Economic factors that influence soybean and canola prices

Lei Cui
ECONOMIC FACTORS THAT INFLUENCE
SOYBEAN AND CANOLA PRICES

By
Lei Cui
B.S. Beijing Normal University, P.R.China, 1997

A THESIS
Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science
(in Resource Economics and Policy)

The Graduate School
The University of Maine
August, 2001

Advisory Committee:
James D. Leiby, Associate Professor of Resource Economics and Policy, Advisor
Mario F. Teisl, Assistant Professor of Resource Economics and Policy
Kathleen P. Bell, Assistant Professor of Resource Economics and Policy
The USDA Agricultural Research Service is examining the feasibility and
profitability of growing Canola and soybeans in potato rotation systems. The study
described in this thesis is part of this research program. The primary objective of this
research is to look for economic factors that influence soybean and Canola prices.

Canola is a new oilseed crop to the U.S. Since the Food and Drug Administration
approved its use as edible food, Canola production in the U.S. has increased
tremendously. Because only 11 years of data are available on Canola consumption and
production in the U.S., it is difficult to empirically analyze Canola prices. Fortunately, we
find that Canola and soybean prices are highly correlated. As a result, if we can explain
the determinants of soybean prices, we can discover information about the determinants
of Canola prices. In this study, we concentrate on soybean price movements, make
inferences about Canola prices.

We establish a simultaneous model of the U.S. soybean market and study factors
affecting soybean prices within this economic structure. Our model is based primarily on
the USDA CROPS Model developed by Houck, et al.
Our results indicate that soybean price is positively affected by a time trend variable, expected wholesale price of corn oil, expected real expenditures spent on food, expected variable cost of growing soybeans, and one-year lagged farm-level corn price, but negatively affected by one-year lagged soybean price, one-year lagged wheat price, and one-year lagged acreage of soybeans. Using the Canola and soybean price relationship, we infer how these economic factors affect Canola price. With the reduced form, we can forecast the future price of both Canola and soybeans.
ACKNOWLEDGMENTS

I would like to thank Dr. Leiby, Dr. Teisl and Dr. Bell for being on this thesis committee. I would especially like to thank my advisor James Leiby, for his remarkable patience and countless hours spent during the process of conducting researches, revising models and writing this thesis.

I would also like to thank my friends, both in China and in the U.S., who encouraged me during the last two years. I especially thank my husband, Hanxing Yu, for his love, understanding and support all the time. Finally, I would like to thank my parents. Their care and love are the sources of my courage.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>ix</td>
</tr>
</tbody>
</table>

Chapter

1. **INTRODUCTION** ................................................................. 1
   - Statement of Problem ....................................................... 1
   - Background ........................................................................ 5
   - Summary ........................................................................... 8

2. **USDA CROPS MODEL FOR THE U.S. SOYBEAN MARKET** ................. 9
   - The Demand Side for Soybeans in the CROPS Model .................. 10
     - Aggregated Demand for Soybeans in the U.S. ..................... 10
     - Variable Definitions .................................................... 11
     - Simultaneous Equations for the Demand Side of Soybean Market ...... 11
   - The Supply Side for Soybeans in the CROPS Model .................. 16
   - Joining Demand and Supply Sides Together .......................... 18
   - Some Criticisms Regarding the CROPS Model ....................... 19
   - Summary ........................................................................... 22
5. **SUMMARY AND CONCLUSIONS** ........................................................... 50

   **Summary** .......................................................................................... 50
   **Conclusions** ...................................................................................... 51

**REFERENCES** .......................................................................................... 52

**APPENDICES** .......................................................................................... 54

   **Appendix A.** Formulation of Distributed Lag Models ..................... 55

   **Appendix B.** Identification in Simultaneous System and
   Two-Stage-Least-Squares Estimates ................................................. 65

   **Appendix C.** Comparison of OLS Estimates with
   Two-Stage-Least-Squares Estimates in the
   Simultaneous Equation Model for the
   Soybean Market ................................................................. 68

   **Appendix D.** Procedure of Obtaining the Reduced Form for
   Farm-Level Soybean Prices (PS) ................................................. 72

**BIOGRAPHY OF THE AUTHOR** ......................................................... 74
LIST OF TABLES

Table 1.1. Regression Results Comparing Soybean and Canola Prices .................. 4

Table 2.1. Variable Definitions in The CROPS Model........................................ 12

Table 3.1. Summary Statistics for Variables Used in the Equation of Soybean Export Demand (Equation 3.1) .................................................. 25

Table 3.2. Summary Statistics for Variables Used in the Equation of Soy Meal Demand (Equation 3.2) ................................................................. 26

Table 3.3. Summary Statistics for Variables Used in the Equation of Soy Meal Export Demand (Equation 3.3) ......................................................... 27

Table 3.4. Summary Statistics for Variables Used in the Equation of Domestic Soy Oil Demand (Equation 3.4) ......................................................... 28

Table 3.5. Summary Statistics for Variables Used in Equation of Soy Oil Export Demand (Equation 3.5) ................................................................. 28

Table 3.6. Summary Statistics for Variables Used in Acreage Response Equation (Equation 3.6) ................................................................. 33

Table 3.7. Summary Statistics for Variables Used in Identity Equations (Equation 3.10 through 3.12) ......................................................... 36

Table 4.1. Parameter Estimates for Equation (4.1)
- U.S. Soy Meal Demand, QMD......................................................... 39

Table 4.2. Parameter Estimates for Equation (4.2)
- U.S. Soy Oil Demand, QOD......................................................... 39
Table 4.3 Parameter Estimates for Equation (4.3)
- U.S. Soybean Export Demand, QSX ........................................ 39

Table 4.4 Parameter Estimates for Equation (4.4)
- U.S. Soy Meal Exports, QMX ..................................................... 40

Table 4.5 Parameter Estimates for Equation (4.5)
- U.S. Soy Oil Exports, QOX ..................................................... 40

Table 4.6 Parameter Estimates for Equation (4.6)
Soy Meal Produced (QMP) Regressed on Soybeans Crushed (QSC)
(with no intercept) ................................................................. 40

Table 4.7 Parameter Estimates for Equation (4.7)
Soy Oil Produced (QOP) Regressed on Soybeans Crushed (QSC)
(with no intercept) ................................................................. 40

Table 4.8 Parameter Estimates for Equation (4.8) – Wholesale Soy Meal Price (PM)
And Wholesale Soy Oil Price (PO) Regressed on Farm-level
Soybean Price (PS) ................................................................. 40

Table 4.9 Parameter Estimate for Equation (4.12)
- Soybean Acreage Response Model ..................................... 41

Table 4.10 Elasticity of Soy Oil Domestic Demand
Computed at Mean Levels ................................................... 42

Table 4.11 Elasticity of Soy Oil Export Demand Computed at Mean Levels.......... 42

Table 4.12 Elasticity Soybean Acreage Computed at Mean Levels .................... 43

Table 4.13 Soybean Farm Level Price With Respect to
Some of the Right-hand-side Variables ..................................... 44
Table A.1. Summary Statistics for Soybean Prices from 1965 to 1998

(Prices are adjusted with CPI, 1999=100) .................................................. 62

Table A.2. Parameter Estimates of Polynomial Distributed Lag Structures

For Soybeans, Soy Oil and Meal

(numbers in parentheses are standard errors) ............................................. 62

Table C.1. Compare OLS Estimates and Two-Stage-Least-Squares Estimates

for Equations in the Simultaneous System In Our Model

(numbers in parentheses are standard errors) ............................................. 71
LIST OF FIGURES

[Figure 1.1] U.S. Average Annual Farm Level Canola Seed Price and Soybean Price
From 1991 to 1999 ................................................................. 2

[Figure 1.2] U.S. Average Annual Canola Oil Price and Soy Oil Price
From 1989 to 1999 ................................................................. 2

[Figure 1.3] U.S. Average Annual Canola Meal Price and Soy Meal Price
From 1989 to 1999 ................................................................. 3

[Figure 1.4] U.S. Canola Production
From 1991 to 2000 ................................................................. 6

[Figure 1.5] U.S. Soybean Production and Exports
From 1970 to 1998 ................................................................. 7

[Figure 3.1] South American Soybean Export Share in the International Soybean Market (1995 – 1999) .................................................. 25

[Figure 4.1] Forecast Farm-Level Soybean Price from 1991 to 1998 ....................... 47

[Figure 4.2] Forecast Canola Price from 1991 to 1998 ................................. 48

[Figure A.1] Forecast Real Soybean Prices from 1990 to 1998
With Polynomial Distributed Lag Structure .................................. 64
In order to identify more profitable potato rotations, the Agricultural Research Service of the USDA is conducting a study of the feasibility of growing soybeans and Canola in a potato rotation system. This thesis is part of that project. Here we attempt to determine what economic factors influence the price of soybeans and Canola and then explain how soybean and Canola prices move with respect to these factors.

Canola is a relatively new oilseed crop in the United States. It was not approved for food use in the U.S. until 1985. Current price data for Canola available from USDA are annual prices from 1989 to 2000, insufficient for a rigorous empirical analysis. Fortunately, we find that the prices of Canola and soybeans move quite similarly. This is displayed in Figure 1.1, Figure 1.2 and Figure 1.3.
Figure 1.1: U.S. Average Annual Farm Level Canola Seed Price and Soybean Price From 1991 to 1999

(Source: USDA Agricultural Statistics, various year)
(Prices are adjusted with Consumer Price Index. 1999=100)

Figure 1.2: U.S. Average Annual Canola Oil Price and Soy Oil Price From 1989 to 1999

(Source: USDA Agricultural Statistics, various years, prices are adjusted with Consumer Price Index. 1999=100)
Figure 1.3: U.S. Average Annual Canola Meal and Soy Meal Prices From 1989 to 1999

We expect Canola and soybeans to be very close substitutes on the demand side since both are crushed to produce oil for the human food market, and meal for the animal feeding market. Similarly, they are competing products on the supply side since they are grown during the same season, in the same regions, and use similar factors of production. Given these commonalities, it is reasonable to expect their prices to move together very closely. Evidence of this is provided by the analysis presented in Table 1.1.

Simple regressions of Canola seed price ($P_{sc}$) onto soybeans price ($P_{s}$), Canola meal price ($P_{cm}$) onto soy meal price ($P_{m}$), and Canola oil price ($P_{co}$) onto soybean oil price ($P_{o}$) yield the results presented in Table 1.1. Note that all the parameter estimates are significantly different from 1 given a 95% significance level.
The results of model one support the hypothesis that the prices of Canola and soybeans move approximately together, and that on average, soybean prices exceed Canola seed price by about 3%. Model two suggest that Canola oil price is about 15% higher than that of soy oil. The result is not surprising because of the different quality of Canola oil and soy oil. Canola and soy oil, as oil products, are mostly consumed by humans. Canola oil is labeled as healthy oil due to the fact that Canola oil has the highest level of monounsaturated fatty acid, oleic acid, which helps reduce serum cholesterol level and LDL cholesterol levels. Canola oil has a saturated fat of 6%, compared to 15% for soybean oil and 89% for palm oil. Canola oil would then be somewhat preferred to soy oil in the cooking oil market. Model three suggest that soy meal price is about 34% higher than that of Canola meal price.

Both Canola meal and soy meal are high-protein livestock feeds or feed supplements. Soy meal has a protein content of about 45% while Canola has about 38% protein. Because of greater protein, soy meal is favored to Canola meal. This is consistent with our result that soy meal price is higher than Canola meal price.

<table>
<thead>
<tr>
<th>Left-Hand-Side Variable</th>
<th>Right-Hand-Side Variable</th>
<th>Coefficient (Standard Error)</th>
<th>Adjusted R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price of Canola Seed (Pcs)</td>
<td>Price of Soybeans (PS)</td>
<td>0.97 (0.01)</td>
<td>0.9970</td>
</tr>
<tr>
<td>Price of Canola Oil (Pco)</td>
<td>Price of Soy Oil (PO)</td>
<td>1.15 (0.01)</td>
<td>0.9982</td>
</tr>
<tr>
<td>Price of Canola Meal (Pcm)</td>
<td>Price of Soy Meal (PM)</td>
<td>0.66 (0.01)</td>
<td>0.9917</td>
</tr>
</tbody>
</table>

All models use real prices per metric ton (adjusted with Consumer Price Index, 1999=100). The number of observations is 11 for all models.
From the discussions above, we see that Canola and soybeans have similar processing, end-use purposes, and are substitutable for each other in the oil and meal markets. So it is likely that their prices are highly correlated. Thus we can understand how one price moves by examining the other. In this way, we can transform our problem of examining Canola price movements to examining soybean price movements. Because data for U.S. Canola are from a very short time series, we will, in this thesis, concentrate on the soy market. From the soy market we can reasonably infer the price movements of Canola.

**Background**

Canola is a genetic variety of rapeseed developed by Canadian plant breeders through traditional plant breeding techniques, specifically for its nutritional qualities. In 1974, Dr. Baldur Stefansson, a University of Manitoba plant breeder, developed the first ‘double low’ variety which reduced both erucic and glucosinolate levels. This Brassica napus variety, Tower, was the first variety to meet the specific quality requirements used to identify a greatly improved crop known as Canola (“The Origin of Canola”, Canola Council, June 24 2001). For rapeseed to be accepted as Canola, it must contain less than two percent erucic acid in the oil and less than 30 micromoles per gram of glucosinolate in the meal. The reasons for this are that high level of erucic, in animal studies, was suspected to cause heart lesions and fat build up around hearts, while high level of glucosinolate, which is sour and bitter tasting, disqualify rapeseed meal as livestock feed.
The United States has relatively a short history of growing and consuming Canola. But since the approval of Canola being used in edible products by the Food and Drug Administration in 1985, domestic production has increased greatly. The implementation of the 1990 Farm Bill also contributed to the increase of Canola growing (Lordkipanidze et al. and Ames et al.). Two aspects of its legislation: planting flexibility and oilseeds marketing provisions encouraged farmers to expand acreage of Canola. In turn, the production of Canola has increased tremendously. According to data from USDA, U.S. Canola production increased by over 950% from 1991 to 2000. (See Figure 1.4)

**Figure 1.4: U.S. Canola Production from 1991 to 2000**

The soybean is often called the miracle crop. It is the world’s foremost provider of protein and oil. Soybean cultivation was first recorded in 2828 B.C. in China (Jordan, Houck, et al.). Soybeans and their products have been important sources of protein for millions of Chinese and other Oriental people for nearly 5000 years.
The Soybean was first introduced into the United States in 1804 (Jordan et al.), primarily for use as a forage crop. In 1921, the growing American soybean industry was provided tariff protection. Since the 1950’s, the U.S. has become the world’s largest soybean producer and exporter. Figure 1.5 shows the production and exports of U.S. soybeans from 1970 to 1998.

