9-1980

An econometric-programming model for agricultural policy analysis

Center for Agricultural and Rural Development, Iowa State University

Wen-yuan Huang
*Iowa State University*

Reuben N. Weisz
*Iowa State University*

Earl O. Heady
*Iowa State University*

Follow this and additional works at: [http://lib.dr.iastate.edu/card_reports](http://lib.dr.iastate.edu/card_reports)

Part of the Agricultural and Resource Economics Commons, Agricultural Economics Commons, Agricultural Science Commons, and the Natural Resources Management and Policy Commons

Recommended Citation

Center for Agricultural and Rural Development, Iowa State University; Huang, Wen-yuan; Weisz, Reuben N.; and Heady, Earl O., "An econometric-programming model for agricultural policy analysis" (1980). CARD Reports. 89.

[http://lib.dr.iastate.edu/card_reports/89](http://lib.dr.iastate.edu/card_reports/89)

This Book is brought to you for free and open access by the CARD Reports and Working Papers at Iowa State University Digital Repository. It has been accepted for inclusion in CARD Reports by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
An Econometric-Programming Model for Agricultural Policy Analysis

CARD Report 95
AN ECONOMETRIC-PROGRAMMING MODEL
FOR AGRICULTURAL POLICY ANALYSIS

by

Wen-yuan Huang
Reuben N. Weisz
Earl O. Heady

CARD REPORT 95

Center for Agricultural and Rural Development
Iowa State University
in Cooperation with Economics, Statistics, and
Cooperatives Services
U.S. Department of Agriculture

September 1980
TABLE OF CONTENTS

INTRODUCTION

I. THE IMPLICATIONS OF STRUCTURAL DIFFERENCES BETWEEN ECONOMETRIC AND PROGRAMMING MODELS

   The Basic Structure of the Econometric Model

   Basic Structure of the Programming Model

   Some Implications

      Economic basis

      Statistical inference

      Modeling flexibility

      Forecasting and structure analysis

      Positive and normative

      Intraregional and interregional analysis

II. NEEDS FOR AND CASES OF HYBRID MODELS

III. SIX METHODS OF COMBINING AN ECONOMETRIC AND A PROGRAMMING MODEL

      The One-Way Communication Model

      The Simultaneous Solution Model

      Recursive Interactive Programming (RIP) Models

         Characteristic of RIP models

      Recursive Adaptive Programming (RAP) Model

         Alternative Feedback Adjustment Producers for RAP Models

IV. MATHEMATICAL FORMULATION AND AN EXPERIMENT TO ILLUSTRATE THE ESSENCE OF HYBRID MODELS
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathematical Exposition</td>
<td></td>
<td>35</td>
</tr>
<tr>
<td>An Experiment to Investigate Characteristics of Hybrid Models</td>
<td></td>
<td>38</td>
</tr>
<tr>
<td>V. NATIONAL AGRICULTURAL ECONOMETRIC MODELS: A SURVEY</td>
<td></td>
<td>43</td>
</tr>
<tr>
<td>Descriptive Survey of the Models</td>
<td></td>
<td>44</td>
</tr>
<tr>
<td>Summary</td>
<td></td>
<td>49</td>
</tr>
<tr>
<td>Selection of an Econometric Model</td>
<td></td>
<td>50</td>
</tr>
<tr>
<td>Theoretical basis</td>
<td></td>
<td>51</td>
</tr>
<tr>
<td>Complementary to programming model</td>
<td></td>
<td>51</td>
</tr>
<tr>
<td>Assessibility</td>
<td></td>
<td>52</td>
</tr>
<tr>
<td>Model's life</td>
<td></td>
<td>52</td>
</tr>
<tr>
<td>Operation and maintenance costs</td>
<td></td>
<td>52</td>
</tr>
<tr>
<td>Empirical validation</td>
<td></td>
<td>53</td>
</tr>
<tr>
<td>VI. NATIONAL AGRICULTURAL PROGRAMMING MODEL: A SURVEY OF CARD-LP MODELS</td>
<td></td>
<td>54</td>
</tr>
<tr>
<td>Descriptive Survey of Benchmark Programming Models</td>
<td></td>
<td>55</td>
</tr>
<tr>
<td>Summary</td>
<td></td>
<td>59</td>
</tr>
<tr>
<td>VII. THE CARD-NRED RECURSIVE ADAPTIVE PROGRAMMING MODEL</td>
<td></td>
<td>62</td>
</tr>
<tr>
<td>Econometric Component</td>
<td></td>
<td>62</td>
</tr>
<tr>
<td>Programming Component</td>
<td></td>
<td>64</td>
</tr>
<tr>
<td>Linkage Variables and Feedback Mechanism</td>
<td></td>
<td>70</td>
</tr>
<tr>
<td>VIII. MODEL VALIDATIONS</td>
<td></td>
<td>74</td>
</tr>
<tr>
<td>Test Model</td>
<td></td>
<td>74</td>
</tr>
<tr>
<td>Test methods</td>
<td></td>
<td>75</td>
</tr>
</tbody>
</table>
IX. SUMMARY AND SUGGESTIONS

REFERENCES CITED
**LIST OF TABLES**

<table>
<thead>
<tr>
<th>Table 1.</th>
<th>Crop production projected by some hybrid models and an econometric model</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 2.</td>
<td>Potential linkage variables between econometric and linear programming components</td>
<td>71</td>
</tr>
<tr>
<td>Table 3.</td>
<td>Crop production supply elasticities with respect to price</td>
<td>78</td>
</tr>
<tr>
<td>Table 4.</td>
<td>Ex-post simulation results</td>
<td>80</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Equilibrium solution ($q_0$) if econometric component falls in the feasible region (a) or outside the feasible region (b)</td>
<td>21</td>
</tr>
<tr>
<td>2</td>
<td>One-way communications model</td>
<td>24</td>
</tr>
<tr>
<td>3</td>
<td>Simultaneous solution model</td>
<td>24</td>
</tr>
<tr>
<td>4</td>
<td>Recursive interactive programming model</td>
<td>24</td>
</tr>
<tr>
<td>5</td>
<td>Recursive adaptive programming model</td>
<td>24</td>
</tr>
<tr>
<td>6</td>
<td>Points C, B, and D are candidates for adjusted values for infeasible solution $A (q_1, q_2)$</td>
<td>33</td>
</tr>
<tr>
<td>7</td>
<td>The 105 production areas</td>
<td>65</td>
</tr>
<tr>
<td>8</td>
<td>The 28 regions with central cities indicated</td>
<td>67</td>
</tr>
<tr>
<td>9</td>
<td>A schematic representing the CARD-ESCS model</td>
<td>69</td>
</tr>
<tr>
<td>10</td>
<td>Feedback process in the recursive adaptive programming model</td>
<td>73</td>
</tr>
<tr>
<td>11</td>
<td>Estimating production flexibility coefficients $\beta_{ij}t$ and $\tilde{\beta}_{ij}t$</td>
<td>75</td>
</tr>
<tr>
<td>12</td>
<td>Corn: comparison of performance of the hybrid and CED-CC models to actual observations</td>
<td>82</td>
</tr>
<tr>
<td>13</td>
<td>Soybeans: comparison of performance of the hybrid and CED-CC models to actual observations</td>
<td>83</td>
</tr>
<tr>
<td>14</td>
<td>Oats: comparison of performance of the hybrid and CED-CC models to actual observations</td>
<td>84</td>
</tr>
<tr>
<td>15</td>
<td>Wheat: comparison of performance of the hybrid and CED-CC models to actual observations</td>
<td>85</td>
</tr>
<tr>
<td>16</td>
<td>Some results of dynamic simulation run 1</td>
<td>88</td>
</tr>
<tr>
<td>17</td>
<td>Some results of dynamic simulation run 2</td>
<td>89</td>
</tr>
</tbody>
</table>
INTRODUCTION

The U.S. Government plays an important role in national and regional agricultural production. The government uses production control programs, price supports, and loans to increase and stabilize national farm commodity prices and invests in research, development, and extension to promote national production. It has invested in various development projects, such as irrigation, to promote regional agricultural production or regulates land use to avoid permanent depletion of resources.

Government programs usually have considerable differential effects over time on regional agricultural production patterns. Completion of individual regional projects frequently has a considerable impact on production in other regions. For example, in implementing the land idling programs during the 1950s, the South was allowed to substitute feed grains for cotton as its contribution to supply control. In contrast, the Corn Belt and Great Plains had to withdraw land from grains without the alternative of replacement by a substitute crop. And, for example, differential regional effects are expected to be extremely important if the nation implements Sections 208 of Federal Water Pollution Control Act (PL92-500) which seeks to achieve swimmable, fishable waters. Regions with much sloping land and large rainfall will have to be farmed less intensively and employ more capital-intensive conservation practices. In contrast, regions of the Great Plains with more level land and less rainfall can intensify production and gain relative economic advantage over time.
Quantitative evaluations of the differential impacts of alternative national and regional programs on production, farm income, and food prices prior to policy enactment and implementation can be useful to policymakers. With advances in econometric modeling and operations research and with increasing capacity and efficiency of computers, numerous econometric and mathematical programming models have been developed and applied to analyze complex agricultural problems. Samples of econometric models are CED-CC, CHASE, CARD-RS, and DRI while programming models can be found in Heady and Srivastava (1975) and Schaller (1968).

