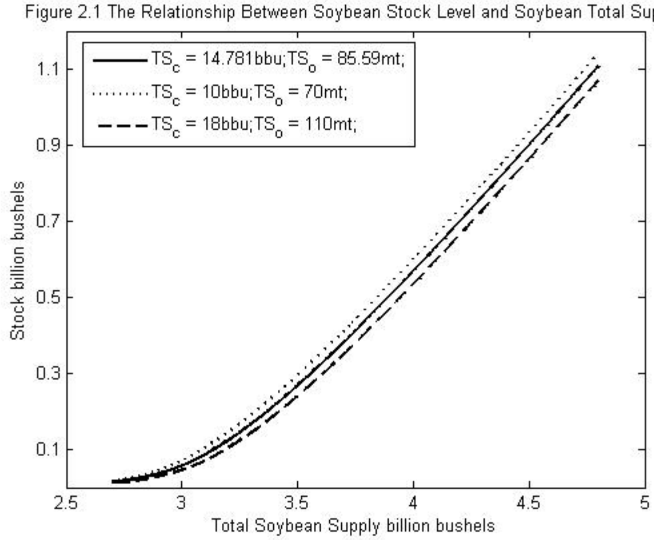


Figure 2.2 The Relationship Between Soybean Stock Level and Soybean Total Supply

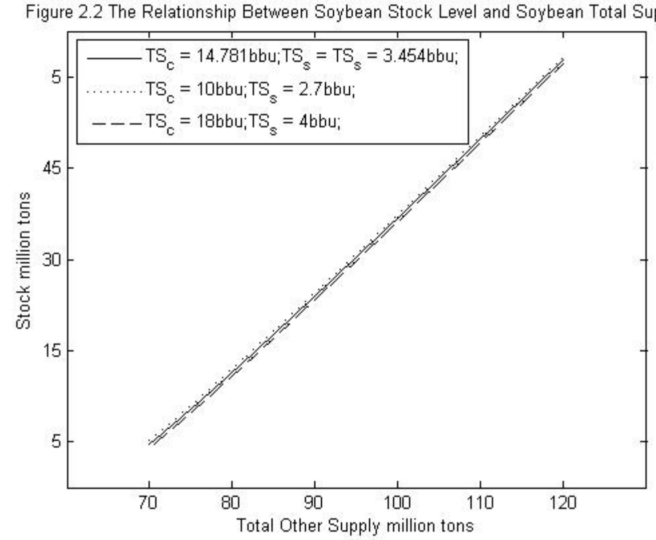


Figure 3.1 Corn Stock Levels with Corn Yield Shock

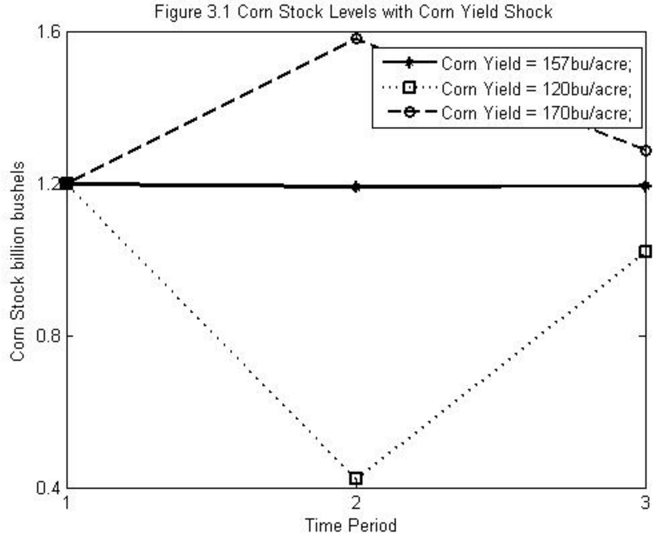


Figure 3.2 Expected Corn Price with Corn Yield Shock