**Figure 1.5: U.S. Soybean Production and Exports from 1970 to 1998**

(Source: USDA Agricultural Statistics various years)
Summary

This chapter provides background information on Canola and soybean production in the United States and stated our objective of studying Canola and soybeans prices. Because the U.S. has relatively a short history of growing Canola, it is difficult to perform a rigorous empirical analysis of Canola prices. Fortunately, we find that Canola and soybean prices are highly correlated. By studying soybean prices, we can also learn about Canola prices. Chapter Two provides a model of soybean prices, developed in the 1960’s. Chapter Three provides our updated model of these prices, based on the information in Chapter Two.
Chapter 2

USDA CROPS MODEL FOR U.S. SOYBEAN MARKET

Preliminary empirical estimates based on pure time series analysis do not provide insight into soybean future prices and cannot pick up the turning points of soybean prices when used for forecasts, although the pure time series structure itself fits the data well (See Appendix A). We will study soybean prices within economic structures. In this Chapter, we first review the work of Houck, Ryan, and Subotnik in their study of a simultaneous equations model for the U.S soybean market. Some criticisms to their model will be raised, which make necessary significant modifications to the model.

Multi-equation econometric models have been used by a number of researchers to analyze the structure and operation of the soybean market. These models have grown in complexity, as their components have become more representative of the total soybean market. The dynamic supply and demand model of the U.S. Soybean Market is one such model. Houck, Ryan and Subotnik (1971) presented a multi-equation model of soybean prices that for the first time took both supply and demand into consideration. In their study, the meshing of both supply and demand relationships is undertaken with special attention to policy variables. Their work is frequently referred to as the USDA CROPS Model (Jordan, et al.). The CROPS model is composed of two “blocks”, which are constituted of the behavioral and technical relationships on the demand supply sides of soybeans.
The Demand Side for Soybeans in the CROPS Model

Aggregated Demand for Soybeans in the U.S.

Meal and oil are joint products from the crushing of soybeans. The ratio of soy oil and meal crushed from soybeans with respect to soybeans crushed is relatively fixed. Thus meal and oil supplies are tightly linked to each other and to the quantities of soybeans crushed.

Soybeans, meal and oil have multiple uses: domestic (crush), export, and inventory. Thus, multiple-market outlets are available for all three products. The three are interdependent with larger economic sectors. Once crushed, the meal and oil components enter market channels that are essentially independent of each other. Each of these is part of a complex economic sector in which competition and substitution among commodities are important. Soy meal is one of several high-protein feed products for the livestock sector. Soy oil is one of many edible vegetable oils in the fats and oils complex. Soybeans are a specific oilseed in a worldwide network of competing oil-bearing products. Prices and product flows of soybeans, meal and oil are determined simultaneously because of the joint-product relationship.

It is indicated in the CROPS model that the total demand of soybeans at the farm level is an aggregated demand of U.S. crushing demand, export demand, and other demands\(^1\). The U.S. soybean crushing demand is a summation of the total wholesale soy meal and soy oil demand. The export soybean demand is reflected by foreign soybean demand, soy oil demand and soy meal demand. The foreign soybean oil demand is

\(^{1}\) Other demands include government purchases of soybeans and demand for stocks of soybeans.
expressed in two parts: P.L. 480 concessional sales and commercial exports through normal trade channels.

**Variable Definitions**

Houck (et al.) formulated a thirteen-equation model for the demand side of soybeans in the U.S. These equations (2.1) through (2.13)) are shown below and discussions of the variables chosen are also presented. Table 2.1 provides variable definitions.

**Simultaneous Equations for the Demand Side of Soybean Market**

Each equation (From equation (2.1) to (2.13)) will be discussed in detail.

\[ Q_{M,D,t} = f_1(P_{M,t}, Q_{P,t}, L_{V,t}, Q_{D,P,t}, e_t) \]  

(2.1)

Equation (3.1) is the domestic soy meal demand. Since the feed outlet overwhelmingly dominates the U.S. soy meal market, the domestic soy meal demand is a total of several derived demands that reflect the variables having a major impact on soybean-meal demand originating in the feed-livestock sector. The quantity of soy meal demanded (QMD) is expressed jointly in a function with the wholesale meal price (PM) and several other predetermined variables. The quantity supplied of other high-protein feeds (QP) is included to capture the substitution effects of other high protein feeds like cottonseed meal, linseed meal, tankage, and meat scraps (Houck, et al.). The livestock production units\(^2\) of hogs, cattle and poultry (LV) is included as these livestock are consumers of soy meal. Their influence in this equation is analogous to the population effect in a primary demand equation (Houck, et al.). The variable QOP represents the estimated percentage of digestible protein in concentrate rations for livestock and poultry (QDP). QOP is an

---

\(^2\) A livestock production unit is approximately 1000 pounds of animal live weight.
indicator of the continuing change in the feeding practices toward higher protein feed sources.

Table 2.1: Variable Definitions in The CROPS Model

<table>
<thead>
<tr>
<th>Variables</th>
<th>Definitions</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM</td>
<td>Average U.S. wholesale price of soy meal</td>
<td>Cents/pound</td>
</tr>
<tr>
<td>PO</td>
<td>Average U.S. price of soy oil</td>
<td>Cents/pound</td>
</tr>
<tr>
<td>PS</td>
<td>Average price received by U.S. farmers for soybeans</td>
<td>$/bushel</td>
</tr>
<tr>
<td>QMD</td>
<td>Quantity of soy meal demanded in the U.S.</td>
<td>Million pounds</td>
</tr>
<tr>
<td>QMP</td>
<td>Quantity of soy meal produced in U.S.</td>
<td>Million pounds</td>
</tr>
<tr>
<td>QMX</td>
<td>Quantity of soy meal exported from U.S.</td>
<td>Million pounds</td>
</tr>
<tr>
<td>QOD</td>
<td>Quantity of soy oil demanded in U.S.</td>
<td>Million pounds</td>
</tr>
<tr>
<td>QOX</td>
<td>Quantity of soy oil commercially exported from U.S.</td>
<td>Million pounds</td>
</tr>
<tr>
<td>QSC</td>
<td>Quantity of soybeans crushed in U.S.</td>
<td>Thousand bushels</td>
</tr>
<tr>
<td>QSX</td>
<td>Quantity of soybeans exported from U.S.</td>
<td>Thousand bushels</td>
</tr>
<tr>
<td>SO</td>
<td>Crop year ending stocks of soybean oil in the U.S.</td>
<td>Million pounds</td>
</tr>
<tr>
<td>SS</td>
<td>Crop year ending stocks of soybeans in U.S. held privately</td>
<td>Million pounds</td>
</tr>
<tr>
<td>CT</td>
<td>Cumulative trend: 1946=1, 1947 = 3, 1948 =6, 1949=10</td>
<td>Million dollars</td>
</tr>
<tr>
<td>DV</td>
<td>Dummy variable: for 1952, DV=1, others, DV=0</td>
<td>Million dollars</td>
</tr>
<tr>
<td>E</td>
<td>Real Expenditures on food in the U.S.</td>
<td>Million dollars</td>
</tr>
<tr>
<td>F</td>
<td>Production of Feed grains in countries importing U.S. Soybeans and soy meal</td>
<td>Thousand metric tons</td>
</tr>
<tr>
<td>GS</td>
<td>Change in stocks of soybeans owned by CCC</td>
<td>Thousand bushels</td>
</tr>
<tr>
<td>I</td>
<td>Index of national income in countries importing U.S. soybeans and soybean product, 1950 = 100</td>
<td>Thousand bushels</td>
</tr>
<tr>
<td>LV</td>
<td>Livestock production</td>
<td>Thousands</td>
</tr>
<tr>
<td>LW</td>
<td>Livestock units in countries importing soy meal from U.S.</td>
<td>Thousands</td>
</tr>
<tr>
<td>OM</td>
<td>World imports of oilseed meal less U.S. exports of soy meal</td>
<td>Thousand metric tons</td>
</tr>
<tr>
<td>FBL</td>
<td>Wholesale price index of butter and lard in U.S.</td>
<td>1957=100</td>
</tr>
<tr>
<td>PG</td>
<td>Price of groundnut oil</td>
<td>Cents/pound</td>
</tr>
<tr>
<td>PLM</td>
<td>Price of linseed meal imported by U.K.</td>
<td>Cents/pound</td>
</tr>
<tr>
<td>QAO</td>
<td>Selected alternative oil supplies in world trade (groundnut, cottonseed, and sunflower-seed oils)</td>
<td>Million pounds</td>
</tr>
<tr>
<td>QCOD</td>
<td>Quantity of cottonseed oil demanded in U.S.</td>
<td>Million pounds</td>
</tr>
<tr>
<td>QCXPL</td>
<td>P.L.480 shipments of cottonseed oil</td>
<td>Million pounds</td>
</tr>
<tr>
<td>QDP</td>
<td>Percentage of digestible protein in concentrates for all U.S. livestock and poultry</td>
<td>Million pounds</td>
</tr>
<tr>
<td>QOOD</td>
<td>Olive-oil production in Mediterranean countries</td>
<td>Million pounds</td>
</tr>
<tr>
<td>QOXPL</td>
<td>P.L.480 shipments of soybean oil</td>
<td>Million pounds</td>
</tr>
<tr>
<td>QP</td>
<td>Production of high-protein feeds other than soy meal in U.S., in soybean-meal equivalent</td>
<td>Thousand short tons</td>
</tr>
<tr>
<td>QSS</td>
<td>Commercial supply of U.S. soybeans</td>
<td>Thousand bushels</td>
</tr>
<tr>
<td>QTXPL</td>
<td>Total P.L.480 exports of vegetable oil (QOXPL+QCXPL)</td>
<td>Million pounds</td>
</tr>
<tr>
<td>SM</td>
<td>Changes in total U.S. stocks of soy meal</td>
<td>Million pounds</td>
</tr>
<tr>
<td>W</td>
<td>Crushing and handling spread</td>
<td>$/bushel</td>
</tr>
</tbody>
</table>
Equation (2.2) represents the domestic demand for soybean oil. This is also a combination of several derived demand functions. The quantity of soybean oil demanded (QOD) is expressed jointly in a function with the price of crude soy oil (PO) and several other predetermined variables. The domestic utilization of cottonseed oil (QCOD) is included to account for the substitution effects. A wholesale price index of butter and lard in the U.S. (PBL) (1957 is used as base year, 1957=100) is included to account for the influence of animal fats and oils on soy oil demand (Houck, et al.). The total yearly-deflated expenditure (E) on food in the U.S. is included to account for the changes in both population and individual incomes (Houck, et al.).

Equation (2.3), the foreign demand for soybeans (QSX) faced by exporters is the sum of the derived demands for soybean-using products by foreign buyers. The major source of demand for soybeans in world markets is for crushing. As foreign buyers can substitute bean purchases for purchases of soy meal and/or soy oil, the ratio of the price of soybeans to the price of soy meal (PSPM) is included to capture this competitive effect (Houck, et al.). The ratio of the livestock units on hand in the importing countries to the quantity of feed grains produced in the importing countries (LW/F) is included to represent the effect analogous to a per capita income effect in a primary demand function (Houck, et al.). We do not agree with this analogy. A time trend variable (T) is in the Model to capture the changes in processing technology (Houck, et al.).

Equation (2.4) is the foreign demand for soybean oil. The quantity demanded for soybean
oil in the international markets (QOX) is a sum of derived demands for oil-using products faced by foreign oil importers. As groundnut oil is a major competitor for soybean oil in the international market, the price ratio of soybean oil to groundnut oil (PO/PG) is included to capture this substitution effect (Houck, et al.). An index of personal income in foreign importing nations (I) is included to capture the income change effects in the soybean oil importing countries. QTXPL represented the quantity of concessional oil exported through P.L. 480. It was hypothesized that some substitution would occur through concessional export (Houck, et al.). Olive-oil production (QOOP) is included to account the competing effects of olive oil from Mediterranean countries in world cooking oil export market (Houck, et al.). Other oil supplies (QAO), groundnut, cottonseed and sunflower-seed oils, are also included to represent substitutes for soy oil in the international market (Houck, et al.). A dummy variable (DV) is included to account for a special trade limitation imposed by the Spanish government in 1952 (Houck, et al.).

\[ Q_{MXt} = f_5\left(\frac{PM}{PL_{Mt}}, \frac{LW}{Ft}, OM_t, CT_t, e_5\right) \]  

Equation (2.5) represents the quantity of meal exports (QMX). The price ratio of soy meal price to linseed meal price (PM, PLM) is to capture the substitutability between soy meal and linseed meal in the international feed market (Houck, et al.). LW/F is the ratio of the livestock units in the importing countries to feed grains produced in these countries (Houck, et al.). OM is the quantity of other oilseed meal imported by the importing countries. It stands for the competing effects of other oilseed-meal imports. CT is a cumulative trend that reflect changes in livestock-feed practices in the importing countries (Houck, et al.).

\[ SS_t = f_6(PS_t - PS_{t-1}, PO_t - PO_{t-1}, SO_t, QSS_t, e_6) \]  

(2.6)
SO_t = f(SS_t, QOP_t, SO_{t-1}, e_t) \tag{2.7}

Equations (2.6) and (2.7) are stock equations for soybeans and soybean oil. The crop year ending stocks of soybeans in the U.S. held privately (SS) is a function of the differences between two successive year soybean prices, soybean oil prices, and crop year ending stocks of soybean oil. Crop year ending stocks of soybean oil is a function of crop year ending stocks of soybeans held privately, the quantity of soybean oil produced, and one-year lagged crop year ending stocks of soybean oil.

\[ QMP_t = 0.0474QSC_t \tag{2.8} \]

\[ QOP_t = 0.0109QSC_t \tag{2.9} \]

Equation (2.8) and (2.9) reflect the average ratio of meal and oil produced with respect to soybeans crushed (outturn rate). In the CROPS model, the technical coefficients of the ratios indicated that one bushel of soybeans yield about 10.9 pounds of oil and 47.4 pounds of meal.

\[ PS_t = 0.474PM_t + 0.109PO_t - W_t \tag{2.10} \]

Equation (2.10) is the price linkage that joins the wholesale value of crushed soybeans to the price received by farmers (PS). The wholesale product value is reflected in the prices per pound of meal and oil, each multiplied by the outturn per bushel. The crushing and handling cost (W) is subtracted. Our preliminary estimation results based on the farm-level soybean price, wholesale soy oil and wholesale soy meal prices (adjusted for inflation with Consumer Price Index, 1999=100) from 1970 to 1999 indicate the relationship between these prices is

\[ 4PS_t = 1.2 + 0.21PM_t + 0.15PO_t \]

\[ PS \text{ is in units of } \$/bushel, \ PM \text{ and } PO \text{ are in units of cents/pound.} \]

---

3 QSC is in units of thousand bushels, QMP and QOP are in unit of million pounds.

4 PS is in units of $/bushel, PM and PO are in units of cents/pound.
\[ QSS_t = QSC_t + QSX_t + SS_t - SS_{t-1} + GS_t \]  (2.11)
\[ QMP_t = QMD_t + QMX_t + SM_t \]  (2.12)
\[ QOP_t = QOD_t + QOX_t + SO_t - SO_{t-1} + QOXPL_t \]  (2.13)

Equations (2.11) through (2.13) are the market equilibrium identities that ensure that the total demand for beans, meal and oil in all outlets are equivalent to total supplies for each crop year. According to the aggregated soybean demand framework, in equation (2.11) the commercial supply of U.S. soybeans (QSS) equals the total of quantity of soybeans crushed in the U.S. (QSC), quantity of soybeans exported as whole beans (QSX), the difference of two successive crop years ending stocks of soybeans (SS_t – SS_{t-1}) and the change in stocks of soybeans owned by the CCC (USDA Commodity Credit Corporation) (Houck, et al.). In equation (2.12), the quantity of soy meal produced in the U.S. equals the summation of quantity of soybean demanded, quantity of soybean export and change in the total U.S. stocks of soy meal. In equation (2.13), the quantity of soybean oil production is a summation of the quantity of domestic soybean oil demanded, soybean oil exported, the differences between two successive crop year ending stocks of soybean oil, and total P.L. 480 exports of vegetable oil.