Because there are policy issues that can be best analyzed by using a combined feature of the econometric and the programming model, interest has been expressed in combining various types of models for policy analysis to complement the uses of each individual model (Boss et al., 1977). Some attempt has already been made to develop hybrid models combining a positive econometric model with a normative model. Three hybrid models incorporating some different quantitative approaches are:

1. Quadratic Programming (Q.P.) Model (Takayama and Judge, 1964; Meister, Chen, and Heady, 1978). This model includes a set of econometric estimated demand and/or supply functions in a resource allocation programming model.

CED-CC: Cross Commodity Forecasting System, developed at ESCS, USDA. CHASE: CHASE Agricultural Model developed by Chase Econometric Associates, Inc. CARD-RS: Recursive Simulation Model developed at the Center for Agricultural and Rural Development, Iowa State University. DRI: Agricultural Model developed by Data Resource, Inc. Descriptions of these models and others are in Chapter 5.

The differences between positive and normative models have been discussed in the literature; for examples, see Friedman (1953), Heady (1961), Kelso (1965), Quance, and Tweeten (1971), and Shumway and Chang (1977). In short, the positive model concerns what was, is, or will be the consequences due to a change in government programs while the normative model concerns what ought to or could be the consequence.
2. Input-Output and Linear Programming (IO-LP) Model (Penn et al., 1976). This model adds an IO model to a linear programming resource allocation model.

3. Recursive Programming (RP) Model (Day, 1961). This model has a set of econometrically estimated flexibility restraints added to a linear programming model.

This study investigates the need and the methods of combining an econometric model with a programming model, and describes a specific hybrid model developed jointly by the Center for Agricultural and Rural Development (CARD) at Iowa State University and the Economics, Statistics, and Cooperatives Service (ESCS) of the U.S. Department of Agriculture.

The study is divided into nine chapters. The first chapter describes the structural differences between econometric and programming models. The second chapter identifies the need for developing a hybrid model which combines an econometric model with a programming model. Subsequently, Chapter 3 investigates methods of linking an econometric model with a programming model and lays out the framework for developing a recursive adaptive programming hybrid model. Mathematical formulation of the basic structure of national regional, and simultaneous national-regional hybrid models is illustrated in Chapter 4. Because it is less expensive to develop a hybrid model by using an existing econometric and a programming model as the components of the hybrid model, Chapter 5 surveys the current econometric model from the available source of information. Chapter 6 surveys the benchmark of LP models and their applications. These survey results are used as the basis for selecting a specific econometric programming model for building the CARD-NRED recursive adaptive programming model. Description of this model is in Chapter 7. Chapter
Chapter 8 discusses some of the testing results of the model. Finally, Chapter 9 summarizes the report and suggests future work.
CHAPTER I. THE IMPLICATIONS OF STRUCTURAL DIFFERENCES BETWEEN ECONOMETRIC AND PROGRAMMING MODELS

Traditionally, economists use either an econometric or a mathematical programming model for policy analysis. This chapter examines structural differences between econometric and programming models. Implication of the functional differences between these two models provides the basis for justifying needs for developing a hybrid model.

The Basic Structure of the Econometric Model

There are various forms of econometric models (see Johnston, 1972) or Intriligator, 1978). A generalized econometric model used at time t (referred to as EMt, hereafter) consists of N independent regression equations:

\[ Y_{1t} = \phi_1(Y, Z, A_1) + e_{1t} \]
\[ \vdots \]
\[ Y_{Nt} = \phi_N(Y, Z, A_N) + e_{Nt} \]

where \( Y \) denotes all possible endogenous variables in equation i at current and lagged time except the one \( Y_{it} \) on the left-hand-side of the equation.

\( Z \) denotes all exogenous variables at current and lagged time.

\( A_n, n=1,...,N \) denotes all regression coefficients in \( \phi_n \).

\( e_{nt}, n=1,...,N \) denotes an error term for regression equation n.
Procedures to find an explicit function in the equations in (1) include the following two steps:

1. Applying economic theory to specify the functional relationships of endogenous and exogenous variables, and
2. Fitting the function to historical data by using econometric techniques of identification and estimation.

The explicit function can be used to investigate relationships of economic variables to forecast and evaluate alternatives. When using EMT for policy analysis, a policymaker can estimate the values of the N endogenous variables, \( y_{nt} \), \( n=1, \ldots, N \), for time period \( t \) from a given set of values of exogenous variables, \( z \), and a set of initial values of the lagged endogenous variables. Some of the exogenous variables, are often called policy variables. The policymaker can assign various values to \( z \) according to the policy and estimate the corresponding effects on \( y_{nt} \).

A simple linear (\( \phi_1 \) is of a linear function) regression equation with one endogenous variable (\( N=1 \)) and one exogenous variable is used to illustrate the essence of using an econometric model for a policy analysis. This simple regression equation is expressed as:

\[
y_{1t} = a_0 + a_1 y_{1,t-1} + a_2 z_{1t} + e_{1t}
\]

where \( y_{1t} \) is the market price of a crop commodity.

\( z_{1t} \) is a policy variable, representing commodity inventory level controlled by a government.

\( e_{1t} \) is an independent error term having a zero mean and a constant variance.
a_0 is a constant, a_1 and a_2 are regression coefficients. a_0, a_1 and a_2 are estimated by fitting the equation to historical data. Let these estimated values be \( \hat{a}_0, \hat{a}_1, \) and \( \hat{a}_2. \)

The estimated value of \( Y_{1t} \) can be expressed as:

\[
\hat{Y}_{1t} = \hat{a}_0 + \hat{a}_1 Y_{1,t-1} + \hat{a}_2 Z_{1t}
\]

Given the estimated value for \( Y_{1,t-1} \), this equation will allow the policymaker to estimate effects of various levels of \( Z_{1t} \) (government inventory) on the commodity \( Y_{1t} \) in each time period \( t. \)

Basic Structure of the Programming Model

There also are many types of programming models.\(^1\) However, all the programming models possess a common structure, that is, an objective function and a set of equality/inequality constraints. To simplify explanations, only a static and deterministic programming model is used for discussion. A generalized programming model (referred to as PM\(_t\), hereafter) can be expressed as:

Objective function:

Optimize \( \pi_t = f(X_{1t}, X_{2t}, \ldots, X_{nt}) \)

Subject to the following constraints

\[
g_i(X_{1t}, X_{2t}, \ldots, X_{nt}) \leq G_{it}
\]

for \( i=1, \ldots, I \)

\[
h_j(X_{1t}, X_{2t}, \ldots, X_{nt}) = H_{jt}
\]

for \( j=1, \ldots, J \)

---

\(^1\)For example, see Hiller and Lieberman (1971), Agrawal and Heady (1972), and Chow (1975).
\[ q_k(X_{1t}, X_{2t}, \ldots, X_{nt}) \geq Q_{kt} \]

\[ k=1, \ldots, k \]

where \( X_{1t}, X_{2t}, \ldots, X_{nt} \) are called decision variables which are to be determined for time period \( t \). \( G_{it}, H_{jt}, \) and \( Q_{kt} \) are constants, while \( f, g_i, h_j, \) and \( q_k \) are functions. Their function forms are assumed to be unchanged with respect to time \( t \). These functions may be either linear or nonlinear. The function of a programming model is to find a set of values \( X_{1t}, X_{2t}, \ldots, X_{nt} \) that gives maximum (or minimum) values for \( \pi_t \) and satisfies the set of inequality and equality constraints (2).

The programming models are often used for finding an optimal policy or for evaluating alternative optimal policies. A policymaker can vary the values of (1) \( G_{it}, H_{jt}, \) or \( Q_{kt} \), and (2) technological coefficients expressed in functions \( f, g_i, h_j, \) and \( q_k \) to investigate the impact of a policy related to commodity demand, resource supply, and production technology change. A policymaker also can delete or include a constraint to investigate the impact of problems relating to a change in the structure of production or regulation.

A simple linear (\( f, g_i, h_i, \) and \( g_k \) are linear functions) programming model is used as an example to illustrate how the model can be used for policy analysis purposes.

minimize \[ Z = C_1X_{1t} + C_2X_{2t} \]

subject to

\[ a_{11} X_{1t} + a_{12} X_{2t} \leq G_{1t} \]
\[ a_{21} X_{1t} + a_{22} X_{2t} = H_{1t} \]
\[ a_{31} X_{1t} + a_{32} \geq Q_{1t} \]
where $X_{1t}$ and $X_{2t}$ are acres of soybeans and corn, respectively. $X_{1t}$ and $X_{2t}$ are to be determined. $C_1$ and $C_2$ are net income per acre to be received from producing soybeans and corn, respectively. Coefficients $a_{11}$ and $a_{12}$ represent technological coefficients per acre of water consumption for growing soybeans and corn, respectively. $G_{lt}$ is total water available. $H_{lt}$ is total land available for production. Coefficients $a_{31}$ and $a_{32}$ denote per acre gross income received from growing soybeans and corn, respectively. $Q_{lt}$ is the minimum gross income that will be achieved in the production.