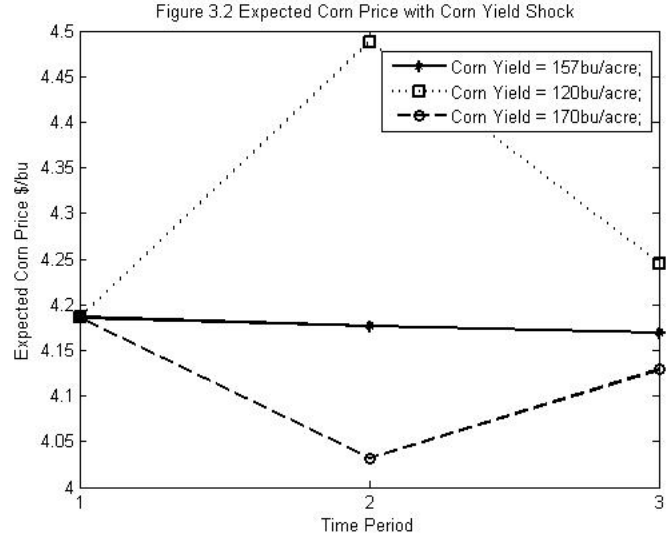


Figure 3.3 Corn Acreage with Corn Yield Shock

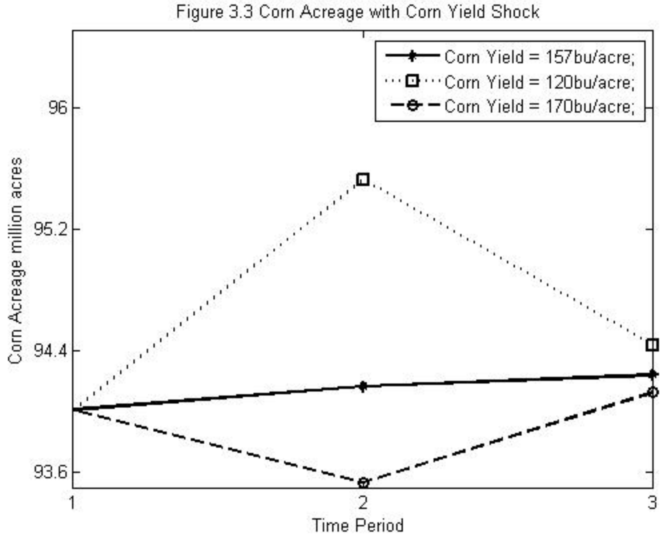


Figure 3.4 Current Corn Price with Corn Yield Shock

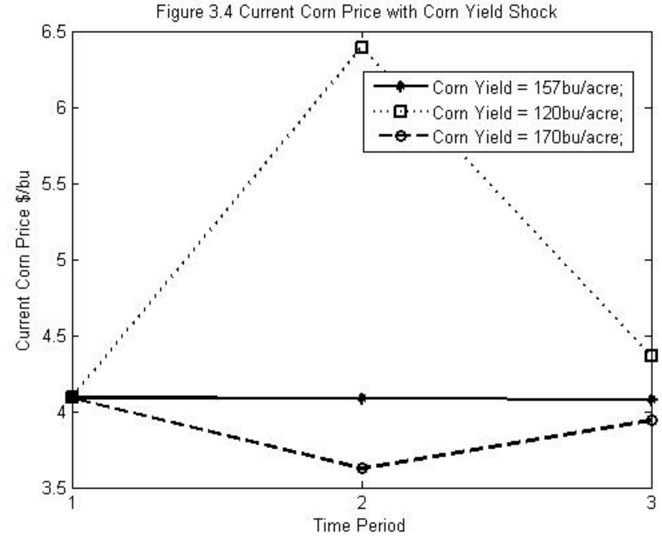


Figure 4.1 Soybean Stock Levels with Corn Yield Shock