**The Supply Side for Soybeans in the CROPS Model**

On the supply side of soybeans in the U.S. market, the Model stressed particular interest on the support prices and acreage restrictions for competing crops. Regional supplies of soybeans are taken into consideration. Different effective support price for
crops that can be grown in different regions will affect the allocation of acreage, thus affecting the supply of soybeans.

Houck, et al. divided soybean-planting areas into six regions: the Lake States, the Corn Belt, the Plains, the Delta States, and some of the Atlantic States, and other States. In the Lake States, corn and wheat are viewed as the competing crops of soybeans. In the Corn Belt, only corn is viewed as the competing crops for soybeans. In the Plains States, some regions in the Atlantic States and other states, corn and oats are viewed as competing crops for soybeans. In Delta States and some of the Atlantic States, oats and cotton are viewed as competing crops for soybeans, so, for the supply block, Houck, et al. established seven equations that reflect the effects of the support price of corn, oats, wheat or cotton on the planted acreage of soybeans in different regions. These equations are shown below from (2.14) through (2.20).

\[ AL_t = f_1(AL_{t-1}, PS_{t-1}, PC_{t-1}, PSS_t, PSC_t, PSW_t) \]  \hspace{1cm} (2.14)

\[ ACB_t = f_2(ACB_{t-1}, PS_{t-1}, PC_{t-1}, PSS_t, PSC_t) \]  \hspace{1cm} (2.15)

\[ AP_t = f_3(AP_{t-1}, PS_{t-1}, PC_{t-1}, PSS_t, PSC_t, PSo_t) \]  \hspace{1cm} (2.16)

\[ AD_t = f_4(AD_{t-1}, PS_{t-1}, POT_{t-1}, PSS_t, PSo_t, PSC_t) \]  \hspace{1cm} (2.17)

\[ AA_t = f_5(AA_{t-1}, PS_{t-1}, POT_{t-1}, PSS_t, PSo_t, PSC_t) \]  \hspace{1cm} (2.18)

\[ AM_t = f_6(AM_{t-1}, PS_{t-1}, PC_{t-1}, PSS_t, PSC_t, PSo_t) \]  \hspace{1cm} (2.19)

\[ AM_t = f_7(M_t, PS_{t-1}, PC_{t-1}, PSS_t, PSC_t, PSo_t) \]  \hspace{1cm} (2.20)

Definitions of the variables above are:

- \( AL_t \): soybean acreage harvested in the Lakes States, in thousand acres;
- \( ACB_t \): soybean acreage harvested in the Corn Belt state, in thousand acres;
- \( AP_t \): soybean acreage harvested in the Plains states, in thousand acres;
AA_t: soybean acreage harvested in the Atlantic states, in thousand acres;
AM_t: soybean acreage harvested in the other states, in thousand acres;
PS_{t-1}: lagged soybean price, in dollars per bushel;
PC_{t-1}: lagged corn price, in dollars per bushel;
POT_{t-1}: lagged oats price, in dollars per bushel;
PSS_t: effective support price of soybeans, in dollars per bushel;
PSC_t: effective support price of corn, in dollars per bushel;
PSW_t: effective support price of wheat, in dollars per bushel;
PSO_t: effective support price of oats, in dollars per bushel;
PSCT_t: effective support price of cotton, in dollars per bushel.

Joining Demand and Supply Sides Together

In the CROPS Model, Houck, et al. estimated the simultaneous equations in the demand side first, and then joined supply side with the demand side. They used related annual data from 1946 to 1966. The supply of soybeans (QSS) entered the demand block as predetermined and influences the level of soybean price (PS) for that crop year. The price of soybeans then influences the supply side in the following crop year, t, through a lagged relationship. This produces a new supply in t+1 which enters the demand side and so on. By these assumptions, Houck, et al. actually impose a cobweb in their model.

The supply and demand sides in the CROPS Model are joined by the basic market equilibrium equation that quantity supplied equals quantity demanded, which is represented as:
\[(Yield_t)(A_t) = QOP_t + QMP_t + QSSt_t\]

where Yield is the average yield of soybeans, A is the aggregated acreage of soybeans planted.

Before they studied the price movements by joining the supply and demand, Houck, et al. ran the simultaneous equations for the demand side and equations for the supply side of soybeans separately. For the demand side, the results showed that, except for variables LV, E, QOCD, T, I, QTXPL, QOOP, QAO, OM, QSS, SS, QOP, and SS\textsubscript{t-1}, other structural estimates are all statistically significant given a 5% level of significance. For the supply side, only one-year lagged acreage in the six regions is not statistically significant at a 5% level.

Houck, et al. then aggregated the supply side to a national level by summing the six regional functions and rearrange the appropriate terms, and joined the supply side to the demand side to study the price of soybeans.

**Some Criticisms Regarding the CROPS Model**

Although the parameter estimates Houck, et al. obtained from the two-stage-least squares estimate method fit the data they used well with high R squares, there are some apparent problems with their model.

First of all, Houck, et al. did not include the supply side within the simultaneous equation system. They only set up the simultaneous equations system to describe the demand side of soybeans market, and obtained the parameter estimates without considering the supply side of soybeans. According to microeconomic theory, since price
is the simultaneous result of supply and demand, it is more appropriate to study the demand side together with the supply side.

Second, some of the variables in the CROPS model do not reflect standard microeconomic theory. In the equation for soy meal demand (equation 2.1), Houck, et al. used the quantity supplied of other high-protein feeds (QP) to account for the substitution effects for soy meal in livestock feeds market. Generally, however, the prices of substitutes, not the quantities, are included in demand equations to account for substitution effects. The same criticism applies also to the quantity of other oil supplies (QAO) in soy oil export demand equation (equation 2.4). As we have noted, that Houck, et al. used price ratios like PS/PM (soybean price with respect to soy meal price), PO/PG (soy oil price with respect to groundnut price) to account for the substitution effects of soy meal for soybeans, gronutnut oil for soy oil, however, this is not a general practice in microeconomic theory. Houck, et al. included the ratio of livestock units to the feeding grains in the soybean importing countries (LW/F) in the equation of soybean export demand (equation 2.3) as a variable analogous to the per capita income effect in a primary demand function Houck, et al. However, this does not seem to be a reasonable analogy. In the soy meal demand equation, the variable QOP (percentage of digestible protein in the concentrate ratios for livestock and poultry) is included as an indicator of the continuing change in the feeding practices toward higher protein feeds sources. This variable makes sense, however, it is not a very good one. A time trend variable may perform better to account for the development of higher quality livestock feeds.

Third, since the CROPS Model was established in the late 1960’s, the U.S. has experienced great changes in its own market, consumption behavior, trade policies,
foreign business partners, and the world economic situation. Thus, some of the variables used in CROPS Model may not be relevant or applicable for soybean study today. In the demand equation of soy oil, Houck, et al. includes the price index of butter and lard to account for an animal fats substitution effect for soy oil. However, during the last three decades, the U.S. consumption pattern has been shifted away from diets rich in animal fat to more health diets of fruits, cereal, vegetables and so on. Thus, the animal fat index may not be an important variable for soybean studies today. The concessional oil exported through P.L.480 was an important variable during the time period when the CROPS Model was established. However oil exported through P.L.480 is now so small a fraction of soy oil exports that it can be safely ignored. Houck, et al. considered stocks in the models for soybeans and soy oil. However, the USDA commodity policies regarding inventories have been far less active than they were for the period of time when this model was established (1946 to 1966), and preliminary estimates suggest that changes in stocks are effectively random. Thus it is reasonable to include stock changes merely as data in market equilibrium equations.

Based on these criticisms, it is necessary to make modifications and establish a more applicable model that adheres better to economic theory.

---

5 P.L. 480 (Public Law 480 is also known as Food for Peace Program.)
Summary

In this chapter, we have reviewed the USDA CROPS Model of the U.S. soybean market developed by Houck, et al. Each equation and variable are discussed in great detail. We then raised some criticisms regarding that model. Based on these criticisms, it is necessary for us to make some reasonable modifications to develop our own model. In the next chapter, such a model is established and each variable will be explained.
Chapter 3

EMPIRICAL MODEL AND DATA

In Chapter Two, we reviewed the CROPS model of the U.S. soybean market from post-World War II to the 1970's. This chapter will modify that model and develop the empirical framework for our study describing the U.S. soybean markets of more recent years. Some new variables are included.

The data used in this study are from USDA Agricultural Statistics (various years). They are annual data that cover 29 years (1970 to 1998). All prices are adjusted for inflation with the Consumer's Price Index (1999=100).

Reasons For Using A Simultaneous Equation System

Four major ideas underpin the simultaneous equation system used for our model. Meal and oil are joint products from the crushing of soybeans, thus meal and oil supplies are tightly linked to each other and to the quantity of soybeans crushed in the United States. Multiple-market outlets are available for all three products. These market channels are essentially independent of each other. However, prices and product flows of soybeans, meal and oil are determined simultaneously because of the joint-product relationship. The market for soybeans is actually the aggregate of the three markets for soybeans, soy meal and soy oil. These markets can also be grouped into domestic and the international markets. Thus, in our analysis, we establish a simultaneous system of
equations to describe soybeans and its products market both in the U.S. and in the rest of the world.

**The Demand Side of Soybean Market**

**Soybean Market**

Our preliminary estimates suggest that soybeans are crushed to about 49% soy meal and 21% soy oil. U.S. domestic demand for soybeans primarily comes from the demand for these products. We will not establish an equation specifically for the domestic soybean market, since it is implicit in the oil and meal demands.

The major source of demand for soybeans in the world market is for crushing, thus foreign buyers can buy soybeans or soy meal/oil. For the equation of the world demand for soybeans, we include variables of soybean price and soy meal price to capture these substitute effects. Another variable included is the South American export of soybeans. South American soybean exports have been increasing steadily. Figure 3.1 shows the increasing share of South American soybean exports in the international soybean export market.

The increase of the South American soybean exports is expected to continue due to the increasing production of soybeans in South American counties. Relatively high yield, cheap labor cost and cheap land have been favorable to the soybeans production in the South American countries since the 1970's (Frederick). Based on our analysis for soybean export market, we establish an equation for soybean export demand equation, as shown in equation (3.1)
(Source: USDA Agricultural Statistics 1995 to 2000)

\[ QSX_t = f_1(PS_t, PM_t, SX_t, e_t), \]  
\hspace{2cm} (3.1)

where, QSX is the quantity of soybeans demanded by foreign buyers, PS is the farm-level soybean price, PM is the wholesale price of soy meal, and SX is the South American soybean exports, \( t \) represents years from 1970 to 1998 (1970=1). Table 3.1 provides the summary statistics for variables used in this equation.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Unit</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soybean Price (PS)</td>
<td>S/bushel</td>
<td>29</td>
<td>10.65</td>
<td>4.06</td>
</tr>
<tr>
<td>Soy Meal Price (PM)</td>
<td>Cents/pound</td>
<td>29</td>
<td>16.17</td>
<td>6.96</td>
</tr>
<tr>
<td>South American Soybeans Export (SX)</td>
<td>000 metric tons</td>
<td>29</td>
<td>24464</td>
<td>14825</td>
</tr>
<tr>
<td>Soybeans Exported (QSX)</td>
<td>Million pounds</td>
<td>29</td>
<td>41127</td>
<td>9253</td>
</tr>
</tbody>
</table>
Soy Meal Market

Domestic demand for soy meal derives from the demand for livestock meat since livestock are the primarily consumers of soy meal. The number of livestock units\textsuperscript{4} in the U.S. (AF) is therefore included as a population shifter. In the livestock feed market, soy meal has some substitutes, for example corn meal, cottonseed meal, etc. We assume only the corn meal to be the substitute for soy meal in this thesis, and use the price of corn meal as an independent variable to capture substitution effects\textsuperscript{5}. A time trend variable is used to capture the effects of the development of higher quality of livestock feeds market. Based on these assumptions, domestic demand for soy meal is represented in equation (3.2):

\[ QMD_t = f_t(PM_t, PCOM_t, AF_t, T_t, e_t), \quad (3.2) \]

where QMD is the domestic demand for soy meal, PM is the wholesale soy meal price, AF is the livestock units, and T is the time trend variable. Summary statistics for variables used in equation (3.2) are shown in Table 3.2.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Unit</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity of Meal Demanded (QMD)</td>
<td>Million pounds</td>
<td>29</td>
<td>25910</td>
<td>5967</td>
</tr>
<tr>
<td>Price of Soy Meal (PM)</td>
<td>Cents/pound</td>
<td>29</td>
<td>16317</td>
<td>6.56</td>
</tr>
<tr>
<td>Price of Corn Meal (PCOM)</td>
<td>Cents/pound</td>
<td>29</td>
<td>13.61</td>
<td>5332</td>
</tr>
<tr>
<td>Livestock Units (AF)</td>
<td>000 animal units</td>
<td>29</td>
<td>66950</td>
<td>5959</td>
</tr>
</tbody>
</table>

World demand for soy meal also derives primarily from the demand for livestock feeds. We assume that corn meal is the only substitute for soy meal in the world livestock

\textsuperscript{4} A livestock production unit is approximately 1000 pounds of animal live weight.\textsuperscript{5} Although other meals are relevant, high correlations among these prices suggests that these effects are captured by the corn meal price and would cause multicollinearity problems if included.
feed market (See Footnote 5). Prior to the 1960’s, some soybean importing counties bought only meal. Since that time, most importers have begun crushing domestically, thus, a time trend variable is used to account for these changes. The export demand for soy meal is therefore represented in equation (3.3).

\[ Q_{MXt} = f_3(P_{Mt}, P_{COMt}, T_t, e_3), \]

(3.3)

where \( Q_{MX} \) is the quantity of soy meal demanded by foreign buyers, \( P_M \) is the wholesale soybean price, \( P_{COM} \) is the wholesale price of corn meal, \( T \) is the time trend variable. Summary statistics for variables used in equation (3.3) are provided in Table 3.3.