A policymaker can use the model for numerous policy analyses: (1) The policymaker can reduce the value of available water $Q_{lt}$ to find the effect of the water supply on reduction of crop production and farm income. (2) In anticipating an increase in the price or yield of soybeans, a policymaker can adjust the value of $a_{31}$ to estimate the effect of a price change of net farm income and on the crop mix. (3) When farm labor becomes a critical factor in production, a policymaker can add another constraint to reflect limitations of labor supply available for production.

Some Implications

Economic basis

Both $EM_t$ and $PM_t$ have a sound economic basis for use as tools to estimate consequences due to a change in an economic setting. In the $EM_t$, a regression equation is specified according to interpretation of economic theory. For example, the market price of a crop commodity frequently is specified as a function of quantity demanded and quantity
supplied. Market equilibrium frequently is assumed. Under this assumption, maximization of consumers' and producers' surplus are implied. In the PM, a decision variable (for example, the production level of a crop) can be interpreted as a function of marginal costs, marginal value products of resources used, and the availability of the resources. The formulation, for example, allows a particular crop to be produced at its maximum production level if its marginal revenue product is greatest relative to the most critical resource in the model.

Statistical inference

The EM uses explicit error terms, , for all n and t in the model. This feature appears to give the EM the advantage of using established statistics theory in making inference from the obtained estimates. This property is lacking in the PM. However, the current available EM is complex and, in reality, making any statistical inference about its estimates is very difficult. This difficulty arises because current econometric theory regarding estimation of a system of simultaneous equations applies only to the large sample properties. Thus, its use for econometric models is very limited because long historical data series in general are not available for building the EM model. Furthermore, the currently available EM often includes conditional variables and nonlinear equations and therefore it becomes difficult to apply statistical inference techniques. Nonparametric measures (Shapiro, 1973) can be used in testing the tracking

\footnote{Sensitivity analysis can be applied for investigating variations of the solution with respect to changes of certain coefficients. However, the analysis results are good only for a specific case and cannot be generalized.}
ability of a large econometric model, but the measures frequently are not without controversy.

**Modeling flexibility**

In the $E_{M_t}$, $\phi_n$ can be a nonlinear function. Existence of the nonlinear function in the model, in general, will not add an additional significant computation effort in using the model for policy evaluation. This feature allows the model more flexibility in formulating a system of equations to simulate more closely to the observed data. On the contrary, when the $P_{M_t}$ is relatively large, functions $f$, $g_i$, $h_i$, and $q_k$, in general have to be in linear form to guarantee a solution if it exists. In a situation where a nonlinear equation has to be used to reflect the actual real world situation, linearization of the nonlinear function has to be done before searching for the solution. Linearization may not be possible for certain types of nonlinear functions. Even if it is possible, considerable effort must be spent to obtain the solution, if it exists.

**Forecasting and structure analysis**

In the $E_{M_t}$ model, coefficients in each function $\phi_n$, $n=1,...,N$ are estimated from observed historical data. The function often is the best regression equation in terms of least square error for explaining a past unknown economic system that generates the observed data. Consequently, $E_{M_t}$ is the proper tool for structure analysis to describe what has actually occurred in the past. It also is a more accurate forecasting tool, if the economic system remains unchanged, to estimate possible consequences of a future policy event which occurred before. However, in light of
changes in technology and institutions or government policy previously unencountered, use of EM_t as a tool becomes difficult because historical data are not available. In contrast, each function in a PM_t model can be estimated from either historical survey data or the data of other sources. For instance, the per acre water consumption, \( a_{11} \), of soybeans in the previous example can be obtained from the historical data, if available, but also can be estimated through observation from a planned small scale field experiment or derived from basic principles of plant physiology. In other words, the PM_t model can be constructed when historical data are absent. This property allows the PM_t model to be a proper tool for analyzing the consequences of an event that has not previously occurred. To build a PM_t model that will explain past national agricultural production is a formidable task. A trial-and-error method and a high level of expert knowledge about the whole agricultural production system are required to find a proper programming model to track time series data. Therefore, PM_t is not an effective tool for structural analysis.

Positive and normative

Traditionally, economists use EM_t for positive economic analysis while they use the PM_t for normative analyses. The positive analysis concerns what is, what was, and what will be the consequence of any change in circumstances (expressed as exogenous variables). The normative concerns what ought to or would be the consequences of any change in circumstances. In building the EM_t, a researcher fits a set of regression equations to a set of historical data to find a "true"
system that can best track actual observed data and also can predict
the future by assuming that the economic system remains unchanged. Information generated under this process is considered to be of a positive
nature. Economists classify the PMt as a normative model, mainly for
the following two reasons: (a) it contains an objective function which
reflects the judgment of the person who uses the model, and is often
criticized for not reflecting the production response of farmers; and
(b) synthesized data and the type of constraints used in the PMt also
are frequently criticized for reflecting the person's judgment rather than
the actual production situation.

But, classification of the PMt as a normative model, while consider­
ing the EMt as a positive model, has increasingly become improper; for
instance, a recursive programming model was an attempt to estimate the
information that could be classified as "positive" type of knowledge
(Day, 1961). Although it may not be economically feasible, it is tech­
cically feasible to construct a PMt model in which the production response
of individual farms is implemented to reflect actual farm operations.
Furthermore, advances in behavioral science allow a person to have more
capability in using the PMt model in a positive economic analysis (Boussard
and Petit, 1967). Similar arguments can be found for using EMt for norm­
mative economic analysis. For instance, Lee and Seaver (1973) used an EMt
to solve a spatial equilibrium problem that usually is solved by a quad­
artic programming model with maximization of net profit as its objective
function.

The objective function in the PMt can serve as a tool for a policy-
maker to reflect his judgment. This makes the PMt a natural tool for
normative economic analysis. Often the objective function is used to reflect an approximation of the real world situation. Whether this assumption is realistic perhaps depends upon the skill of the model builder and cost effectiveness of introducing a more complex function.

Whether a PM$_t$ should be considered and used for a normative or positive economic analysis should not be judged from its structure but rather from the intention of the person who built the model, that is, how the data are used and the way objective functions and constraints are set up. At the present state of art in model building, EM$_t$ is generally a better tool for the positivistic analysis to achieve a given level of accuracy, while the PM$_t$ is a natural tool for normative economic analysis.

**Intraregional and interregional analysis**

Use of an objective function in the PM$_t$ as a performance measurement for crop production and resource use in each region allows competition among regions for crop production and crop competition in each region for limited resources. Therefore, regional production adjustments and regional resource use can be examined in detail and for economic environments not previously experienced. Regional competition can also be examined by using the EM$_t$ only if historical data are available. One approach in the latter is to use regional production trends and production shares of the region. This approach explains a fact historically but lacks a mechanism to investigate the regional adjustment under a different economic setting. Another approach is the simultaneous equation model (Lee and Seaver, 1973). The approach needs to solve $2^n(n-1)$ solutions
for \( n \) regions. Extensive computation is required when the number of regions is relatively large. Thus, for investigating resource use for competition between regions and within regions, the \( PM_t \) is a better tool.
CHAPTER II. NEEDS FOR AND CASES OF HYBRID MODELS

Although there are policy issues that can be best analyzed either through an econometric or a programming model, there also are policy issues that can be analyzed through a linked or hybrid model which combines an econometric model (or component) with a programming model (or component). Linking an econometric component with a programming component allows one component to provide policy variables, information, and analytical structure that could not be specified in the other component. Requirements for a hybrid model can be identified according to their emphasis of model uses. The hybrid model (a) can provide detailed and complementary information through the outputs from both econometric and programming models, (b) has a new analytical structure that neither an econometric nor a programming model can provide, and (c) improves the predictive accuracy of an econometric model or transforms a programming model into a predictive model.

The hybrid model belonging to the first emphasis is characterized by linking an econometric (or programming) model to a programming (or econometric) model. Output from one model becomes the exogenous values to the other model. This group of models is frequently used for three types of studies.

The first type of study estimates regional production patterns and production related activities (output of a programming model) from a given set of commodity demands (projected by an econometric model). The econometric component provides information on market activities at the
national level while the programming component generates information on production activities and resource use at the regional level. A typical example is the use of consumption demands projected by the NIRAP econometric model (Quance, 1976) as constraints in the CARD interregional linear programming model (Meister and Nicol, 1975). The purpose of using the model is not to predict a production pattern, but to estimate regional production capability for given national market demands of crop commodities.

A second type of study, using output of one model as input into another, estimates the production potential under different policies or resource capabilities and then estimates its market impact if the potential were attained. The linear-programming, input-output (LP+IO) model used by Sonka and Heady (1973) is a typical example. The nature of the model is to estimate the likely market responses if the production change is realized.

Finally, another type of study examines the optimal designs or plans in regional production practices to accomplish specific objectives under different national economic situations. For instance, Saygideger (1977) used the modified NWA model (Meister and Nicol, 1975) with the NIRAP econometrical project demands (Quance, 1976) to investigate production activities from several possible alternatives to curb problems related to environmental quality. A large number of potential studies falls in this group. The main purpose of this application is to determine what production ought to occur under different possible market demands generated from different scenarios for the national economy.