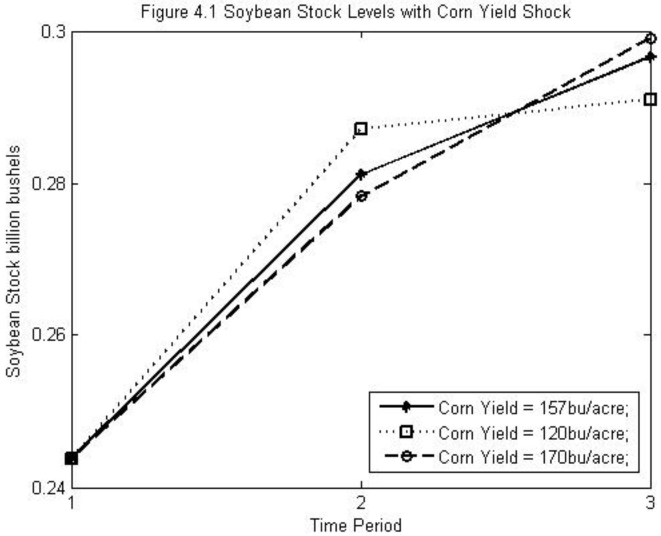


Figure 4.2 Expected Soybean Price with Corn Yield Shock

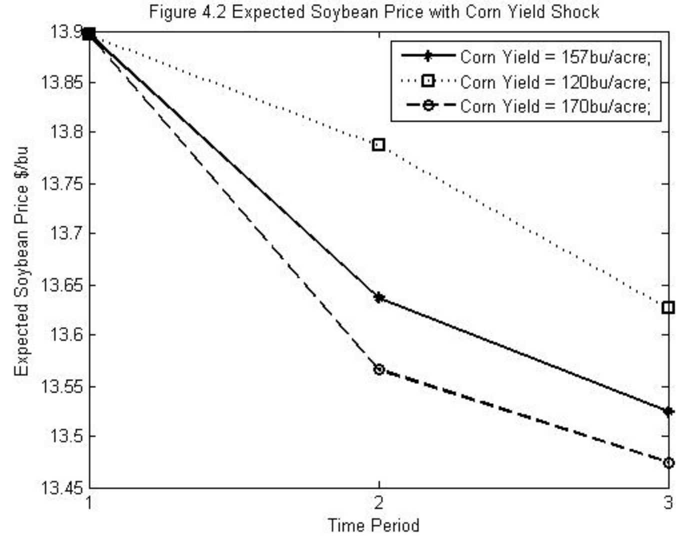


Figure 4.3 Soybean Acreage with Corn Yield Shock

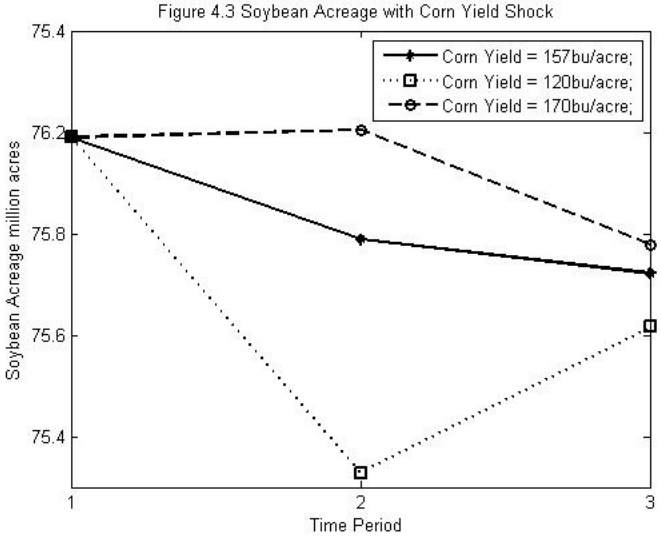
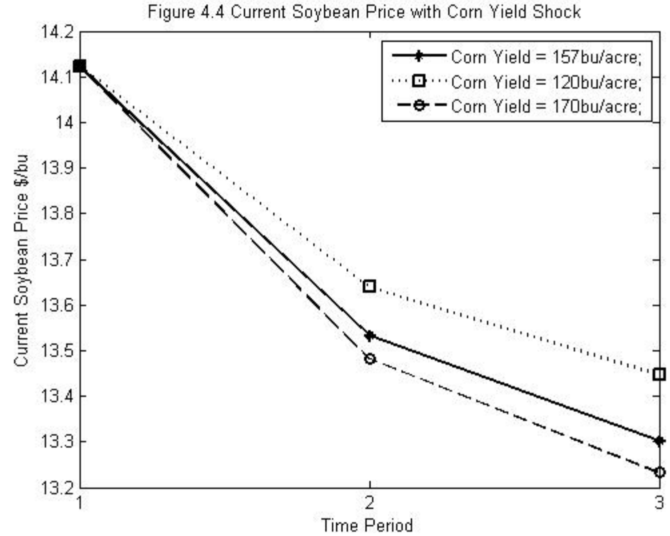