**Table 3.3: Summary Statistics for Variables Used in the Equation of Soy Meal**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Unit</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Export Demand for Soy Meal Demanded (QMX)</td>
<td>Million pounds</td>
<td>29</td>
<td>601.1</td>
<td>1181</td>
</tr>
<tr>
<td>Price of Soy Meal (PM)</td>
<td>Cents/pound</td>
<td>29</td>
<td>16.17</td>
<td>6.56</td>
</tr>
<tr>
<td>Price of Corn Meal (P_{COM})</td>
<td>Cents/pound</td>
<td>29</td>
<td>13.61</td>
<td>5.32</td>
</tr>
</tbody>
</table>

(Source: USDA Agricultural Statistics, various years)

**Soy Oil Market**

In both the domestic and world markets, soy oil is used primarily as cooking oil. Although soy oil also has nonfood uses, such as soap, paints, drying oils, and plastics, these outlets are a very small share of the total soy oil market, thus, are not considered here.

For the equation of domestic soy oil demand, we assume that corn oil\(^6\) is the only substitute for soy oil, thus corn oil price is included to account for the substitution effect in the cooking oil market. A variable of the total annual real expenditures on food in the

---

\(^6\) Although other oils are relevant, high correlations among these prices suggests that these effects are captured by the corn oil price and would cause multicollinearity problems if included.
U.S is also included to account for income and population effects. Equation (3.4) represents the domestic soy oil demand.

\[ QOD_t = f_4(PO_t, PCOO_t, E_t, e_4) \]  

(3.4)

where QOD is the quantity of soy oil demanded in domestic market, PO is the wholesale soy oil price, PCOO is the wholesale corn oil price, E is real expenditures on food.

The summary statistics for variables used in this equation are shown in Table 3.4.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Unit</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domestic Soy Oil Demand (QOD)</td>
<td>Million pounds</td>
<td>29</td>
<td>12021</td>
<td>2897</td>
</tr>
<tr>
<td>Soy Oil Price (PO)</td>
<td>Cents/pound</td>
<td>29</td>
<td>41.8</td>
<td>19.92</td>
</tr>
<tr>
<td>Corn Oil Price (PCOO)</td>
<td>Cents/pound</td>
<td>29</td>
<td>24.43</td>
<td>5.61</td>
</tr>
<tr>
<td>Real Expenditures on Food (E)</td>
<td>Billion dollars</td>
<td>29</td>
<td>334.41</td>
<td>164.56</td>
</tr>
</tbody>
</table>

For the world cooking oil market, we assume that soy oil export demand is only affected by its substitute, corn oil. Thus the export demand for soy oil is represented in equation (3.5) as:

\[ QOX_t = f_5(PO_t, PCOO_t, e_s) \]  

(3.5)

where QOX is the quantity of soy oil exported from U.S., PO is the wholesale price of soy oil, and PCOO is the wholesale price of corn oil. Summary statistics for variables used in equation (3.5) are provided in Table 3.5.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Unit</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soy Oil Exported (QOX)</td>
<td>Million Pounds</td>
<td>29</td>
<td>1738</td>
<td>523.33</td>
</tr>
<tr>
<td>Price of Soy Oil</td>
<td>Cents/pound</td>
<td>29</td>
<td>41.8</td>
<td>19.92</td>
</tr>
<tr>
<td>Price of Corn Oil</td>
<td>Cents/pound</td>
<td>29</td>
<td>24.44</td>
<td>5.61</td>
</tr>
</tbody>
</table>
Stock Demand

As we have mentioned in Chapter Two, in the CROPS Model, Houck, et al. considered the demand for soybean, soy oil and soy meal stocks as endogenous variables. However, the USDA commodity policies regarding inventories have been far less active than they were for the period of time for which the CROPS Model was estimated (1946 to 1966). Preliminary estimates suggested that changes in stocks are effectively random. We therefore use stocks for soybeans, soy oil and soy meal merely as data in the equilibrium equation.

The Supply Side of Soybeans

On the supply side of soybeans, we use an acreage response model, just as did Houck, et al., but we make some changes to his original model. Since the implementation of the 1990 Farm Bill, farmers have been granted planting flexibility so some variables in his model are not as important as they were during the late 1960’s when Houck, et al. established the CROPS Model. To begin, we first review some theories on product supply and acreage response.

Product Supply and Acreage Response

The traditional production function can be derived from firm’s profit maximizing rule. Assume that the firm’s economic profit is a function of output price, inputs and their wages and other nonprice factors:

$$\pi = P \cdot Q - (La \cdot A + \sum LiXi),$$
where $P$ is the output price, $A$ is the acreage, $X_i$ is the $i$th input employed, $L_a$ is the wage for crop land, $L_i$ is the wage for the $i$th other input. By transforming the equation above, we may represent $Q$ as:

$$Q = f_q(A, X, Z),$$

where $Z$ is the vector of other factors.

We obtain the factor demand functions from the first order conditions of the profit function and substitute it into the production function above and get the firm’s supply function as follows:

$$Q_s = f_s(P, C, Z),$$

where $C$ represents a generalized vector of factor costs.

Farmers are faced with uncertainty regarding the quantities supplied ($Q_s$ in the above equation). There are at least two sources of the uncertainty, the unpredictable variation of yield and differences between planned and planted acreage. The yield is uncertain because it is affected by weather and other uncontrollable factors, so it is hard for farmers to adjust their production precisely. However, since they have more control over the acreage than production, farmers usually can adjust their acreage allocated to a crop. In this sense, acreage is a better index of the producers’ reactions to price changes than is production. Thus acreage planted, rather than production, is often used to formulate the farmer’s response in both theory and application. The acreage response model is:

$$A = f_A(P, C, Z),$$

Planted acreage is not a perfect proxy for production. Cassels describes two problems with the approximation. First, the weather and other problems may cause farmers to be unable to fully plant their planned acreage or to fully harvest planted
acreage. Second, the acreage gives no indication of the response that is made through increased or decreased intensity of cultivation. Nerlove (1958) and later Behrman specify two additional inadequacies of the approximation. First, the land is an heterogenous factor. Each piece of land drawn into production of a particular crop or released from production has either higher or lower fertility than the acres already in production or remaining in production. Second, the land is only one of many inputs in agriculture. All these inputs might not increase or decrease proportionately, thus the output per unit of any one input changes. Nerlove and Behrman also compare the elasticity of planned output and planned acreage to illustrate the discrepancy between planned output and planned acreage. They find that relative magnitudes of these two elasticities are hard to identify.

In spite of this imperfect approximation, the planted acreage may be the best available indication of supply response because acreage changes are the primary means to control production while yields may vary through changes in intensity or uncontrollable factors. It is thus reasonable to focus on an acreage model of a crop supply, rather than the more standard direct supply response.

**Factors Relating to Acreage Response Model**

In our model on the supply side of soybean beans, the acreage response model is a function expressed in equation (3.6) as

\[
A_t = f_6 (A_{t-1}, PS_{t-1}, COST_t, PW_{t-1}, PCORN_{t-1}, e_6),
\]

where \( A \) is the acreage allocated to planting soybean, \( PS \) is the farm-level soybean price, \( COST \) is the production cost, \( PW \), and \( PCORN \) are farm-level prices of wheat and corn. The subscript \( t \) represents time period \( t \), and \( t-1 \) represents one-year lag.
One-year lagged acreage is included to account for the effects of asset fixity on production. Farmers tend to establish a fixed ratio of equipment to acreage. As this equipment is of low salvage value, it makes no sense for farmers to decrease acreage often; on the other hand, when the equipment is used at capacity, one unit increase in acreage will require large amount of increase in equipment. Thus asset fixity inhibits increases in acreage also.

Cassels, in his supply response literature in 1933, concludes that the acreage response to price approach based on the past experience with respect to prices is more practical than other approaches. The output price of the current year is unknown to farmers when they make their planting decisions. So farmers respond to price signals other than current year’s actual observed price. For simplicity reasons, we assume that farmers have naïve expectations for farm-level soybean prices.

Production Cost (COST) is another important factor in production. U.S. Department of Agriculture estimates of variable costs of production for selected field crops are considered the best available estimates for production costs for corn and soybean (Shideed and White). The index of prices paid by farmers for production items, interest, taxes, and wage rates was used to adjust the cost values for the study period in this study. We assume that cost is known when the planting decisions are made.

Farm level expected prices for corn, and wheat are included here to capture the competing crop effects. Opportunity cost is another important determinant in production. Some previous studies included the prices of competing crops when they analyzed the production. Love and Willette, for example, include competing crops in their model of potato acreage response. In our model, we include corn and wheat as the competing crops
or joint crops with soybeans. The effects are reflected by the changes of acreages allocated to soybeans with respect to the changes of the price of corn and wheat. For simplicity, we assume that farmers make acreage decisions based on last years’ corn and wheat farm-level prices. Summary statistics of the variables we include in the acreage response model are provided in Table 3.6.

**Table 3.6: Summary Statistics for Variables Used in Acreage Response Equation (Equation 3.6)**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Unit</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable Cost</td>
<td>COST $/acre (adjusted by producer price index, 1990–1992 =100)</td>
<td>29</td>
<td>3285</td>
<td>461</td>
</tr>
<tr>
<td>Yield</td>
<td>Y Pounds/acre</td>
<td>29</td>
<td>1893</td>
<td>271</td>
</tr>
<tr>
<td>Farm-level soybean price</td>
<td>PS $/bushel</td>
<td>29</td>
<td>10.62</td>
<td>4.12</td>
</tr>
<tr>
<td>Farm-level corn price</td>
<td>PCORN $/bushel</td>
<td>29</td>
<td>3.38</td>
<td>1.18</td>
</tr>
<tr>
<td>Farm-level wheat price</td>
<td>PW $/bushel</td>
<td>29</td>
<td>4.76</td>
<td>1.88</td>
</tr>
<tr>
<td>Acres</td>
<td>A Million acres</td>
<td>29</td>
<td>159.10</td>
<td>7.64</td>
</tr>
</tbody>
</table>

(Source: USDA Agricultural Statistics various years)

**Technical and Physical Relationships**

That, discussed above, forms the behavioral equations in our model of the soybean demand side. These equations indicate the formal constraints on the variables we use in the model. Besides these, we also must identify the technical and physical relationships to ensure that market supply of soybeans equals market demand.

**Price Linkage Equation**

A price linkage equation between the farm-level price of soybean and wholesale price and soy meal and soy oil takes the form of:

\[
PS_t = f(\text{PM}_t, \text{PO}_t, \epsilon_t), \quad (3.7)
\]
where PS is the farm-level price of soybeans, PM is the wholesale soy meal price, and PO is the wholesale soy oil price. This price linkage links the farm level soybean price to the wholesale soy meal price and wholesale soy oil price. Our preliminary estimation results based on the soybean, soy oil and soy meal prices (adjusted for inflation with consumer price index, 1999=100) from 1970 to 1999) indicate that the relationship between these prices is:

\[ PS_t = 1.2 + 0.21 PM_t + 0.15 PO_t \]

**Ratio of Soy Oil/Meal Produced to Soybeans Crushed**

Two equations linking the soybeans crushed to soy meal produced and soy oil produced take the forms of

\[ QMP_t = f(QSC_t, e_8) \] (3.8)

\[ QOP_t = f(QSC_t, e_9) \] (3.9)

They represent the fixed ratio of soy oil and meal crushed from soybeans. Quantities of soy meal and soy oil produced are tightly linked to the quantity of soybeans crushed. Our preliminary estimate results indicate that soybeans crush to about 49% meal and 21% oil.

**Market Equilibrium Identities**

Three identity equations are included to ensure the quantities of soybeans supplied equal to the quantities of soybeans demanded, as in the CROPS Model. These are represented in equation (3.10) through (3.12).

\[ QSS_t = QSC_t + QSX_t + SS_t - SS_{t-1} \] (3.10)

7 PS is in units of $/bushel, PM and PO are in units of cents/pound.
\[ QMP_t = QMD_t + QMX_t + SM_t - SM_{t-1} \]  
\[ QOP_t = QOD_t + QOX_t + SO_t - SO_{t-1} \]

Equation (3.10) represents soybeans market equilibrium. On the left hand side, \( QSS \) is the market supply of soybeans, which is expressed by

\[ QSS_t = (A_t)(Yield_t), \]

in which \( \text{Yield} \) is assumed to be constant and take value of the average yield of soybeans in the U.S. from 1970 to 1998 (1891 pounds/acre). On the right-hand-side of equation (3.11), \( QSC \) is quantity of soybeans crushed; \( QSX \) is the quantity of soybeans exported.

\( SS_t - SS_{t-1} \) represents the stock changes. Equation (3.12) represents the market equilibrium for soy meal. On the left-hand-side of the equation, \( QMP \) is the quantity of soy meal produced. It equals the sum of the quantity of meal domestically demanded (\( QMD \)), quantity of meal exported (\( QMX \)), and the changes of soy meal stocks (\( SM_t - SM_{t-1} \)). Equation (3.12) represents market equilibrium for soy oil. The quantity of soy oil produced (\( QOP \)) is equal to total of the quantity of soy oil domestically demanded (\( QMD \)), quantity of soy oil exported (\( QOX \)), and soy oil stock changes (\( SO_t - SO_{t-1} \)).

Summary statistics for the variables used in the identity equations are provided in Table 3.7.