The second emphasis on using the hybrid model is to have a new analytical structure resulting from the combination of an econometric
model with a programming model. The model is characterized by the outputs of both the econometric and programming components being determined either simultaneously or recursively. This group of models allows an interfacing of market activities at the national level and production activities on the regional level. The simultaneous solution model utilizes equations derived from an econometric model as identities (rather than inequality constraints) within the programming model. A typical simultaneous model is the IO-LP model used by Penn, et al. (1976) for evaluating the impacts of energy shortages on the U.S. economy. The Quadratic Programming (QP) model with econometrically estimated market behavior restraints also belongs in this latter group. Typical studies are the QP model by Plessner (1965) and Meister, Chen, and Heady (1978). The model allows interaction between market mechanisms with regional production activities and resource uses.

Another way to capture interaction between national market activities and regional production is through a recursive modeling structure. At time period t, an econometric (programming) model provides input data to run the programming (econometric) model at the next time period, t+1. The recursive interactive programming models (Baum, 1978; Schaller, 1968; Sharples and Schaller, 1968) are examples.

The third emphasis of using a hybrid model comes from attempts to improve prediction performance, characterized by including a set of econometrically estimated equations in a programming model to bound regional production responses, or by using a LP model to generate data as input to an econometric model to improve accuracy of prediction.
The recursive programming (RP) model (Day, 1961) and the LP-time series approach (Shumway and Chang, 1977) are two examples. The RP model uses a set of econometrically estimated flexibility restraints to limit the production shift from one time period to another to tract more closely the observed data. This feature allows the use of a programming model to produce positive estimates. The LP-time series approach uses LP generated output as prior information to an econometric model for estimating crop supply elasticities.

Consequently, need for detailed information, a better analytical structure, or an improvement of prediction accuracy leads to the development of various hybrid models. By linking an econometric model with a programming model, analytical capability is extended from a pure normative or positive economic analysis to a wide range of analyses through a combination of unique features available in each of the models. Many positive analyses with prior normative assumptions can be analyzed through the hybrid model. Similarly, normative analysis with prior positive assumptions can be performed.

Several features of the hybrid or linked model have importance for the analysis of agricultural, environmental, soil conservation, and land and water allocation policy. The model can be used for policy analysis at both regional and national levels. It can estimate regional differential impacts due to a given national policy and estimate impacts at the national level due to a given regional policy. It has a dynamic structure and can estimate temporal effects of interrelated events. It provides the time path of various impacts. It has a structure that is
related spatial patterns of supply, resource use, and the technical structure of commodity production to the national market processes and price levels. It has other versatile applications. It can track historical events and give a positive prediction. It also can be used for analysis of policy issues involving a normative analysis with prior positive assumptions or positive analysis with prior normative assumptions.

Among possible types of hybrid models, the simultaneous solution model with a time recursive structure, conceptually, appears to be the best approach. It gives solutions that simultaneously satisfy the assumption of both econometric and programming components in each time period. Figure 1 illustrates solutions of the simultaneous model when the equilibrium solution of the econometric component is in or outside the feasible region (dashed lines) defined by the programming component. The model uses a set of production restraints and uses maximization of consumers' and producers' surplus as its objective function. The model will have the equilibrium solution \( q_0 \) as its solution when the equilibrium solution is inside the feasible region (Figure 1a), and will have a disequilibrium solution \( q_1 \) as its solution when the equilibrium solution \( q_0 \) is outside the feasible region \(^1\) (Figure 1b). Since the solutions in both situations will satisfy the assumption of the two components, a consistent result is obtained. In practice, however, the simultaneous approach is difficult to use. In many cases both the supply and demand equations may be nonlinear with lagged variables. Under

\(^1\)It is assumed that the feasible region defined by the programming component can be formulated more accurately through engineering data, interpretation of government regulation and other information.
Figure 1. Equilibrium solution \( (q_0) \) if econometric component falls in the feasible region (a) or outside the feasible region (b)
these conditions, computation problems may arise. For this reason, an alternative approach is needed to give consistent estimates regardless of the location of the equilibrium solution. The approach we use is a recursive adaptive programming (RAP) model.
CHAPTER III. SIX METHODS OF COMBINING AN ECONOMETRIC AND A PROGRAMMING MODEL

The One-Way Communication Model

This section contains a discussion of six alternative approaches for linking econometric (EM) and mathematical programming (PM) models, and subsequently describes the conceptual framework for building a RAP hybrid model.

The One-Way Communication Model (Figure 2) is so named because the information flow is one way--from the econometric model to the programming model (or vice versa). This hybrid is most easily characterized by a single-period and an interregional programming model with fixed demands which are determined by a set of econometric equations. For example, a one-way communication model was used to analyze alternative future potentials for U.S. agriculture as defined for the National Water Assessment conducted under the auspices of the U.S. Water Resources Council (Meister and Nicol, 1975). At each point in time (i.e., the years 1985 and 2000) in this analysis, the quantities of agricultural products demanded (U.S. Water Resources Council, 1974 and 1975) were projected by the NIRAP econometric model (Quance, 1976). These demands were used as constraints in a linear interregional programming model (Meister and Nicol, 1975). The linear programming model was then used to project the least-cost (competitive equilibrium) spatial pattern of agricultural production and resource use subject to these minimum fixed demands.
The solution of the hybrid model for time period $t; t=1,2,...,T$

The solution of a set of equations component of the hybrid model

$P$ A set of equations obtained from a programming model
$E$ A set of equations obtained from an econometric model

The transfer of information between components and/or models
An example of linking programming to other models is provided by Sonka and Heady (1973). They used an LP-IO model to analyze impacts of farm policy on rural income and employment. The model's solutions were summarized into 10 farm production regions. The summary was linked to an input-output (I-O) analysis to evaluate secondary impacts on rural income and employment. The basic coefficients in the I-O analysis were derived from Schluter (1971).

One-way communication models are currently available for this type of long-range analysis. But, the ability of these models to simulate the short-run behavior (which is not their original purpose) of the agricultural sector is limited by the fact that the models do not have a feedback from the programming to the econometric model within or between time periods. The model \( (EM_t \rightarrow PM_t) \) will encounter a problem of non-feasible solutions when the econometrically estimated values of the linkage variables are outside of the feasible region defined by the restraints in the programming model. Proper adjustment of the econometric or programming model is required to obtain a feasible solution. The model \( (PM_t \rightarrow EM_t) \), however, is likely to overestimate production and thus underestimate prices if a set of production flexibility constraints is not added to the programming component.

The Simultaneous Solution Model

The simultaneous solution model (Figure 3) uses equations derived from an econometric model as identities (rather than inequality constraints) within the programming model. The conceptual appeal of this hybrid is that the solution to the model will simultaneously satisfy the assumptions of both parent models.
Penn, et al. (1976) used this approach to evaluate the short-run impacts of energy shortages on the U.S. economy. This Simultaneous Solution Model incorporated input-output data developed by the U.S. Department of Commerce (1974) for 85 sectors into a linear programming model that contained two energy constraint equations.

Quadratic programming models also belong to the simultaneous solution category. The QP incorporates demand and supply equations into a linear programming model (Takayama and Judge, 1964) and has been applied extensively to study the problem of U.S. agricultural production (e.g., Plessner, 1965; Meister, Chen, and Heady, 1978).

Problems will arise in future applications of a Simultaneous Solution Model under any one of the following three conditions:

1. where the feasibility region defined by the equations derived from the positive model is smaller than the computational errors inherent in the linear programming software package;
2. where a static equilibrium solution is imposed on a dynamic disequilibrium system (see Baumol, 1951); and
3. where nonlinear equations derived from the econometric component result in prohibitive computational costs when cast within a mathematical programming framework.

For example, a Simultaneous Solution Model constructed using equations from the CED-CC Forecasting System and the CARD-NWA programming model would contain thousands of equations and tens of thousands of variables. A Simultaneous Solution Model of this size would be computationally infeasible and/or prohibitively costly (particularly if bounding procedures are used).
Recursive Interactive Programming (RIP) Models

Characteristic of RIP models

The basic features which characterize Recursive Interactive Programming (RIP) models (Figure 4) are: (a) The hybrid model develops a unique description of the agricultural economy for each stage in a sequence of time periods (t=1, 2,...,T). (b) The hybrid model consists of at least one programming and at least one econometric component. (c) Within each stage the individual components are solved once in a prespecified sequence. The former component (the programming model or the econometric model in Figure 4) is solved before the latter-component (the econometric model or the programming model in Figure 4) is run. (d) For the purpose of this discussion, each component has three categories of variables. These categories, as follows, are not necessarily mutually exclusive: Exogenous variables are not determined within either component. The "explanation" for their behavior resides outside the component. When the equations within a component are solved, these variables are taken as given. Endogenous variables are those whose values are "explained" or determined by the operation of the component model. Linkage variables are the exogenous variables in one component whose values are determined by the operations of the preceding component. (e) The solution procedure for the RIP model begins by running the former component for the first stage. The input data consist of a set of initial values for the lagged endogenous variables; the solution vector produced by this run contains the values of the former component's endogenous variables; this is used to determine the values of the linkage variables in the latter component.
Subsequently, the latter component is solved. The solution of the latter component in this stage is used to determine the starting values of the linkage variables in the former component of the next stage. (f) Using the recursive relationship identified in "e", above, the RIP model moves forward stage by stage.