Figure 4.4 Current Soybean Price with Corn Yield Shock



Another condition for making the second method more favorable is the assumption of convenience yield without which there will be stock out conditions and a kink in the storage function. The difficulties of approximating a function with a kink will bring the accuracy for all the second algorithm down to an unacceptable level. In this case, the 3rd algorithm that approximates the smoother expected price functions could be used instead.

References

[1] Adjemian and Smith. (2012) "Using USDA Forecasts to Estimate the Price Flexibility of Demand for Agricultural Commodities." American Journal of Agricultural Economics, 94(4): 978-995.

[2] Arfini, F., Donati, M., & Paris, Q. (2008, August). Innovation in estimation of revenue and cost functions in PMP using FADN information at regional level. In 12th Congress of the EAAE, Belgium, Ghent.Arfini, F., Donati, M., & Paris, Q. (2008, August). Innovation in estimation of revenue and cost functions in PMP using FADN information at regional level. In 12th Congress of the EAAE, Belgium, Ghent.

[3] Arfini, F., & Donati, M. Endogenous prices in mathematical programming models for policy analysis.

[4] Cafiero, C., Bobenrieth H.E. S.A., Bobenrieth H.,J. R.A. and Wright, B. D. 2011. "The empirical relevance of the competitive storage model." Journal of Econometrics, 162(1), 44-54.

[5] Den Haan, W. J., & Marcet, A. (1990). Solving the stochastic growth model by parameterizing expectations. Journal of Business & Economic Statistics, 8(1), 31-34.

[6] Gohin, A., & Chantreuil, F. (1999). La programmation mathématique positive dans les modèles d'exploitation agricole: principes et importance du calibrage. Cahiers d'économie et sociologie rurales, (52), 59-78.

[7] Gouel, C. (2013a). Optimal food price stabilisation policy. European Economic Review, 57, 118-134.

[8] Gouel, C. (2013b). Comparing numerical methods for solving the competitive storage model. Computational Economics, 41(2), 267-295.

[9] Hasanhodzic, J., & Kotlikoff, L. J. (2013). Generational risk–is it a big deal?: Simulating an 80-period olg model with aggregate shocks (No. w19179). National Bureau of Economic Research.

[10] Heckelei, T., & Britz, W. (2005). Models based on positive mathematical programming: state of the art and further extensions. Modelling Agricultural Policies: State of the Art and New Challenges, Parma, Italy, 48-73.

[11] Howitt, R. E. (1995). Positive mathematical programming. American journal of agricultural economics, 77(2), 329-342.

[12] Jansson, T., & Heckelei, T. (2011). Estimating a primal model of regional crop supply in the European Union. Journal of Agricultural Economics, 62(1), 137-152.

[13] Joseph, K., Irwin, S. H., & Garcia, P. (2011, May). Commodity Storage under Backwardation: Does the Working Curve Still Work?. In Unpublished paper presented at Agricultural and Applied Economics Association's 2011 AAEA and NAREA Joint Annual Meeting. Pittsburgh, Pennsylvania, USA.

[14] Jin, N., Lence, S., Hart, C., & Hayes, D. (2012). The long-term structure of commodity futures. American Journal of Agricultural Economics, 94(3), 718-735.

[15] Judd, K. L. (1992). Projection methods for solving aggregate growth models. Journal of Economic Theory, 58(2), 410-452.

[16] Judd, K., L. Maliar, S. Maliar. (2011). "Numerically stable and accurate stochastic simulation approaches for solving dynamic economic models." Quantitative Economics, 2, 173-210.