---

\(^8\) Soybean yield in the U.S. has increased from 1970 to 1998. However, yields vary slightly since the 1990’s, thus, we will take the average of the yields from 1990 to 1998 as the constant (2204 pound/acre).
Table 3.7: Summary Statistics for Variables Used in Identity Equations

<table>
<thead>
<tr>
<th>Variables</th>
<th>Unit</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soybeans Crushed (QSC)</td>
<td>Million pounds</td>
<td>29</td>
<td>65422</td>
<td>14928</td>
</tr>
<tr>
<td>Soy Meal Produced (QMP)</td>
<td>Million pounds</td>
<td>29</td>
<td>31917</td>
<td>6892</td>
</tr>
<tr>
<td>Soy Oil Produced (QOP)</td>
<td>Million pounds</td>
<td>29</td>
<td>13741</td>
<td>3157</td>
</tr>
<tr>
<td>Soy Meal Domestically Demanded (QMD)</td>
<td>Million pounds</td>
<td>29</td>
<td>25910</td>
<td>5967</td>
</tr>
<tr>
<td>Soy Oil Domestically Demanded (QOD)</td>
<td>Million pounds</td>
<td>29</td>
<td>12021</td>
<td>2897</td>
</tr>
<tr>
<td>Soybeans Stock (SS)</td>
<td>Million pounds</td>
<td>29</td>
<td>14461</td>
<td>6552</td>
</tr>
<tr>
<td>Soy Meal Stock (SM)</td>
<td>Million pounds</td>
<td>29</td>
<td>246</td>
<td>93</td>
</tr>
<tr>
<td>Soy Oil Stock (SO)</td>
<td>Million pounds</td>
<td>29</td>
<td>1196</td>
<td>501</td>
</tr>
</tbody>
</table>

**Summary**

In this Chapter, we establish a simultaneous equations model for U.S. soybeans Market. The model is based on the USDA CROPS Model developed by Houck, et al., however, we have made some necessary modifications. Variables in each equation are explained in detail, and associated data are also described. Note that some of the variables in the USDA CROPS model do not appear in our model. The reason for this is that since the CROPS Model was established in the 1960’s, there have been tremendous structural changes in the world economy. As a result, we include some new variables in our model. In the next chapter, we estimate our empirical model with data from 1970 to 1998. After we find the parameter estimates, we then define the reduced form equation for soybean prices and find the economic factors that influence soybean prices.
Chapter 4

EMPIRICAL RESULTS

In Chapter Three, a simultaneous equation model describing the U.S. soybean market was developed and each explanatory variable was discussed in some detail. In Chapter Four, we report the estimation results and discuss the reduced form estimates of the model. Finally, these results are applied to the U.S. Canola market.

Statistical Methods And Estimation Results

In Chapter Three, we develop a simultaneous model for both demand and supply sides of the soybean market. This simultaneous system is presented below, where variables marked by * are exogenous variables.

\[ Q_{MDt} = \alpha_{11} + \alpha_{12} P_{Mt} + \alpha_{13} P_{COM*} + \alpha_{14} A_{F*} + \alpha_{15} T_{t} + e_{1}, \]

\[ Q_{ODt} = \alpha_{21} + \alpha_{22} P_{Ot} + \alpha_{23} P_{COO*} + \alpha_{24} E_{t} + e_{2}, \]

\[ Q_{SXt} = \alpha_{31} + \alpha_{32} P_{S} + \alpha_{33} P_{Mt} + \alpha_{34} S_{X*} + e_{3}, \]

\[ Q_{MXt} = \alpha_{41} + \alpha_{42} P_{Mt} + \alpha_{43} P_{COM*} + \alpha_{44} T_{t} + e_{4}, \]

\[ Q_{OXt} = \alpha_{51} + \alpha_{52} P_{Mt} + \alpha_{53} P_{COO*} + \alpha_{54} T_{t} + e_{5}, \]

\[ Q_{MPt} = \alpha_{61} Q_{Sc} + e_{6}, \]

\[ Q_{Opt} = \alpha_{71} Q_{Sc} + e_{7}, \]

\[ P_{St} = \alpha_{81} P_{Mt} + \alpha_{82} P_{Ot} + e_{8}, \]

\[ Q_{SS} = Q_{Sc} + Q_{SX} + S_{S} + S_{S*} + S_{S*}^{t-1}, \]

\[ Q_{MP} = Q_{MD} + Q_{MX} + S_{M*} + S_{M*}^{t-1}, \]
\[ QOP_t = QOD_t + QOX + SO^*_t - SO^*_{t-1}, \]  
\[ A_t = \alpha_9 + \alpha_{10} A^*_t + \alpha_{11} PS^*_t - 1 + \alpha_{12} PCORN^*_t - 1 + \alpha_{13} PW^*_t - 1 + \alpha_{14} COST^*_t + \epsilon_9, \]  
\[ QSS_t = 2204 A_t, \]  
\[ (4.11) \]
\[ (4.12) \]
\[ (4.13) \]

All error terms in the above equations are assumed to be normally distributed with zero means and constant variances. Recall that endogenous variables in the system are QMD (quantity of soy meal demanded in the domestic market), QOD (quantity of soybean oil demanded in the domestic market), QSX (quantity of soybeans demanded by foreign buyers), QMX (quantity of soy meal exported), QOX (quantity of soy oil exported), QMP (quantity of meal produced), QOP (quantity of oil produced), PS (farm level soybean price), PM (wholesale soy meal price), and PO (wholesale soy oil price), and QSS (soybeans supply).

The exogenous variables are, \( AF_t \) (animal units), \( T_t \) (a time trend variable), \( PCOO_t \) (the price of corn oil), \( PCOM_t \) (the price of corn meal), \( E_t \) (real expenditures on food), \( SX_t \) (South American soybeans exports), \( SS_t \) (soybean stocks), \( SM_t \) (soy meal stocks), and \( SO_t \) (soy oil stocks), \( A_t \) (soybean acreage), \( PCORN_t \) (farm level price of corn), \( PW_t \) (farm level price of wheat), \( COST_t \) (Variable cost of growing soybeans), \( A_{t-1} \) (one-year lagged acreage of soybeans) and \( P_{t-1} \) (one-year lagged farm-level soybean price).

As the simultaneous system is over-identified (See Appendix B), OLS estimates are biased and inconsistent. They tend to overestimate the parameter estimates, but underestimate the intercepts, however, by using two-stage-least squares method, we can estimate parameters that are more likely to be consistent and efficient. A two-stage least square method is used to estimate the coefficients in this simultaneous system (See Appendix B).
In the following tables showing the estimation results. Numbers marked with * are statistically significant at 5% level of significance; numbers marked with ** are statistically significant at 10% level.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficients</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1571</td>
<td>10945</td>
</tr>
<tr>
<td>Price of Soy Meal (PM)</td>
<td>-167.89</td>
<td>209</td>
</tr>
<tr>
<td>Price of Corn Meal (PCOM)</td>
<td>230</td>
<td>332</td>
</tr>
<tr>
<td>Time Trend Variable (T)</td>
<td>579*</td>
<td>168</td>
</tr>
<tr>
<td>Animal Units (AF)</td>
<td>0.18</td>
<td>0.19</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.89</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficients</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>6756*</td>
<td>1210</td>
</tr>
<tr>
<td>Price of Soy Oil (PO)</td>
<td>-31.2</td>
<td>25</td>
</tr>
<tr>
<td>Price of Corn Oil (PCOO)</td>
<td>86.27</td>
<td>55</td>
</tr>
<tr>
<td>Real Expenditures on Food (E)</td>
<td>13.28*</td>
<td>2.95</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.90</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficients</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>37836*</td>
<td>11889</td>
</tr>
<tr>
<td>Price of Soybeans (PS)</td>
<td>-242</td>
<td>838</td>
</tr>
<tr>
<td>Price of Soy Meal (PM)</td>
<td>-26</td>
<td>398</td>
</tr>
<tr>
<td>South American Soybean Export (SX)</td>
<td>0.26</td>
<td>0.19</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>Variables</td>
<td>Coefficients</td>
<td>Standard Error</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>--------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Intercept</td>
<td>2712</td>
<td>2006</td>
</tr>
<tr>
<td>Price of Soy Meal (PM)</td>
<td>-97.16</td>
<td>108.36</td>
</tr>
<tr>
<td>Price of Corn Meal (PCOM)</td>
<td>182.95</td>
<td>173.27</td>
</tr>
<tr>
<td>Time Trend Variable (T)</td>
<td>119.11*</td>
<td>50.1</td>
</tr>
<tr>
<td>Adjusted R^2</td>
<td>0.29</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.6: Parameter Estimates for Equation (4.6) – Soy Meal Produced (QMP)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficients</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>998*</td>
<td>444</td>
</tr>
<tr>
<td>Price of Soy Oil (PO)</td>
<td>-10.87*</td>
<td>5.05</td>
</tr>
<tr>
<td>Price of Corn Oil (PCOO)</td>
<td>47.9*</td>
<td>18.3</td>
</tr>
<tr>
<td>Adjusted R^2</td>
<td>0.21</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.7: Parameter Estimates for Equation (4.7) – Soy Oil Produced (QOP)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficients</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soy Oil Crushed</td>
<td>0.49*</td>
<td>0.003</td>
</tr>
<tr>
<td>Adjusted R^2</td>
<td>0.99</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.8: Parameter Estimates for Equation (4.8) – Wholesale Soy Meal Price (PM) And Wholesale Soy Oil Price (PO) Regressed on Farm-level Soybean

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficients</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.2**</td>
<td>0.62</td>
</tr>
<tr>
<td>Price of Soy Meal (PM)</td>
<td>0.21*</td>
<td>0.046</td>
</tr>
<tr>
<td>Price of Soy Oil (PO)</td>
<td>0.15*</td>
<td>0.016</td>
</tr>
<tr>
<td>Adjusted R^2</td>
<td>0.91</td>
<td></td>
</tr>
</tbody>
</table>
Table 4.9: Parameter Estimates for Equation (4.12)-Soybean Acreage

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficients</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>84858*</td>
<td>12576</td>
</tr>
<tr>
<td>One-year lagged Acreage (At.1)</td>
<td>0.24*</td>
<td>0.10</td>
</tr>
<tr>
<td>One-year lagged Farm level price of soybeans (PS$_{t-1}$)</td>
<td>443.19*</td>
<td>162.14</td>
</tr>
<tr>
<td>One-year Lagged Farm level price of wheat (PW$_{t-1}$)</td>
<td>1580.79*</td>
<td>487.86</td>
</tr>
<tr>
<td>One-year lagged Farm level price of corn (PCORN$_{t-1}$)</td>
<td>-1103*</td>
<td>763</td>
</tr>
<tr>
<td>Variable cost (COST)</td>
<td>-14.79*</td>
<td>2.47</td>
</tr>
<tr>
<td>Adjusted R$^2$</td>
<td>0.90</td>
<td></td>
</tr>
</tbody>
</table>

**Results Discussion**

**Elasticity Analysis of Soybean Demand and Supply**

The elasticities of soybean demand and supply are calculated at mean levels. The general formula we use is

$$E_{x_1,x_2} = \frac{(\partial x_1/\partial x_2)(\text{Mean of } x_2/\text{Mean of } x_1)}$$

in which, the elasticity of $x_1$ with respect to $x_2$ ($E_{x_1,x_2}$) equals to the percentage changes of $x_1$ with respect to $x_2$ ($\partial x_1/\partial x_2$) times the ratio of the mean of $x_2$ to the mean of $x_1$.

Table 4.10 presents the elasticity of domestic soy oil demanded with respect to real expenditures on food computed at mean level. One percentage increase in the real expenditures on food (E) will increase the quantity of soy oil demanded by 0.36%. The sign is of our expectation. Increases in expenditures on food are expected to increase the quantities of soy oil demanded.
Table 4.1 presents the elasticity of soy oil export demand with respect to the price of soy oil and corn oil. The soy oil exported from the U.S. has a positive elasticity with wholesale soy oil price and negative elasticity with wholesale corn oil price. One percentage increase in soy oil price will decrease soy oil exported by 0.26%, while one percentage increase in corn oil price will increase soy oil exported by 0.66%. This result suggests that corn oil is a substitute for soy oil in the world cooking oil market. The signs correspond to our expectation, as increases in corn prices will decrease the quantity demanded for corn, but increase the quantity of soy oil demanded.

**Table 4.11: Elasticity of Soy Oil Export Demand Computed at Mean Levels**

<table>
<thead>
<tr>
<th></th>
<th>Soy Oil Export Demand (QOX)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price of Sov Oil (PO)</td>
<td>-0.26</td>
</tr>
<tr>
<td>Price of Corn Oil (PCOO)</td>
<td><strong>0.66</strong></td>
</tr>
</tbody>
</table>

Table 4.12 presented the elasticity of the supply side (using acreage as a proxy) with respect to one-year lagged acreage ($A_{t-1}$), one-year lagged farm level soybean price ($PS_{t-1}$), one-year lagged corn price ($PCORN_{t-1}$), one-year lagged wheat price ($PW_{t-1}$) and the variable cost of growing soybeans ($COST$). One percentage increase in last year acreage will increase this year’s acreage by 0.24%. One percentage increase in last year’s soybean price will increase this year’s acreage by 0.07%. One percentage increase in last year’s corn price will decrease this year’s soybean acreage by 0.06%. One percentage increase in last year wheat price will decrease this year’s soybean acreage by 0.13%. One
percentage increase in the variable cost of growing soybeans will decrease the soybean acreage by 0.82%. These signs correspond to our expectations. Last year soybean price increases will encourage farmers to expand soybean acreage this year. Last year variable cost increases tend to increase last year’s soybean price, thus the quantity demanded for soybeans last year will have decreased; therefore, in this year, soybean farmers will decrease the acreage of soybeans due to the shrinking market. Corn is considered a competing product for soybeans. Last year’s corn price increase encourages farmers to allocate more acreage to growing corn this year, while decreasing the acreage allocated to soybeans. Wheat is considered as a joint-crop with soybeans. Last year’s wheat price increase will encourage farmers to expand acreage of wheat this year, which, in turn, increases the acreage of soybeans this year.

<table>
<thead>
<tr>
<th></th>
<th>Acreage (A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>One year lagged acreage (A_{t-1})</td>
<td>0.24</td>
</tr>
<tr>
<td>One year lagged soybean price (PS_{t-1})</td>
<td>0.06</td>
</tr>
<tr>
<td>One year lagged wheat price (PW_{t-1})</td>
<td>0.13</td>
</tr>
<tr>
<td>One year lagged corn price (PCORN_{t-1})</td>
<td>-0.82</td>
</tr>
<tr>
<td>Variable cost of growing soybeans (COST)</td>
<td>-0.24</td>
</tr>
</tbody>
</table>

**Reduced Form of Soybean Farm Level Prices**

By substituting the parameter estimates into the reduced form of soybean farm level price (See Appendix D for the procedure) and rearranging the terms, we obtain the reduced form of soybean farm level price, as shown below:

\[
PS_t = 3.2 - 0.000013A_{t-1} - 0.022PS_{t-1} + 0.057PCORN_{t-1} - 0.082PW_{t-1} + 0.0008COST_t
+ 0.006PCOM_t + 0.000003AF_t + 0.011T_t + 0.000015(SM_t - SM_{t-1}) + 0.000012SX_t
+ 0.000026*(SS_t - SS_{t-1}) + 0.012PCOO_t + 0.0011E_t + 0.000088(SO_t - SO_{t-1})
\]

PS_t is the farm level price of soybeans of year t, A_{t-1} is the one-year lagged acreage,
PS_{t-1} is the one-year lagged farm-level soybean price, PCORN_{t-1} is the one-year lagged farm-level corn price, PW_{t-1} is the one-year lagged farm level wheat price, COST\_t is the variable cost of growing soybeans, PCOM\_t is the wholesale price of corn meal, AF is animal units in year t, T\_t is the time trend variable, (SM\_t-SM_{t-1}) is the changes of soy meal stocks from year t-1 to year t, SX\_t is South American soybeans exports, (SS\_t-SS_{t-1}) is the soybean stock changes from year t-1 to year t, PCOO\_t is the wholesale price of corn oil, (SO-SO_{t-1}) is the changes of soy oil stocks from year t-1 to year t, E\_t is the real expenditures on food.