Two examples of the RIP model are described: First: The national model (Shaller, 1968 and Sharples and Shaller, 1968), developed by the Farm Production Economics Division, Economic Research Service, is one example of the RIP model. It consists of about 90 profit-maximization linear programming submodels. At year t, these submodels estimate planned acreage which is used to estimate the planned production as well as the quantity of production input used. These estimates are then fed into a national econometric model to calculate equilibrium prices and then expected prices. These prices, estimated costs, production, and input uses are fed into each programming submodel to estimate the production for year t+1. Second: Baum (1978) built and empirically tested a national recursive programming model for U.S. agriculture. He used the crop sector of the abridged version of the CARD-NWA linear programming model with a revised econometric simulation model based on one developed by Ray and Heady (1974). Within each stage of his analysis, the profit-maximizing linear programming model was run first to estimate national crop acreage and production. (In the first example, 90 profit-maximization submodels were used.) The values of these linkage variables were then passed to the simulation model. The simulation model was subsequently run to estimate values of market sector variables for the same stage in time. The output of the simulation component was used to revise
the coefficients in the linear programming model in the following stage. These coefficients include the net return coefficients in the objective function, the values of the activity flexibility restraints, i.e., the upper and lower limits on crop acreage response by region, and the values of input-output coefficients, optimal nitrogen fertilizer rates, and crop yields.

The RIP models have many advantages over those described earlier. They allow for a two-way flow of communication—one way within each stage and the other between stages. This is a higher degree of interaction between components than is achieved by the one-way communication models. They present less of a computational problem than the simultaneous solution models because the feasibility set is not restricted to equality solutions of the econometric model. Finally, they dynamically simulate a sequence of events over space and through time in a nonsimultaneous, or cobweb solution framework.

The RIP approach also has limitations. An RIP hybrid begun by running with the LP model tends to overestimate total production and underestimate prices when the interest is in positive predictions. This is because the linear programming component produces an economically efficient use of resources. This overestimates production as input to the econometric model and therefore, underestimates the prices.

The RIP hybrid, which begins by running with the Econometric model, may encounter the infeasible solution problem as described in the one-way communication hybrid model. The econometric component may give an estimated production that exceeds the capacity of regional production. If either of the components has been specified incorrectly, the recursive
nature of the model may result in a propagation of errors over time, between stages.

The first problem can be ameliorated by introducing pseudo behavioral constraints into the programming component; Shaller's and Baum's procedure of adjusting upper and lower bounds on regional acreage limitations in response to the price impacts produced by the econometric component is an appropriate methodology. Additional research is needed to improve the accuracy of regionally specific acreage (or production) response equations. The second and third problems can be addressed in part by incorporating a two-way flow of communications between components within each stage of the analysis. This concept of a corrective adjustment within the stage feedback mechanism is similar to a self-adaptive control system (D'Azzo and Houpis, 1966); it is defined as a model which has the capability of changing values of linkage variables through an internal process of estimation evaluation and adjustment according to a pre-setup rule. It forms the basis for the Recursive Adaptive Programming (RAP) models which are described in the following discussion.

Recursive Adaptive Programming (RAP) Model

The RAP model (Figure 5) is constructed from the RIP model by including a feedback structure in each stage. A RAP model is constructed by using an econometric model as the former component and a linear programming model as the latter component. The econometric model is used as the former component based on the following reasoning. For evaluating the short-run impacts of agricultural policies, it is natural to relate the econometric model as the principal component in the hybrid model and
to use the linear programming model to act in the following subordinate
and complementary role: (a) For each commodity, the LP contains an
accounting row that measures the deviation between aggregate production
as forecast by the econometric component and the aggregate contained
in the LP solution. Large penalty costs have been assigned to the
deviational variables in the profit maximizing objective function to
force the LP solution to come as close as possible to the econometric
solution.

If all the deviational (production) variables in the LP solution
vector are equal to zero (that is, the econometrically estimated value
which falls inside the feasible region in Figure 1.(a)), the solution
produced by the two components is assumed to be consistent. In this
case, the LP has validated the results of the econometric component,
and the RAP model begins the computations for the next stage in time.
(b) But, if any of the deviational variables in the LP solution vector
are not equal to zero (Figure 1.(b)), then the production possibilities
region defined by the feasibility constraints in the LP component is
actively determining the solution. In this case, the pre-set-up adaptive
feedback mechanism is invoked. Within this stage, the production vari-
ables become linkage variables from the programming component to the
econometric component; they are set equal to the LP solution values.
The econometric component is solved, producing a new set of prices. Then
the RAP model goes forward to the next stage of analysis.
Alternative Feedback Adjustment Procedures for RAP Models

The key problem in building the RAP hybrid model is to find the best pre-set-up procedure to adjust the production (or acreage) when the equilibrium solution is outside the feasible region.

Harrison (1976) suggested an iteration procedure to find the equilibrium prices in a hybrid model. Convergence of shadow prices from linear programming is used as the criterion. This procedure is appropriate when both the econometric and the programming components in the hybrid model have a simple model structure and are inexpensive to run. This is not the case for integrating large-scale models. Four potential procedures are examined. Figure 6 illustrates four procedures that can be used to adjust the econometrically estimated equilibrium solution $A(q_1, q_2)$ which is outside the feasible region in a two-crop model.

**Shortest Distance Approach (SD)**

This approach will give the final solution indicated by B in Figure 6. AB is the shortest distance as measured by the sum of squares of the quantity of the two crops to be adjusted from A to the feasible region.

**Independent Adjustment Approach (IA)**

This approach gives the solution indicated by C. This approach adjusts only the crop production which is larger than the feasible range (OE). The approach, however, does not adjust the crop production which is less than the feasible range (OH).

**Maximization Approach (MA)**

This approach ignores the equilibrium solution generated by the econometric model. The approach obtains the adjusted value simply by running the programming component. One of the likely solutions is indicated by D in Figure 6.
Figure 6. Points C, B and D are candidates for adjusted values for infeasible solution A \((q_1, q_2)\)
**Minimum Absolute Distance Approach (MAD)**

This approach gives the solution which has the least adjustment of absolute values of the production. The solution may be either C or B, depending on the slope of DE as shown in Figure 6.

Each approach has its own appeal: It should be noted that it is possible that the SD, MA, and MAD approaches may obtain the same adjusted values when the econometric estimated value is at A' in Figure 6. The SD has the least squares of quantity of production to be adjusted; the IA has only to adjust the crop which is outside the feasible region; the MA approach gives the adjusted value which maximizes net return from the production; the MAD has the least absolute value in adjustment. Furthermore, the MAD gives a solution either the same as the solution of the SD or the solution of the IA. Because of this characteristic, the MAD approach is used in building the RAP hybrid model.
CHAPTER IV. MATHEMATICAL FORMULATION AND AN EXPERIMENT TO ILLUSTRATE THE ESSENCE OF HYBRID MODELS

A mathematical summary is employed to illustrate the linkage between an econometric and a programming component in the RAP hybrid model. Use of the hybrid model as a regional, national, or regional-national model also is illustrated. A simple experiment is conducted to show solutions resulting from various methods of linkage.

Mathematical Exposition

An econometric component and a programming component are used to illustrate the essence of constructing the hybrid models. Assuming an econometric component, which consists of $N$ equations in the hybrid model, is expressed as

$$\begin{align*}
Y_{nt} &= \sum_{i=1}^{I} a_i Y_{it} + \sum_{j=1}^{J} Y_{jt-1} + \sum_{k=1}^{K} c_k Z_{kt} + e_{nt} \\
&\text{for } n=1,2,\ldots,N; \text{ for } i=n
\end{align*}$$

(1)

where $Y_{nt}$ and $Z_{kt}$ denote endogenous and exogenous variables respectively; $a_i$, $b_j$, and $c_k$ are coefficients; $e_{nt}$ is an error term. The first $I$ ($I<N$) endogenous variables are linking variables to a programming component which is expressed as

$$\begin{align*}
\text{Maximize } & \sum_{ij} (P_{ij} - C_{ij})X_{ij} - M_1 (\sum V_i^+ + V_i^-) - M_2 (\sum W_i^+ + W_i^-) \\
\text{Subject to: } & (1) \text{ National Production Balance Restraints} \\
& \sum_{j=1}^{J} X_{ij} + V_i^+ - V_i^- = Y_{it}, \text{ for } i=1,2,\ldots,I
\end{align*}$$

(2)
(2) Regional Production Response Balance Restraints

\[ X_{ijt} + W_{ij}^+ - W_{ij}^- - \beta_{ijt} X_{ijt-1} \]  

for \( i=1,2,\ldots,I \)
\( j=1,2,\ldots,J \)

(3) Production Resource Restraints

\[ \sum_{i=1}^{I} \sum_{j=1}^{J} T_{ijt} X_{ijt} \leq R_{lt} \]  

for \( \ell=1,2,\ldots,L \)

Where:
- \( P_{ijt} \) = farm price for crop \( i \) in producing region \( j \) in time period \( t \);
- \( C_{ijt} \) = cost of production for crop \( i \) in producing region \( j \) in time period \( t \);
- \( X_{ijt} \) = quantity of production of crop \( i \) in region \( j \) in time region \( t \);
- \( M_1 \) and \( M_2 \) = two arbitrary large constant values satisfying the following conditions: \( M_1, M_2 > (P_{ijt} - C_{ijt}) \) for all \( i, j, \) and \( t \);
- \( V_i^+, V_i^- \) = positive or negative deviation from econometric estimated production of crop \( i \) \( (V_i^+, V_i^- \geq 0) \);
- \( W_{ij}^+, W_{ij}^- \) = positive or negative deviation from econometric estimated production of crop \( i \) in region \( j \) \( (W_{ij}^+, W_{ij}^- \geq 0) \);
- \( T_{ij} \) = technological coefficients for using resource \( \ell \) by crop \( i \) in region \( j \);
- \( R_{\ell t} \) = maximum amount of resource \( \ell \) available in time \( t \);
- \( \beta_{ijt} \) = coefficient used to predict the products \( X_{ijt} \) from
The value of $X_{ijt-1}$ is estimated from a regression equation which has independent variables such as expected price and other variables.