[17] Judd, K. L., Maliar, L., Maliar, S., & Valero, R. (2013). Smolyak method for solving dynamic economic models: Lagrange interpolation, anisotropic grid and adaptive domain (No. w19326). National Bureau of Economic Research.

[18] Kaldor, N. (1939). Speculation and economic stability. The Review of Economic Studies, 7(1), 1-27.

[19] Klimke, W. A. (2006). Uncertainty modeling using fuzzy arithmetic and sparse grids. Shaker.

[20] Klimke, W. A. (2007). Sparse Grid Interpolation Toolbox–user's guide. IANS report, 17.

[24] Kollmann, R., Maliar, S., Malin, B. A., & Pichler, P. (2011). Comparison of solutions to the multi-country real business cycle model. Journal of Economic Dynamics and Control, 35(2), 186-202.

[21] Krueger, D., & Kubler, F. (2004). Computing equilibrium in OLG models with stochastic production. Journal of Economic Dynamics and Control, 28(7), 1411-1436.

[22] Malin, B., Krueger, D., & Kubler, F. (2007). Computing stochastic dynamic economic models with a large number of state variables: a description and application of a Smolyak-collocation method (No. w13517). National Bureau of Economic Research.

[23] Malin, B. A., Krueger, D., & Kubler, F. (2011). Solving the multi-country real business cycle model using a Smolyak-collocation method. Journal of Economic Dynamics and Control, 35(2), 229-239.

[24] Mérel, P., S. Bucaram. 2010. "Exact calibration of programming models of agricultural supply against exogenous supply elasticities." European Review of Agricultural Economics, 37(3), pp. 395-418.

[25] Mérel, P., Simon, L. K., & Yi, F. (2011). A fully calibrated generalized constant-elasticity-of-substitution programming model of agricultural supply. American Journal of Agricultural Economics, 93(4), 936-948.

[26] Miranda, M. J. (1997). Numerical strategies for solving the nonlinear rational expectations commodity market model. Computational Economics, 11(1-2), 71-87.

[27] Miranda, M. J., & Glauber, J. W. (1993). Estimation of dynamic nonlinear rational expectations models of primary commodity markets with private and government stockholding. The Review of Economics and Statistics, 75(3), 463-70.

[28] Paul, A. B. (1970). The pricing of binspace—a contribution to the theory of storage. American Journal of Agricultural Economics, 52(1), 1-12.

[29] Peterson, H. H., Tomek, W. G., 2005. "How much of commodity price behavior can a rational expectations storage model explain?" Agricultural Economics, 33, 289-303.

[30] Roberts, M. J., and W. Schlenker. 2013. "Identifying Supply and Demand Elasticities of Agricultural Commodities: Implications for the US Ethanol Mandate." American Economic Review, 103(6): 2265-95.

[31] Roberts, M. J., & Tran, A. N. (2012, May). Commodity Price Adjustment in a Competitive Storage Model with an Application to the US Biofuel Policies. In 2012 Annual Meeting, August 12-14, 2012, Seattle, Washington (No. 124869). Agricultural and Applied Economics Association.

[32] Roberts, M. J., & Tran, A. N. (2013). Conditional Suspension of the US Ethanol Mandate using Threshold Price inside a Competitive Storage Model. In 2013 Annual Meeting, August 4-6, 2013, Washington, DC (No. 150717). Agricultural and Applied Economics Association.

[33] Rui, X., & Miranda, M. J. (1995). Commodity storage and interseasonal price dynamics. In NCR-134 Conference Proceedings (pp. 44-53).

[34] Smolyak, S. A. (1963). Quadrature and interpolation formulas for tensor products of certain classes of functions. In Dokl. Akad. Nauk SSSR (Vol. 4, No. 240-243, p. 123).

[35] Thompson, W., S. Meyer, and P. Westhoff. (2008). "Mandate economics applied to US biofuel policies." Paper presented at the Annual Meeting of the International Agricultural Trade Research Consortium held at Scottsdale, Arizona, on December 7-9.

[36] Villena-Roldán, B. (2013). Hiring Selectivity and Business Cycle Dynamics in the Labor Market.

[37] Williams, J. C., & Wright, B. D. (2005). Storage and commodity markets. Cambridge University Press.