We present the soybean price elasticity with some of the right-hand-side variables in Table 4.13.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Farm-Level Price of Soybeans (PS_t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-year lagged Acreage (A_{t-1})</td>
<td>-0.07</td>
</tr>
<tr>
<td>One-year lagged Farm-level Soybeans Price (PS_{t-1})</td>
<td>-0.022</td>
</tr>
<tr>
<td>One-year lagged Farm-level Corn Price (PCORN_{t-1})</td>
<td>0.02</td>
</tr>
<tr>
<td>One-year lagged Farm-level Wheat Price (PW_{t-1})</td>
<td>-0.04</td>
</tr>
<tr>
<td>Variable Cost of Growing Soybeans (COST_t)</td>
<td>0.25</td>
</tr>
<tr>
<td>Wholesale Corn Oil Price (PCOO_t)</td>
<td>0.03</td>
</tr>
<tr>
<td>Real Expenditures on Food (E_t)</td>
<td>0.035</td>
</tr>
</tbody>
</table>

Our results suggest that the farm-level price has positive elasticity with respect to one-year lagged farm-level corn price, variable cost of growing soybeans, wholesale corn oil price and real expenditures on food. Farm-level soybean price has negative elasticity with respect to one-year lagged acreage, one-year lagged farm-level wheat price, and one-year lagged soybeans price. These signs correspond to our expectations.

A one percent increase in one-year lagged farm-level corn price results in 0.02% increase in the farm level soybeans price, which suggests that corn and soybeans are
competing crops. The increases in last year’s corn price will encourage farmers to allocate more acreage to corn this year, while less for soybeans, thus the production of soybeans of this year is expected to decrease which, in turn, decreases the quantities of soybean supplied this year. The farm-level soybean prices of this year will increase as a result.

A one percent increase in corn oil price increases the farm-level soybean price by 0.03%. As corn oil and soy oil are substitutes in the cooking oil market, increases in the price of corn oil will decrease the quantities of corn oil demanded, but increase the quantity of soy oil demanded, which in turn increases the soy oil price. As soy oil is a joint product of soybeans and soybean farm-level price is positively related to soy oil price, the soybean farm-level price will increase as a result.

A one percent increase in the variable cost of growing soybeans will increase expected farm-level soybean price by 0.25%. Variable cost increases decreases the soybeans supplied, which increases the soybean prices, assuming a finite negative demand elasticity.

A one percent increase in real expenditures on food will increase the soybean farm-level price by 0.035%. Increase in food expenditures increases the quantities of soy oil demanded, which in turn increases the soy oil price. As the farm-level soybean price is positively related to soy oil wholesale price, the soybean price is also driven up.

A one percent increase in acreage of soybeans of last year decreases the farm-level soybean price this year by 0.07%. Increases in last year’s acreage increase the expected quantity of soybeans supplied, which decreases the expected farm-level soybean price.
A one percent increase in last year’s soybean farm level price will decrease the farm-level soybean price this year by 0.022%. An increase of last year’s soybean price encourages farmers to expand the acreage of soybeans this year, which increases soybeans supplied this year. The soybean farm-level price is then decreased as a result. This is consistent with the cobweb assumptions of Houck, et al.

A one percent increase of last year’s farm-level wheat price will decrease farm-level soybean price by 0.04%, which suggests that wheat and soybeans are joint-crops. This result corresponds to the common growing practices that soybeans and wheat are double cropped in the same rotation system.

**Forecast Soybean Prices**

With the reduced form of farm-level soybean price

\[
PS_t = 3.2 - 0.000013A_{t-1} - 0.022PS_{t-1} + 0.057PCORN_{t-1} - 0.082PW_{t-1} + 0.0008COST_t \\
+ 0.006PCOM_t + 0.000003AF_t + 0.011T_t + 0.000015(SM_t - SM_{t-1}) + 0.000012SX_t \\
+ 0.000026*(SS_t - SS_{t-1}) + 0.012PCOO_t + 0.0011E_t + 0.000088(SO_t - SO_{t-1})
\]

we see that we need forecast some variables on the right-hand-side before we forecast the soybean farm-level price. These variables include: variable cost of growing soybeans (COST), wholesale corn meal price (PCOM), animal units (AF), South American soybean exports (SX), price of corn oil (PCOO), and real expenditures on food (E). For simplicity, we assume that forecasts are naïve for these variables, that is, the forecast soybean price in 1999 is based on 1998 information. With this assumption, we forecast the farm-level price of soybeans from 1991 to 1998 and the results are presented in
Figure 4.1. Note that both the observed and forecast prices are real prices adjusted for inflation with CPI (1999=100).

**Figure 4.1: Forecast Farm-Level Soybean Price From 1991 to 1998**

Applying the Forecast Results to Canola Prices

Recall in Chapter One, we found the relationship between Canola and soybean prices as:

\[ P_{sc,t} = 0.974P_{s,t} \]

in which \( P_{sc} \) is the Canola farm-level price and \( PS \) is the soybean farm-level price. Nearly 99% of the variation in Canola farm-level price can be explained by the soybean farm-level price; if we can explain how the soybean farm-level price moves, we can
explain how the Canola farm-level price moves. We can substitute the forecast result for soybeans into the relationship between Canola and soybean prices; thus, each parameter in the soybean price reduced form is proportioned by 0.974. Figure 4.2 presents the Canola forecast price from 1991 to 1998. Observed Canola price of the same years are also presented in the same figure. Note that prices here are real prices adjusted with CPI (1999=100).

**Figure 4.2: Forecast Canola Price from 1991 to 1998**

Our results suggest that the forecast results farm-level soybean price and Canola price roughly captures the movements of the two prices.
Summary

In this chapter, we estimated a simultaneous system of soybeans using two-stage-least-squares estimate method. We find some variables that significantly influence the soybean market and the supply of soybeans. By substituting the parameter estimates into the reduced form of farm-level soybean price, we are able to forecast soybeans price and the Canola prices. Our forecast results suggest that with economic structure, we can roughly capture the movements for soybean price from 1991 to 1998.
Chapter 5
SUMMARY AND CONCLUSIONS

Summary

The focus of this study is to identify economic factors that influence soybean and Canola prices. These results are expected to be used to answer whether it is feasible to grow soybeans or Canola as rotation crops in a potato rotation system. Because the United States has limited data on Canola, we attempt to characterize the determinants of Canola prices by studying soybean prices. The reason for this is that Canola prices and soybean prices are highly correlated. Through analysis, we find that Canola seed price is about 97% soybean price; Canola oil price is about 1.15 times the soy oil price and Canola meal price is about 66% soy meal price. We consider an economic structure to study how the soybean prices are affected by economic factors.

Houck, Ryan, and Subotnik established a famous USDA CROPS Model for the U.S. soybean and its products market in the late 1960’s. This model basically studies the soybean market from the demand side. Market supply of soybeans is considered as an exogenous variable. Houck, et al. imposed a cobweb model for soybeans prices in their model. Primarily based on the CROPS Model, we made some necessary modifications and established our own model for this thesis. In our model, the stock demands for soybeans, soy meal and soy oil are considered constant due to the fact that the USDA commodity policies regarding inventories have been far less active than they were for the period of time when Houck, et al, established his model (1946 to 1966). Stock demand is
used in the identity equation to ensure that the supply of soybeans equals the demand. By a two-stages-least-squares estimation method, we obtain the parameters for a reduced form for soybean price. With these results and the former relationship between soybean price and Canola price, we are able to explain how the canola price is affected by economic factors and to forecast Canola price.

**Conclusions**

Several conclusions can be drawn from this study. First, soybean prices and Canola prices are highly correlated and move closely. Second, we can use a simultaneous equation system for studying the farm-level soybean prices. We find that soybean prices are positively affected by: animal units (AF), wholesale corn oil price (PCOO), and real expenditures on food (E), time trend variable (T), the variable cost of growing soybeans and stock changes of soybeans, soybean oil and soy meal. Farm-level price of soybeans is negatively affected by one-year lagged farm level soybean price, one-year lagged farm-level wheat price and one-year lagged soybean acreage. Third, assume that the relationship of these economic factors between soybean farm-level price also hold for Canola, with the relationship of Canola and soybean price we obtained in Chapter One, we can explain how Canola prices are affected by these same economic factors. Finally, based on one-year lagged information of the economic variables on the right hand sides of the reduced form we obtain for soybeans, we can forecast soybeans and Canola future prices. Our forecast results suggest that with economic structure, we may roughly capture the price movements of soybean price and Canola prices.
REFERENCES


Frederick, R. “South American Competition Likely Here to Stay” Department of Agricultural Economics, University of Nebraska-Lincoln http://ianrhome.unl.edu/static/0105030.shtml. (May 03, 2001)


USDA. Agricultural Statistics. 1960 to 1990


Appendix A

FORMULATION OF DISTRIBUTED LAG MODELS

In this Appendix, we first review some of the expectation theories. We then use Polynomial distributed lag structures to study the soybean price movements. We will present the results of forecasting annual soybean prices from 1970 to 1998 by using second degree, 4-year lag polynomial distributed lag structures.

Expectation Theories

Extrapolative Expectations

In 1941, Metzler presented a model using a coefficient of expectation, as an alternative to a naive expectation model. His model later was used by Goodwin (1947) and subsequently analyzed by Enthoven and Arrow (1956). Their work developed the theory of extrapolative expectation, for which they defined the expected price as

\[ \hat{p}^* = p_{t-1} + \eta(p_{t-1} - p_{t-2}), \]

where \( \hat{p}^* \) is the expected price for period \( t \) at period \( t-1 \), \( p_{t-1} \) the observed price in period \( t-1 \), \( p_{t-2} \) the observed price in period \( t-2 \), and \( \eta \) Metzler’s coefficient of expectation.

The purpose of the extrapolative expectations model is to modify the cobweb theory to take into account the most recent trends in prices. Metzler compares his coefficient of expectation to Hick’s elasticity of expectation and concludes that his coefficient of expectation plus one is exactly Hick’s elasticity. However, this comparison lacks theoretical appeal.
Adaptive Expectations

Based on Hicks’ definition of elasticity of expectation, the adaptive expectations models first appear in Cagan (1956). Under the adaptive expectations hypothesis, the individuals are assumed to revise their expectations according to their most recent experience:

$$P^*_t - P^*_{t-1} = \beta (P_{t-1} - P^*_{t-1})$$

where $P^*$ is the expected price for period $t$ at period $t-1$, $P^*_{t-1}$ the expected price for period $t-1$ at period $t-2$, $P_{t-1}$ the observed price in period $t-1$ and $\beta$ the coefficient of expectation and $0<\beta<1$. Rearrange the terms of both sides and we can get:

$$P^*_t - (1-\beta)P^*_{t-1} = \beta P_{t-1}.$$  

Replacing $(1-\beta)$ by $\lambda$ and introducing the lag operator $V$ so that $V^j x_t = x_{t-j}$, we obtain

$$(1-\lambda V)P^*_t = (1-\lambda)P_{t-1},$$

so that

$$P^*_t = [(1-\lambda)/(1-\lambda V)]P_{t-1}.$$  

So long as $-1<\lambda<1$, we can expand $1/(1-\lambda V)$ as $1+\lambda V+\lambda^2 V^2+\ldots$ and thus express $P^*_t$ as

$$P^*_t = (1-\lambda)\sum_{k=0}^{\infty} \lambda^k P_{t-1-k}.$$  

By transformation, the expected price in adaptive expectation can be expressed by an infinite weighted average of past-realized prices with weights that decline geometrically with the lag.

Nerlove (1958) developed the idea of an expected price using the adaptive expectations model. Adaptive expectations have been popular for their simplicity since maximum-likelihood estimates for $\lambda$ can be obtained easily and because such models appear to work well in a number of empirical studies.
Although adaptive expectation is widely used, there are still several unsolved problems that make the concept questionable. The first problem is that there is no real theoretical justification for the model. Besides, much of the criticism of the adaptive expectation theory has to do with its implication of a geometrically decaying lag structure.

**Muth’s Rational Expectations**

As we have indicated above that the extrapolative expectations and adaptive expectations models lack theoretical justification. Muth developed an expectations model that eliminates the theoretical weakness common to the previous theories of expectation formation.

Muth’s theory is based on three hypotheses about individual behavior:

“1). Information is scarce, and the economic system generally does not waste it. 2). The way expectations are formed depends specially on the structure of the relevant system describing the economy. 3) A ‘public prediction,’ in the sense of Grunberg and Modigliani…, will have no substantial effect on the operation of the economic system (unless it is based on inside information)” (Muth, p.3 16)

Muth’s theory implies that expectations are based on information, which is assumed to be costless to obtain and to be generated according to perceptible forces. For example, if a producer operating under free competition has some idea of market conditions, he will use the information available to him about demand and supply conditions in generating his expectations about the relevant variables for decision purposes.

By applying his theory to price expectations, Muth introduced three simplifying assumptions: random disturbances are normal; certainty equivalents exist for the variable
to be predicted; and the equations of the system, including the expectation formulas are linear. He assumed that the market equations take the form

\begin{align*}
P_D &= -\beta p_t \quad (1) \\
P_S &= \gamma p^e_t + e_t \quad (2) \\
P_D &= P_S \quad (3)
\end{align*}

Where \( P_S \) represents the number of units produced in a period lasting as long as the production lag, \( p_t \) is the market price in the \( t \)th period, \( p^e_t \) is the market price expected to prevail during the \( t \)th period on the basis of information available through the \((t-1)\)'s \( t \) period, \( e_t \) is an error term which represents all kinds of undetermined variations in yields.