The objective function (2) is to maximize production net returns and minimize the absolute deviation between the programming solution and values of national and regional econometric estimates. The formulation of minimizing the absolute deviation is described in Sposito (1975). Another form of the objective function, e.g., minimization of production costs and the deviation or minimization of the deviation also can be formulated.

Properly assigning values for $V_i^+, V_i^-, W_{ij}^+, \text{ and } W_{ij}^-$ the model ((1) to (5)) can be transformed into a national, regional, and a simultaneous national-regional hybrid model. The model becomes a one-way (N) national hybrid model by setting $V_i^+$ and $V_i^-$ equal to zero. The model is a national model because the production in the solution from the programming component is set to the value estimated by the econometric component. The N model usually is used to analyze regional production response to meet a national target quantity of the production. Similarly, the model can become a one-way regional (R) hybrid model by setting $W_{ij}^+$ and $W_{ij}^-$ equal to zero. The model is regional because national production is determined by summing all the regional production. This R model is useful to investigate possible impact from regional production expansion. When the R model is structured in such a way that the sum of all the regional production is used in the next time period (stage), the model becomes a two-way communication or RIP model as mentioned earlier.
The model also can be transformed into a jointly determined national
and regional (NR) hybrid model when $v_1^+, v_1^-, w_{ij}^+, \text{ and } w_{ij}^-$ are not set
equal to zero or when ranges are used to replace these variables. One
variation of the NR simultaneous model is to employ a pair of flexibility
restraints to give a range for each crop production in each region. The
model is a simultaneous one because the final solution production in
the programming component is jointly determined by national and regional estimates. This is the basic structure of the RAP model explained earlier.

At each time period, the RAP model checks whether $\sum_{j=1}^{I} x_{ijt}$ is equal to $y_{it}$. If not, (either $v_1^+$ or $v_1^-$ is not equal to zero) the value of $\sum_{j=1}^{I} x_{ijt}$ replaces the value of $y_{it}$ and is fed back to the econometric model to adjust the values of all endogenous variables before the RAP model starts for next time period (stage t+1).

If the first term (the net profit) in the equation (2) is set at zero, this RAP model is equivalent to a statistical model which is fitted with least-absolute-deviation to a series of production data generated by the econometric component.

An Experiment to Investigate Characteristics of Hybrid Models

An experiment is conducted to show solution results from various
methods of linkage. Although the experiment uses a specific simple model,
the test results demonstrate several fundamental characteristics resulting
from each method of linkage. The experiment uses a simple two-crop
econometric model which is formulated as:
\[
P_1 = 250 - 0.003 q_1 \quad \text{(P_1 and q_1 are price and quantity of crop 1)}
\]

\[
C_1 = -1125 + 0.250 q_1 \quad \text{(C_1 is cost of production of crop 1)}
\]

\[
P_2 = 400 - 0.080 q_2 \quad \text{(P_2 and q_2 are price and quantity of crop 2)}
\]

\[
C_2 = -1143 + 1.4 q_2 \quad \text{(C_2 is cost of production of crop 2)}
\]

The programming model in the experiment uses only land resource constraints and is expressed as:

Maximize

\[
f(P_1, P_2, C_1, C_2, q_1, q_2)^{1/}
\]

1/ The objective functions used in the experiment are formulated as:

a. Surplus Simultaneous Model

\[
G \int [(250-0.003q_1)-(-1125 \text{ to } 25q_1)]dq_1 + \int [(400-0.08q_2)-(-1143) + 1.4q_2]dq_2
\]

b. One-Way, or Recursive Model

\[
(p_1-c_1)q_1 + (p_2-c_2)q_2
\]

where \(p_1, c_1, p_2,\) and \(c_2\) are predetermined by an econometric model (component)

c. Net Income Simultaneous Model

\[
[(250-0.003q_1)-(-1115 + 0.25q_1)]q_1 + [(400-0.080q_2^2)-(-1143) + 1.4q_2]q_2
\]

d. Recursive Adaptive Model

\[
(p_1-c_1)q_1 + (p_2-c_2)q_2 - M_1(v_1^+ + v_1^- + v_2^+ + v_2^-)
\]

where \(p_1, c_1, p_2,\) and \(c_2\) are predetermined by the econometric component. \(M\) is arbitrarily assigned to a value of 1000.
Subject to:

\[
\frac{q_1}{50} + \frac{q_2}{40} \leq L
\]

\[
P_1 > C_1, \quad 0 > C_2
\]

Table 1 shows quantities of crop 1 and crop 2 produced in each case. The Surplus Simultaneous model is used as the reference model because this model gives maximum consumer's and producer's surplus among all the cases considered. All other cases (except the EMt model) use maximization of net profit as their objective function.

Several conclusions can be drawn from these results:

(a) In the abundant land resource case, only the model beginning with the EMt model gives the correct solution (same as the solution of the reference model). This implies than any hybrid model beginning with the LPt model will give a wrong solution. To avoid this shortcoming, production flexibility constraints are frequently added to the LP model.

(b) In the tight land resource case, none of the model yield the correct solution. The one-way and recursive interacting programming (RIP) models, begun by running the EMt model give a nonfeasible solution, while models begun by running the LPt model give results with extreme values. The simultaneous model with a net profit objective function underestimates the production. The solution from the recursive adaptive programming (RAP) model, although incorrect, is relatively close to the correct solution -- especially so when the Marginal Revenue Produce (MRP) of land can be estimated accurately. The RAP solution is (4697,1042) as compared with the true solution of (4844,925) because the RAP model is formulated to minimize the absolute deviation between the econometrically
Table 1. Crop production projected by some hybrid models and an econometric model

<table>
<thead>
<tr>
<th>Models</th>
<th>Abundant Land Resource</th>
<th>Tight Land Resource</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. SURPLUS SIMULTANEOUS</td>
<td>(5434, 1042)</td>
<td>(4844, 925)</td>
<td>Assuming true solutions</td>
</tr>
<tr>
<td>$EM_t + LP_t$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. ECONOMETRIC</td>
<td>(5434, 1042)</td>
<td>(5434, 1042)</td>
<td>Correct solution for L=150, Invalid solution for L=120</td>
</tr>
<tr>
<td>$EM_t$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. ONE-WAY</td>
<td>(5434, 1042)</td>
<td>None</td>
<td>Correct solution for L=150, Nonfeasible solution for L=120</td>
</tr>
<tr>
<td>$EM_t \rightarrow LP_t$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. ONE-WAY</td>
<td>(7500, 0)</td>
<td>(6000, 0)</td>
<td>Incorrect solution for both L=150 and L=120</td>
</tr>
<tr>
<td>$LP_t \rightarrow EM_t$</td>
<td>(0, 6000)</td>
<td>(0, 4800)</td>
<td></td>
</tr>
<tr>
<td>$MRP_1 &gt; MRP_2$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$MRP_2 &gt; MRP_1$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. RECURSIVE INTERACTIVE (RIP)</td>
<td>(5434, 1042)</td>
<td>None</td>
<td>Correct solution for L=150, Nonfeasible solution for L=120</td>
</tr>
<tr>
<td>$EM_t \rightarrow LP_t$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. RECURSIVE INTERACTIVE (RIP)</td>
<td>(7500, 0)</td>
<td>(6000, 0)</td>
<td>Incorrect solution for both L=150 and L=120</td>
</tr>
<tr>
<td>$LP_t \rightarrow EM_t$</td>
<td>(0, 6000)</td>
<td>(0, 4800)</td>
<td></td>
</tr>
<tr>
<td>$MRP_1 &gt; MRP_2$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$MRP_2 &gt; MRP_1$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. NET INCOME SIMULTANEOUS</td>
<td>(2691, 521)</td>
<td>(2691, 521)</td>
<td>Incorrect solution for both L=150 and L=120</td>
</tr>
<tr>
<td>$EM_t + LP_t$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. RECURSIVE ADAPTIVE (RAP)</td>
<td>(5434, 1042)</td>
<td>(4697, 1042)</td>
<td>Correct solution for L=150, Incorrect solution for L=120</td>
</tr>
<tr>
<td>$EM_t \rightarrow LP_t$</td>
<td>(5434, 453)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$EM_{t+1}$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note
- Assuming true solutions
- Correct solution for L=150, Invalid solution for L=120
- Correct solution for L=150, Nonfeasible solution for L=120
- Incorrect solution for both L=150 and L=120
- Correct solution for L=150, Nonfeasible solution for L=120
- Incorrect solution for both L=150 and L=120
- Incorrect solution for both L=150 and L=120
- Correct solution for L=150, Incorrect solution for L=120
estimated value and solution point in the feasible region. The critical production resource is distributed accordingly to the crop production with the highest value of net profit.