By solving the equilibrium of supply and demand, Muth generated the relationship between \( p_t \), market price in the \( t \)th period and \( p^e_t \), the market price expected to prevail during the \( t \)th period on the basis of information available through the \((t-1)\)'s \( t \) period, in the form of:

\[ p_t = (-\gamma/\beta)p^e_t - (1/\beta)e_t. \]

The error term is unknown at the time the production decisions are made, but it is known at the time the commodity is purchased in the market. The prediction of the model is found by replacing the error term by its expected value, conditional on past events. Assume there is no serial correlation among errors, we have the expected value of the error term, \( E(e_t) = 0 \) and we can obtain the expected value of price in the \( t \)th period

\[ E(p_t) = (-\gamma/\beta)p^e_t. \]

This result is of great importance because if the predictions of the theory were substantially better than the expectations of the firms, then there would be opportunities for the “insider” to profit from the knowledge, by possible inventory speculation, or by
selling a price forecasting service to the firms. The profit opportunities would no longer exist if the aggregate expectation of the firms is the same as the prediction of the theory, which is:

\[ E(p_t) = E(p_t^n) \]

From an economic view of point, Muth’s rational expectations, compared with other expectation theories, are more consistent with the underlying structure of economic behavior, while alternative models of expectations (cobweb, extrapolative, and adaptive) are not necessarily compatible with the economic behavior implied by the underlying economic structure. If forecast efficiency is the criterion, rational expectations are always empirically better than any other expectation theories (Wallis). Rational expectations are attractive in that economic behavior is directly incorporated in their definition, and it is that makes expectations depend on the parameters of the model itself.

Theoretically satisfactory as it is, rational expectations also present problems in use. From the technical viewpoint, rational expectations are more difficult to estimate than other alternatives. To obtain rational expectations in our model, it must be solved for the expected values of the uncertain variables. Although this solution is linear in the exogenous variables, the coefficients are combinations of the structural parameters that are generally not linear. The identification problem is also associated with rational expectations and poses a difficulty for using rational expectations models. A third possible problem with rational expectations is the evidence of serial correlation in the structural disturbance. Imposing the assumption that random disturbances are not enough to simplify estimation. If there is serial correlation, then the estimate of the error term \((e)\) is not rational.
Almon Polynomial Distributed Lag

Due to the difficulty of estimation or identification of the expectation theories we discussed above, we will focus on polynomial distributed lags in this part and investigate its significance to our study.

Almon used nonsample information about the distributed lag weights to improve the precision of estimation. In her study of the relationship between capital appropriation and capital expenditures, Almon assumed that a smooth pattern of lag weights could be approximated by a polynomial of relatively low order. Her insight indicates that the lag weights don’t necessarily decline geometrically with time, but can be specified by a continuous function.

The general polynomial structure can be expressed as

\[ Y_t = \alpha + \beta (w_1 X_{t-1} + w_2 X_{t-2} + \ldots + w_t X_{t-n}) + \varepsilon_t, \quad t = n \ldots T \]

where, \( w \) represents polynomial distributed lags weights in the function of the lag index-variable \( u \),

\[ w_u = \lambda_0 + \lambda_1 u + \lambda_2 u^2 + \lambda_3 u^3 + \cdots + \lambda_q u^q, \]

in which, \( u \) is the lag index (e.g. 1 lag, \( u=0 \), 2 lags, \( u=1 \)), \( q \) is the degree of polynomial.

The error term \( \varepsilon_t \) is assumed to be normal distributed with zero means.

Assume a second degree polynomial (\( q=2 \)) distributed lag structure. The weight is then represented as:

\[ w_u = \lambda_0 + \lambda_1 u + \lambda_2 u^2 \quad u=0, 1, 2, \ldots, n \]

We have considerable flexibility involved in specifying a polynomial distributed lag model. Usually a third-degree or a fourth degree polynomial will provide a sufficiently accurate approximation to the lag structure. Almon assumed that a second-degree
A polynomial distributed lag can sufficiently approximate the lag structure in her study of capital appropriations. The choice of length of lag depends more on the nature of the problem being specified, so that useful rules of thumb are not available. (Pindyck, 211)

A polynomial distributed lag provides a more flexible lag structure. And the decision regarding the lag length is a subjective decision. We can choose appropriate order and lags for the convenience of estimation. So it is a feasible model for our empirical study in this thesis.

**A Pure Time Series Analysis of Soybean Prices**

We assume that a second degree, four-year lag polynomial distributed lag structure sufficiently approximates the lag structures for studying soybean prices. In the polynomial distributed lag structure in our study, we will define the weights as

\[ w_t = \lambda_0 + \lambda_1 t + \lambda_2 t^2, \quad (t=0,1,2,3), \]

in which 

\[ t \] is the lag index (e.g. 1 lag, \( t=0 \), 2 lags, \( t=1 \)) and \( w \) is the weight.

The soybean prices can be expressed as

\[ PS_t = \alpha_1 + \beta_{11} \lambda_{03} PS_{t-1} + \beta_{21} (\lambda_{03} + \lambda_{13} + \lambda_{23}) PS_{t-2} \]

\[ + \beta_{31} (\lambda_{03} + 2\lambda_{13} + 4\lambda_{23}) PS_{t-3} + \beta_{41} (\lambda_{03} + 3\lambda_{13} + 9\lambda_{23}) PS_{t-4} + e_1 \]

\[ PM_t = \alpha_2 + \beta_{12} \lambda_{02} PM_{t-1} + \beta_{22} (\lambda_{02} + \lambda_{12} + \lambda_{22}) PM_{t-2} \]

\[ + \beta_{32} (\lambda_{02} + 2\lambda_{12} + 4\lambda_{22}) PM_{t-3} + \beta_{42} (\lambda_{02} + 3\lambda_{12} + 9\lambda_{22}) PM_{t-4} + e_2 \]

\[ PO_t = \alpha_3 + \beta_{13} \lambda_{01} PO_{t-1} + \beta_{23} (\lambda_{01} + \lambda_{11} + \lambda_{21}) PO_{t-2} \]

\[ + \beta_{33} (\lambda_{01} + 2\lambda_{11} + 4\lambda_{21}) PO_{t-3} + \beta_{43} (\lambda_{01} + 3\lambda_{11} + 9\lambda_{21}) PO_{t-4} + e_3 \]

---

9 An F-test can be employed to test the length of lags (Judge et al. 1982)
in which,

**PS:** farm-level soybean bean price

**PM:** wholesale soybean meal price

**PO:** wholesale soybean oil price.

Assume that all the error terms are normally distributed with zero means and constant variances.

We use the annual data (1965 to 1998) from USDA Agricultural Statistics for regressions of the above models. The summary statistics for these data are reported in Table A.1.

<table>
<thead>
<tr>
<th></th>
<th>Total Observations</th>
<th>Mean ($/metric ton)</th>
<th>Std. Deviation ($/metric ton)</th>
<th>Maximum ($/metric ton)</th>
<th>Minimum ($/metric ton)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soy Oil</td>
<td>34</td>
<td>42.44</td>
<td>18.65</td>
<td>103.78</td>
<td>20.32</td>
</tr>
<tr>
<td>Soy Meal</td>
<td>34</td>
<td>16.46</td>
<td>6.46</td>
<td>43.47</td>
<td>7.07</td>
</tr>
<tr>
<td>Soybeans</td>
<td>34</td>
<td>10.8</td>
<td>3.78</td>
<td>19.25</td>
<td>5.03</td>
</tr>
</tbody>
</table>

Using the **SAS** procedure PDLREG, we estimate the above three models. The results as provide in Table A.2.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Soy Oil Price</th>
<th>Soy Meal Price</th>
<th>Soybean Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.27 (1.29)</td>
<td>-0.0627 (1.21)</td>
<td>-0.086 (0.44)</td>
</tr>
<tr>
<td>One-year Lag</td>
<td>0.919805 (0.0201)</td>
<td>0.934311 (0.0378)</td>
<td>0.955366 (0.0363)</td>
</tr>
<tr>
<td>Two-year Lag</td>
<td>0.140771 (0.0414)</td>
<td>0.144876 (0.0271)</td>
<td>0.147208 (0.0276)</td>
</tr>
<tr>
<td>Three-year Lag</td>
<td>-0.138571 (0.0144)</td>
<td>-0.144133 (0.0273)</td>
<td>-0.152493 (0.0277)</td>
</tr>
<tr>
<td>Four-year Lag</td>
<td>0.081778 (0.0206)</td>
<td>0.067285* (0.0386)</td>
<td>0.056261* (0.0382)</td>
</tr>
<tr>
<td>Regression $R^2$</td>
<td>0.9939</td>
<td>0.9742</td>
<td>0.9872</td>
</tr>
<tr>
<td>Durbin-Watson</td>
<td>2.82</td>
<td>3.52</td>
<td>3.53</td>
</tr>
</tbody>
</table>
Given a 5% level, all parameter estimates are significant except for four-year lag terms for soy meal and soybeans. Also, Durbin-Watson values indicate that autocorrelation does not seem to be a problem for these estimators. As all of the regression $R^2$s are as high 99%, we can say that nearly all of the variation of the left-hand-side variables is explained by the variations in the right-hand-side variables.

From the above estimation results, we can weigh the soybean prices on their four-year lag prices as shown below.

\[
\begin{align*}
\text{PO}_t &= -0.27 + 0.9198\text{PO}_{t-1} + 0.1408\text{PO}_{t-2} - 0.1386\text{PO}_{t-3} + 0.082\text{PO}_{t-4} \\
\text{PM}_t &= -0.0627 + \text{PM}_{t-1} + 0.1449\text{PM}_{t-2} - 0.1441\text{PM}_{t-3} + 0.0673\text{PM}_{t-4} \\
\text{PS}_t &= -0.086 + 0.9554\text{PS}_{t-1} + 0.1472\text{PS}_{t-2} - 0.1525\text{PS}_{t-3} + 0.056\text{PS}_{t-4}
\end{align*}
\]

(A.1) (A.2) (A.3)

Note that the forecast procedure and results are similar for the three prices, we will only present soybean price analysis in this appendix.

Based on the estimation results, we use equation (A.3) to forecast soybean prices from 1990 to 1998. The forecast results indicate only slight variations and provides little insight into future price. Figure A.1 presented the forecasting results for farm-level soybean bean prices.

Since forecast results indicated that pure time series analysis doesn’t capture the structure of movements of soybean prices. It is necessary to consider an economic structure to study the economic factors that influence the soybean prices.
Figure A.1: Forecast Real Soybean Prices from 1990 to 1998 With Polynomial Distributed Lag Structure

(Sources: USDA Agricultural Statistics, various years)
(Prices are adjusted with CPI 1999=100)
Appendix B

Identification in Simultaneous System and Two-Stage Least Square Estimates

Counting Rules and Rank Rules

Counting rules (Order rules) state that in the \( t \)th equation of a simultaneous system, if the sum of the number of endogenous variables in this equation and the number of exogenous variables is less than the total number of exogenous variables in the whole system plus 1, the equation is over-identified; if it equals to the total number of exogenous variables plus 1, the equation is just identified; if it is larger than the total number of exogenous variables plus 1, the equation is under-identified (Griffiths, et al.).

Recall that the simultaneous equation system in our model is:

\[
\begin{align*}
QMD_t &= \alpha_{11} + \alpha_{12} PM_t + \alpha_{13} PCOM_t + \alpha_{14} AF_t + \alpha_{15} T + \alpha_{16} e_1, \\
QOD_t &= \alpha_{21} + \alpha_{22} PO_t + \alpha_{23} PCOO_t + \alpha_{24} E_t + \alpha_{25} e_2, \\
QSX_t &= \alpha_{31} + \alpha_{32} PS + \alpha_{33} PM_t + \alpha_{34} SX_t + \alpha_{35} e_3, \\
QMX_t &= \alpha_{41} + \alpha_{42} PM_t + \alpha_{43} PCOM_t + \alpha_{44} T + e_4, \\
QOX_t &= \alpha_{51} + \alpha_{52} PM_t + \alpha_{53} PCOO_t + \alpha_{54} T + e_5, \\
QMP_t &= \alpha_{61} QSC_t + \alpha_{61} e_6, \\
QOP_t &= \alpha_{71} QSC_t + \alpha_{71} e_7, \\
PS_t &= \alpha_{81} + \alpha_{82} PM_t + \alpha_{83} PO_t + \alpha_{81} e_8, \\
QSS_t &= QSC_t + QSX_t + SS_t - SS_{t-1}, \\
QMP_t &= QMD_t + QMX_t + SM_t - SM_{t-1}, \\
QOP_t &= QOD_t + QOX + SO_t - SO_{t-1}, \\
A_t &= \alpha_{91} + \alpha_{92} A_{t-1}^* + \alpha_{93} PS^*_{t-1} + \alpha_{94} PCORN^*_{t-1} + \alpha_{95} PW^*_{t-1} + \alpha_{96} COST^*_{t-1} + e_9
\end{align*}
\]
Endogenous variables are QMD, QOD, QSX, QMX, QOX, QOP, QMP, QSC, PS, PM and PO. Other variables are all exogenous variables and the total number of exogenous variables in the system is fifteen. For equation (1), number of endogenous variables in this equation is two; number of exogenous variables in this equation is three. As five \((2+3=5)\) is less than fourteen \((15-1=14)\), equation (1) is over-identified. Applying the same rule to the following equations in the system (except for the identity equation), we see that all the equations in the system are over-identified.

Counting rules are only necessary conditions for identification. Sufficient (rank) conditions require that for the ith equation, no linear combination of the other equations in the simultaneous system can produce the ith equation. In our simultaneous system, this is apparently true. So the simultaneous equations of our model are over-identified.

**Two Stage Least Square Estimates**

Griffin, Hill, and Judge (1993) explained that for the over-identified simultaneous system, indirect least squares and the instrumental variable method do not yield unique estimates; these estimates are consistent but not efficient due to the problem of correlation between the random variables and the error terms in the over-identified system. Two-stage-least-squares estimates may be used to provide a consistent and unbiased estimator. They use an example to state the procedure.

Suppose a statistical model for demand and supply takes the forms:

\[
P_t = \beta_{11} + \gamma_1 Q_t + \beta_{12} P_{st} + \beta_{13} \text{dit} + \epsilon_{1t}, \text{and}
\]

\[
Q_t = \beta_{21} + \gamma_2 P_t + \beta_{24} P_{ft} + \epsilon_{2t}.
\]
Let $P_t = y_{1t}$, $Q_t = y_{2t}$, $ps_t = x_{2t}$, $dit = x_{3t}$, $pf_t = x_{4t}$ and $x_{1t}$ represent the intercept variable, they rewrite the equations for a sample of $T$ observations as:

$$y_1 = y_{21}y_{12} + x_1\beta_{11} + x_2\beta_{12} + x_3\beta_{13} + e_1,$$

and

$$y_2 = y_{11}y_{12} + x_1\beta_{21} + x_4\beta_{24} + e_2.$$  

(Example 1)

(Example 2)

They first concentrate on the supply equation (example 2). Since the endogenous variable $y_1$ is correlated with $e_2$ and thus yield biased and inconsistent estimates by using least squares estimator method. From the reduced form equation for $y_1$, they get the least square estimator for parameters in $y_1$, and they substitute the predictions of $y_1$ back into $y_2$, and get the new formulation of $y_2$ that takes forms of:

$$y_2 = Z'\delta_2 + e'_2,$$

where $Z'\delta_2$ is the vector of the right-hand-side variables in $y_2$ after the replacement. By the replacement, a new error term is obtained in $y_2$ ($e'_2$) that depends on the reduced form residuals for $y_1$ and the original error term in $y_2$. The advantage of the new formulation is that when observation sample $T$ becomes very large, new right-hand side variables (variables in vector $Z'\delta_2$) and error terms ($e'_2$) in $y_2$ become uncorrelated. Consequently, the least squares estimator in the new formulation may be used to provide a consistent estimator of $\delta_2$. 
Appendix C

Comparison of OLS Estimates with Two-Stage-Least-Squares Estimates in the Simultaneous Equation Model for The Soybean Market

Recall that the simultaneous equation system for our soybean market is:

\[ \begin{align*}
QMD_t &= \alpha_{11} + \alpha_{12}P_M + \alpha_{13}PCOM_t + \alpha_{14}AF_t + \alpha_{15}T + \varepsilon_1, \\
QOD_t &= \alpha_{21} + \alpha_{22}PO_t + \alpha_{23}PCOO_t + \alpha_{24}E_t + \varepsilon_2, \\
QSX_t &= \alpha_{31} + \alpha_{32}PS + \alpha_{33}PM_t + \alpha_{34}SX_t + \varepsilon_3, \\
QMX_t &= \alpha_{41} + \alpha_{42}PM_t + \alpha_{43}PCOM_t + \alpha_{44}T + \varepsilon_4, \\
QOX_t &= \alpha_{51} + \alpha_{52}PM_t + \alpha_{53}PCOO_t + \alpha_{54}T + \varepsilon_5, \\
QMP_t &= \alpha_{61}QSC_t + \varepsilon_6, \\
QOP_t &= \alpha_{71}QSC_t + \varepsilon_7, \\
PS_t &= \alpha_{81} + \alpha_{82}PM_t + \alpha_{83}PO_t + \varepsilon_8, \\
QSS_t &= QSC_t + QSX_t + SS_t - SS_{t-1}, \\
QMP_t &= QMD_t + QMX_t + SM_t - SM_{t-1}, \\
QOP_t &= QOD_t + QOX_t + SO_t - SO_{t-1}, \quad \text{and} \\
A_t &= \alpha_{91} + \alpha_{92}A^* + \alpha_{93}PS^* + \alpha_{94}PCORN^* + \alpha_{95}PW^* + \alpha_{96}COST^* + \varepsilon_9.
\end{align*} \]


If we had not written equations in the simultaneous system, but focused only on each equation, it would be natural to use an OLS rule to estimate the parameter coefficients for the behavioral equations. Table C.1 presents the comparison of the OLS estimates vs. Two-Stage-Least-Squares estimates.

68
Table C.1: Compare OLS Estimates and Two-Stage Least Square Estimates for Equations in the Simultaneous System In Our Model (Numbers in the Parentheses Are Standard Errors)

<table>
<thead>
<tr>
<th></th>
<th>OLS Estimates</th>
<th>2SLS Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Constant</strong></td>
<td>2280</td>
<td>1571</td>
</tr>
<tr>
<td></td>
<td>(10743)</td>
<td>(10945)</td>
</tr>
<tr>
<td><strong>Price of Soy Meal (PM)</strong></td>
<td>-118</td>
<td>-167</td>
</tr>
<tr>
<td></td>
<td>(157)</td>
<td>(209)</td>
</tr>
<tr>
<td><strong>Price of Corn Meal (PCOM)</strong></td>
<td>158</td>
<td>230</td>
</tr>
<tr>
<td></td>
<td>(264.9)</td>
<td>(332)</td>
</tr>
<tr>
<td><strong>Time Trend Variable (T)</strong></td>
<td>574.94</td>
<td>579*</td>
</tr>
<tr>
<td></td>
<td>(167.35)</td>
<td>(168)</td>
</tr>
<tr>
<td><strong>Number of Animal Units (AF)</strong></td>
<td>0.176</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>(0.195)</td>
<td>(0.19)</td>
</tr>
<tr>
<td><strong>Adjusted R²</strong></td>
<td>0.8909</td>
<td>0.89</td>
</tr>
</tbody>
</table>

**Domestic Demand for Soy Oil (QOD)**

<table>
<thead>
<tr>
<th></th>
<th>OLS Estimates</th>
<th>2SLS Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Constant</strong></td>
<td>6412.14*</td>
<td>6759*</td>
</tr>
<tr>
<td></td>
<td>(1089.97)</td>
<td>(1210)</td>
</tr>
<tr>
<td><strong>Price of Soy Oil (PO)</strong></td>
<td>-21.2</td>
<td>-31.2</td>
</tr>
<tr>
<td></td>
<td>(20)</td>
<td>(25)</td>
</tr>
<tr>
<td><strong>Price of Corn Oil (PCOO)</strong></td>
<td>68.4</td>
<td>86.27</td>
</tr>
<tr>
<td></td>
<td>(48.64)</td>
<td>(55)</td>
</tr>
<tr>
<td><strong>Real Expenditures on Food (E)</strong></td>
<td>14.37*</td>
<td>13.28*</td>
</tr>
<tr>
<td></td>
<td>(2.45)</td>
<td>(2.95)</td>
</tr>
<tr>
<td><strong>Adjusted R²</strong></td>
<td>0.9017</td>
<td>0.90</td>
</tr>
<tr>
<td>Table C1: (Continued)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>U.S. Soybean Export</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>41681* (11212)</td>
<td>37836* (11889)</td>
</tr>
<tr>
<td>Price of Soybeans (PS)</td>
<td>-558.92 (729.14)</td>
<td>-242 (838)</td>
</tr>
<tr>
<td>Price of Soy Meal (PM)</td>
<td>29.05 (355.76)</td>
<td>-26 (398)</td>
</tr>
<tr>
<td>South American Soybean Export (SX)</td>
<td>0.2155 (0.1849)</td>
<td>0.26 (0.19)</td>
</tr>
<tr>
<td>R²</td>
<td>0.2274</td>
<td>0.21</td>
</tr>
</tbody>
</table>

| **U.S. Soy Meal Export Demand (QMX)** |
| Constant | 2923.12 (1930) | 2712 (2006) |
| Price of Soy Meal (PM) | -68.84 (81.6) | -97.16 (108.36) |
| Price of Corn Meal (PCOM) | 141.52 (138.31) | 182.25 (173.27) |
| Time Trend Variable (T) | 113.88* (48.42) | 119.11* |
| Adjusted R² | 0.248 | 0.29 |

| **(QOX)** |
| Constant | 975.66* (442.92) | 998* (444) |
| Price of Soy Oil (PO) | -9.06** (4.85) | -10.87* (5.05) |
| Price of Corn Oil (PCOO) | 45.84* (18.19) | 47.9* (18.3) |
| Adjusted R² | 0.17 | 0.21 |

| **Soy Meal Produced (QMP)** |
| QMP | 0.48628* | 0.49* |

| QOP | 0.2096* (0.00186) | 0.21* (0.0016) |
| Adjusted R² | 0.9978 | 0.99 |
### Table

<table>
<thead>
<tr>
<th>Parameter</th>
<th>OLS Estimate</th>
<th>2SLS Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>84859*</td>
<td>84858*</td>
</tr>
<tr>
<td></td>
<td>(12577)</td>
<td>(12576)</td>
</tr>
<tr>
<td>Acreage</td>
<td>0.245*</td>
<td>0.24*</td>
</tr>
<tr>
<td></td>
<td>(0.109)</td>
<td>(0.10)</td>
</tr>
<tr>
<td>Soybean Price</td>
<td>434.1957*</td>
<td>443.1958*</td>
</tr>
<tr>
<td></td>
<td>(162)</td>
<td>(162.14)</td>
</tr>
<tr>
<td>Wheat Price</td>
<td>1580.798*</td>
<td>1580.79*</td>
</tr>
<tr>
<td></td>
<td>(487.86)</td>
<td>(487.86)</td>
</tr>
<tr>
<td>Corn Price</td>
<td>-1103.55*</td>
<td>-1103*</td>
</tr>
<tr>
<td></td>
<td>(763.7)</td>
<td>(763)</td>
</tr>
<tr>
<td>Variable Cost</td>
<td>-14.78863*</td>
<td>-14.7886*</td>
</tr>
<tr>
<td></td>
<td>(2.477)</td>
<td>(2.47)</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.8997</td>
<td>0.90</td>
</tr>
</tbody>
</table>

Comparison of the OLS and **2SLS** parameter estimates indicates that OLS estimates overestimate the parameter coefficients, but underestimate the intercepts.
Appendix D

Procedure of Obtaining the Reduced Form for Farm-level Soybean Prices (PS)

Recall that the simultaneous equations for soybean and its product market are listed as follows:

\[
\begin{align*}
Q_{MDt} &= \alpha_{11} + \alpha_{12}P_{Mt} + \alpha_{13}P_{COMt} + \alpha_{14}A_{Ft} + \alpha_{15}T + \epsilon, \\
Q_{ODt} &= \alpha_{21} + \alpha_{22}P_{Ot} + \alpha_{23}P_{COOt} + \alpha_{24}F_{t} + \epsilon, \\
Q_{SXt} &= \alpha_{31} + \alpha_{32}PS + \alpha_{33}PM_{t} + \alpha_{34}SX_{t} + \epsilon, \\
Q_{MXt} &= \alpha_{41} + \alpha_{42}PM_{t} + \alpha_{43}P_{COMt} + \alpha_{44}T + \epsilon, \\
Q_{OXt} &= \alpha_{51} + \alpha_{52}PM_{t} + \alpha_{53}P_{COOt} + \alpha_{54}T + \epsilon, \\
Q_{MPt} &= \alpha_{61}Q_{St} + \epsilon, \\
Q_{OPt} &= \alpha_{71}Q_{Sc} + \epsilon, \\
PS_{t} &= \alpha_{81} + \alpha_{82}P_{Mt} + \alpha_{83}P_{Ot} + \epsilon, \\
Q_{SS_{t}} &= Q_{Sc} + Q_{SX_{t}} + S_{St} - S_{St-1}, \\
Q_{MP_{t}} &= Q_{MD_{t}} + Q_{MX_{t}} + S_{Mt} - S_{Mt-1}, \\
Q_{OP_{t}} &= Q_{OD_{t}} + Q_{OX_{t}} + S_{Ot} - S_{Ot-1}, \text{ and} \\
A_{t} &= \alpha_{91} + \alpha_{92}A_{t-1} + \alpha_{93}PS_{t-1} + \alpha_{94}P_{COM_{t-1}} + \alpha_{95}PW_{t-1} + \alpha_{96}COST_{t} + \epsilon.
\end{align*}
\]

To derive the reduced form equation for PS, we first make use of the relationships between QMP and QSC, QOP and QSC, and obtained

\[
-(\alpha_{12} + \alpha_{52} + \alpha_{61}\alpha_{34})/\alpha_{61}P_{Mt} - \alpha_{32}PS_{t} = -Q_{SS_{t}} + (\alpha_{11} + \alpha_{51} + \alpha_{6}A_{31})/\alpha_{61} + (\alpha_{13} + \alpha_{53})/\alpha_{61}P_{COM_{t}} + \\
(\alpha_{15} + \alpha_{54})/\alpha_{61}T_{t} + (\alpha_{46}/\alpha_{61})A_{F_{t}} + \\
+1/\alpha_{61}(S_{Mt} - S_{Mt-1}) + \alpha_{35}SX_{t}
\]

and
\[-(\alpha_{22} + \alpha_{42})/\alpha_{71}POt - \alpha_{32}PS_r - \alpha_{34}PM_r = -QSS_t + (\alpha_{21} + \alpha_{41} + \alpha_{71}\alpha_{31})/\alpha_{71} + (\alpha_{23} + \alpha_{43})/\alpha_{71}PCOO_t + (\alpha_{14}/\alpha_{61})E_t + (1/\alpha_{71})(SO_t - SO_{t-1}) + \alpha_{35}SX_t + (SS_t - SS_{t-1}).\]

These two equations, together with equation (D.8) will result the reduced form for PS, which is shown in the following equation

\[PS_r = m + n(\alpha_{21} + \alpha_{41} + \alpha_{71}\alpha_{31})/\alpha_{71} - g(\alpha_{11} + \alpha_{51} + \alpha_{61}\alpha_{31})/\alpha_{61} - (n-g)QSS_t + n(\alpha_{23} + \alpha_{43})/\alpha_{71}PCOO_t + n(\alpha_{24}/\alpha_{71})E + n/\alpha_{61}(SO_t - SO_{t-1}) + \alpha_{35}(n-g)SX_t + n(SS_t - SS_{t-1}) - g(\alpha_{13} + \alpha_{53})/\alpha_{61}PCOM_t - g(\alpha_{15} + \alpha_{54})/\alpha_{61}T_r - g(\alpha_{14}/\alpha_{61})A{Fr},\]

in which,

\[m = \alpha_{81} + (d/a)\alpha_{82}\alpha_{32}\alpha_{83} - (d/a)b\alpha_{82} - (b/a)d\alpha_{83} - (dc/a)\alpha_{83}\]
\[n = (d/a)\alpha_{82}\alpha_{83}\]
\[g = (b\alpha_{82} + c\alpha_{83})/a\]

where,

\[a = ((\alpha_{22} + \alpha_{42} + \alpha_{71}\alpha_{32}\alpha_{82})(\alpha_{12} + \alpha_{52} + \alpha_{61}\alpha_{34} + \alpha_{61}\alpha_{32}\alpha_{82}) - \alpha_{71}\alpha_{61}\alpha_{32}\alpha_{83}(\alpha_{32}\alpha_{82} + \alpha_{34})) / \alpha_{71}\alpha_{61}\]
\[b = (\alpha_{22} + \alpha_{42} + \alpha_{71}\alpha_{32}\alpha_{82}) / \alpha_{71}\]
\[c = \alpha_{32}\alpha_{82} + \alpha_{34}\]
\[d = \alpha_{32}\alpha_{82}\]
Lei Cui was born in Nanjing, Jiangsu, P.R. China on July 6, 1975. She graduated from the Affiliated Middle School of Nanjing Normal University in 1993 and attended Beijing Normal University in the same year. She received her Bachelor of Science in Economics in July 1997. She was enrolled for graduate study in agricultural production and marketing at the University of Maine in September 1999 and served as a research assistant in the Department of Resource Economics and Policy.

Lei is a candidate for a Master of Science degree in Resource Economics and Policy from the University of Maine in August, 2001.