(c) A correct formulation of the objective function to reflect behavior of production response is the key to success in using the simultaneous model. The poor performance of the simultaneous model (with maximization of net profit as its objective function and with demand and supply constraints) in estimating production should not be overlooked. This model does not give the maximum value in the objective function and use of the model will underestimate production. The reference model, Table 1, is also a simultaneous model but with maximization of the surplus as its objective function.

(d) As mentioned in (a), any hybrid model begun by running a programming model will give poor results in simulating production response. In practice, using a programming model (PM\textsubscript{t}) as a predictive model is very difficult. Unlike an EM\textsubscript{t}, the PM\textsubscript{t} does not have the capabilities of a search method such as the least-square-error method to fit the model to the observed time series data. The PM\textsubscript{t} approach depends on the trial and error method and a high level of expert knowledge about the whole production system to find a proper model to simulate time series data. This approach is inefficient in achieving the accuracy that is readily obtained, with less effort, by an EM\textsubscript{t}. At this stage of research, any hybrid model intended for use in prediction should use the EM\textsubscript{t} as its main structure.
CHAPTER V. NATIONAL AGRICULTURAL

ECONOMETRIC MODELS: A SURVEY

Efforts to develop a RAP model for U.S. agriculture can be minimized by using the existing national and agricultural policy econometric and programming models. For this reason, it is imperative to investigate the models currently available. In this and the next chapters a descriptive survey method is used to examine some of the existing econometric and programming models. Results from the survey provide a basis for selection of a specific econometric and a programming model to build the RAP model.

Since early 1960, several econometric models for national agricultural policy evaluation and production projection have been developed by various research organizations. Some of these models are: POLYSIM, NIRAP, CARD-RS, CED-CC, CHASE, AGRIMOD, DRI, and WHARTON.

In this section the fundamental characteristics of the crop sector in each model will be described. Each model will be described according to the following items: (a) Primary use - The purpose of using a model can be divided into three categories: structure analysis, policy evaluation, and forecasting (or projection of baseline information). Each model's main purpose will be identified. (b) Aggregation level of input data - There are three dimensions of input data aggregation: crop commodity aggregation, time aggregation, and geographical aggregation. The lowest aggregation level used in each dimension will be indicated. (c) Farm price determination - The method of farm price determination will
be described. (d) Basic structure - The following subjects will be examined: 1) the time structure, either recursive or nonrecursive. A model has recursive structure if the outputs of the model in time period \( t \) are used as inputs to the model for time period \( t+1 \), 2) the interdependence among endogenous variables in the same time period, and 3) any unique feature of specification and estimation used in each model. (e) Yield function - Type of yield function used in each model will be investigated. (f) Short description of the model organization - General organization of each model is briefly described. (g) Validation - Procedures and criteria used in each model will be mentioned.

The respective models will not be quantitatively evaluated\(^1\) because of a lack of research resources. Some of the models are proprietary and their specific structure is not available for distribution. Thus, no attempt will be made to give a detailed description of these models. The descriptions presented here are a digest of the limited model documentations listed in the reference and may differ from the current situation since some of the models are still undergoing revision.

Descriptive Survey of the Models

A. POLYSIM (Ray and Richardson, 1977)

1) Purpose: short-run (0-5 years) and intermediate (5-20 years) policy evaluation. 2) Aggregation level: Four crop commodities (one of which includes corn, barley, oats, sorghum) national and annual. 3) Price determination: A cobweb type approach is used to determine the price, i.e., the farm price at time \( t-1 \) determines the quantity supply at time \( t \); the supply then determines the price

\(^1\)Some attempts have been made to compare some econometric models (Just and Rausser, 1979).
at time t, which subsequently determines quantity demand at time t.

4) Basic structure: The model has a recursive structure. Cross elasticities of harvested acreage with respect to lagged prices help determine harvested acreage and allow for interdependence among commodities. The model extensively uses published elasticities of demand and supply. Basic equations in the model are synthesized by the elasticities, baseline data, and other factors instead of regression estimation. The endogenous variables are sequentially determined. 5) Yield function: Baseline data, elasticities with respect to price and cost, and long-run adjustment are key factors that determine yields. 6) Brief description: POLYSIM was developed to estimate the effects of a policy change on agricultural production from a given set of baseline data. The effect is measured in terms of estimated deviation from given baseline data. Since baseline data of previous years of the projected year are used in estimating the effects, the estimated values are affected by these exogenous baseline data. Polysim requires baseline data which have to be obtained from other sources. The model is executed sequentially for a given year, starting with the livestock sector then sequentially crop, farm income, and consumer expenditure sectors are estimated. 7) Validation: Mean of actual data, mean absolute percentage simulation error, root-mean-square percentage error, and the Theil inequality coefficient were used to compare the actual and the predicted values of the major endogenous variables in the model.

B. NINAP (Jaske, 1977)

1) Purpose: Intermediate and long-run (more than 20 years) projection.
2) Aggregation: Thirty-one crop commodities, annual and national.
3) Price determination: Among these 31 crops, prices and quantities of 21 major crop commodities are simultaneously determined and then adjusted, while the price of each of the remaining crops is sequentially determined in time period t. 4) Basic structure: It does not have a recursive structure. Interdependence among crops is captured through constant cross elasticities. In time period t endogenous variables are sequentially determined except the prices and quantities of 21 major crop commodities. Static market equilibrium of major crops is assumed. 5) Yield function: Previous year yield, fertilizer application, and a productivity index are the key variables for determining yields of the major commodities. Simple Spillman and Cobb-Douglas functions are used for predicting yields of the remaining commodities. 6) Brief description of the model: The organization of the model can be seen by examining the listings of model components, which are: scenario component, productivity simulation component, foreign trade component, general economy component, aggregate farm output component, commodity production and utilization component, equilibrium adjustment component, regional model component, yield component, and land availability component. These components are to be sequentially executed with the model divided into two phases.
The first phase estimates national commodity equilibrium prices and production and adjusts these estimates using the Newton-Ralpson search technique to find the equilibrium of the prices and qualities. The equilibrium solutions then are adjusted to be consistent with the aggregate output from aggregate farm output components. The second phase allocates these national production figures to states and regions but has not yet been fully developed. 7) Validation: Individual components of the model were validated. For instance, goodness of fit between predicted and observed was used to validate aggregate farm output components of the model (Yeh, 1976).

C. CARD-RS² (Ray and Heady, 1972; Reynolds, Heady, and Mitchell, 1975)

1) Purpose: Policy evaluation, short-run and intermediate projection. 2) Aggregation level: Four crop commodities (corn, sorghum, barley, and oats are grouped into feed grains), national and annual. 3) Price determination: A typical approach is: Price at time t is determined by a regression equation with key variables: ending inventory and support price at time period t, and price of last time period t-1. 4) Yield function: A simple yield trend function is used for baseline projection. A production function expressed in terms of input factor changes and their elasticities is used for estimating yield changes from a baseline projection because of a policy change (Heady et al., 1977). 5) Basic structure: The model uses an extensive time recursive structure. Lagged values of many of the endogenous variables are used as independent variables in succeeding years. Interdependence among commodities enters through price determination. For example, the lagged price of soybeans is one of the factors determining the price of feed grain. Two-stage least squares and auto-regressive least squares estimation methods are used. 6) Brief description of the model: The livestock submodel is solved first to estimate livestock demands for crop commodities. Export and industrial demands are added to livestock demand to obtain total demands which are fed into the crop submodels to determine crop prices. These prices are then fed into the livestock submodel to determine livestock prices and livestock demands for crop commodities for the next year. The model is divided into three sequentially executed sections: pre-input, input, and output sections. Prices of crop commodities are recursively determined; for example, soybean price of year t-1 is a factor that determines the feed grain price of year t. 7) Validation: Theil-U coefficients were used as indicators of the model's performance (Ray and Heady, 1972).

D. DRI Agricultural Model (DRI, 1977a, b, c)

1) Purpose: Short-run and intermediate policy evaluation and projection. 2) Aggregation level: Twelve crop commodities; regional and

²The content of this section is based on the report by Reynolds, Heady, and Mitchell (1975).
quarterly.  3) Price determination: All the prices and quantities and other endogenous variables are simultaneously determined.
4) Basic structure: It determines endogenous variables and has a recursive time structure between time periods. Interrelation among the crop commodities is built through prices of their substitutes. An accounting identity is used as a standard procedure to compute domestic disappearance which then is used to determine the price level.  5) Yield function: Input cost, expected output prices, weather condition, and fertilizer utilization are the key independent variables in determining yield.  6) Brief description of the model: The crop sector starts by determining acreage planted and ends by determining farm incomes. Domestic production is determined and then, using an accounting identity, determines domestic disappearance (consumption) which subsequently determines the cash and farm price. The farm price then is multiplied by total market volume to determine farm income (cash receipts). The cash and farm price are fed back to determine domestic production and domestic disappearance.
7) Validation: Correlation analysis of actual and fitted values.


1) Purpose: Short-run and intermediate policy evaluation and forecasting.  2) Aggregation level: Five crop commodities; state and quarterly.  3) Price determination: Various forms of regression equations are used. Supply (stock) and demand (domestic disappearance and export) or previous quarters t-i (i=1,2,3) and other factors (such as per capita disposal income or price substitutes) are the key independent variables in determining price at time t.
4) Basic structure: The model has a recursive time structure and makes extensive use of independent variables of quarterly (seasonal) data. Interdependence among crops is established through the price and acreage planted or acreage harvested equations. For example, soybean acreage planted is a function of cotton acreage harvested, while cotton acreage planted is a function of average price of soybeans.
5) Yield function: Not available.  6) Brief description of the model: The model starts from estimation of acreage planted and harvested, and continues to compute domestic disappearance, net stock, price of commodity, livestock quantity and price, wholesale price index of processed foods and feed, consumer price index of processed foods and feed, consumer price index, and farm income.
7) Validation: Not available.


1) Purpose: Short-run and intermediate policy evaluation and projection.  2) Aggregation level: Ten crop commodities; regional and annual.  3) Price determined from the demand and supply functions.
4) Basic structure: The model has a recursive time structure. Crop
commodities are interdependent because prices and quantities of all crop commodities are simultaneously determined by solving demand and supply equations. The model uses a mathematical programming technique to allocate production resources for each crop among regions.

5) **Yield function:** A generalized function of Mitscherle–Baule–Spillman yield function is used. 6) **Brief description of the model:** The model has a time recursive structure. It consists of 17 submodels integrated into four sectors linked by three markets. The four sectors are a) the input sector, b) the farm production sector, c) the food supply sector, and d) the food consumption sector. The three markets are for farm inputs, agricultural commodities, and consumption. A mathematical programming technique is employed in the input market to determine the planned allocation of cropland, fertilizer, and machinery used for each crop in each region. 7) **Validation:** Actual and fitted values are displayed on a chart.

---

**G. CED-CROSS COMMODITY (Teigen, 1977)**

1) **Purpose:** Forecasting and policy evaluation. 2) **Aggregation level:** Six crop co-modities; national and quarterly. 3) **Price determination:** All the commodity prices, quantities, and other endogenous variables are simultaneously determined. 4) **Basic structure:** The model is recursive with respect to time and the endogenous variables are interdependent. The Gauss Seidal algorithm is used to determine the values of all the endogenous variables simultaneously. 5) **Yield function:** Yield function is used for some of the commodities. 6) **Brief description of model:** Each individual model interfaces with other models through livestock prices and outputs. Farm price of typical grain model used in the CED is determined by supply and demand together. 7) **Validation:** Their-U coefficient and actual and fitted values were displayed.

---

**H. Wharton Agricultural Model (Chen, 1976)**

1) **Purpose:** Short-term econometric forecasting, policy evaluation, and structural analysis. 2) **Aggregate price level:** Thirteen commodities; quarterly, national projection. 3) **Price determination:** Stock-demand ratio and other factors determine price t. 4) **Basic structure:** The model is recursive. Interrelation among crops is built through a quarterly and simultaneous simulation block which includes the livestock sector and the crop sector. Stock demand ratio is used as a principal explanatory variable in price equation. 5) **Yield function:** The yield is a function of total acreage planted, crop weather condition proxy, farm prices, crop index of fertilizer prices, crop productivity trends and other disturbance terms. 6) **Brief description of model:** The model is subdivided into four blocks of equations. The intercommodity block, the annual crop production block, the income-expenditure block, and the micro-macro linkage block. Of these blocks, the intercommodity block is the heart of the model. It consists of
livestock and crop sections and is operated on a quarterly basis with simultaneous structure. Individual crop demand is determined in this block. The crop production block is operated on an annual basis with recursive structure. The supply of crop production is determined and used in the intercommodity blocks to determine prices. The income expenditure block, operated on a quarterly basis with recursive structure, computes production expense and farm incomes. The micro-macro linkage block serves a feedback relationship between the agricultural sector and other production sectors in the economy. 7) Validation: Root-mean-square errors as a percentage of mean resulting from ex-post tests are presented.

Summary

NIRAP is designed for the use of long-run projection. POLYSIM is structured for short-run and intermediate policy evaluation. The remaining models are used for the purpose of either the short-run or intermediate policy evaluation or projection. The level of aggregation in crop commodities ranges from four commodities (CARD-RS) to 31 individual crop commodities (NIRAP). The level of geographical aggregation varies from state levels (CHASE) to the national level (such as CED-CC and CARD-RS). The lowest level of time aggregation varies from quarterly (CED-CC), CHASE, and DRI) to annual levels (POLYSIM, NIRAP, AGRIMOD, and CARD-RS).

Price determination varies greatly from one model to another. To determine prices in time period t, POLYSIM uses a cobweb approach. NIRAP solves supply and demand equations simultaneously; CARD-RS uses a regression equation which uses ending inventory and lagged price; DRI solves all endogenous variables simultaneously; CHASE uses a regression equation in which estimated quantity of supply and quantity of demand are included; AGRIMOD solves a set of supply and demand equations simultaneously; CED-CC solves all endogenous variables simultaneously;
Wharton uses stock-demand ratio as the key determinant. Except for NIRAP, all the models have a recursive structure, but there are differences in degree of recursiveness. Most of the models use one-year and two-year lagged variables. Interdependence among crop production is established through lagged variables through a programming model (AGRIMOD) or through livestock prices (CED-CC). Salient features of each model are: POLYSIM is built mainly on synthesized equations constructed from published elasticities; NIRAP assumes a partial static market equilibrium for some crop commodities; CARD-RS and CHASE extensively use generated data as input to determine outputs; DRI and CED-CC each solve all endogenous variables simultaneously at each time period; AGRIMOD uses both econometric and programming approaches. The CHASE model uses a production function without specifying a yield function. The remaining models have yield functions ranging from a trend (CARD-RS) approach to a sophisticated approach such as the use of weather variables in the DRI and use of Mitscherly - Baule - Spillman production function in the AGRIMOD model.

Various methods were used to indicate the performance of each model. These methods include using a chart to display the actual and the fitted values and using statistical methods to test the deviation of the actual and the fitted values.

Selection of an Econometric Model

Most of the models are designed for policy evaluation and projection of national agricultural production and related activities and they are different in specification and structure. Ideally, a model should be judged by the following criteria: theoretical basis, policy variables
that are complementary to the programming model, accessibility, model life, operation and maintenance costs, and empirical validation.

Theoretical basis

There are various forms of approximation to explain an economic event. These lead to variation in specification of models which often reflect the modeler's judgment. Variation includes differences in the use of key explanatory variables and the functional form used to relate variables. For example, the CARD-RS model uses ending inventory and lagged price as key explanatory variables in determining commodity price, while the CED-CC model determines the price by solving simultaneously all the endogenous variables in the model. Another reason contributing to the model's variation is that the estimation theory for dealing with simultaneous models estimated from a small number of observations is far from complete. A standard estimation procedure is yet to be developed. Consequently, each modeler uses a different estimation procedure. Because there is no standard rigid procedure in specification and estimation, it is difficult to determine if one model has superiority over any other.

Complementary to programming model

The LP model is a crop production and distribution model. Policy variables in the LP model can be categorized as: (a) production technology, (b) resource supply, and (c) other physical and institutional variables.

In the LP model it is not proper to have a policy variable dealing with market price and consumption, while an econometric model ($EM_t$)
should include these variables in its model structure. Most of the EM's surveyed meet this criteria. The CED-CC model can exogenize any commodity sector or endogenous variable with very little effort. This feature allows a policymaker to manipulate large numbers of variables very easily.

Accessibility

Current econometric models are developed by both private and public institutions. The privately owned models (AGRIMOD, DRI, CHASE, and WHARTON) are less accessible than those owned publicly (POLYSIM, CARD-RS, CED-CC, NIRAP). A complete list of regression equations and programs can be obtained from the public institutions giving them an advantage over privately owned models for inclusion in the hybrid model.

Model's life

An econometric model, in general, has a very short life and usually requires annual updating of the equations and the exogenous variables. A model developed by a commercial institution is updated more frequently than one developed at a public institution because the public model frequently is not updated once the researcher has graduated or the project is terminated. The models are not usually updated until a need to again use them arises. Based on updating frequency, the commercial models have the advantage.

Operation and maintenance costs

It does cost more to run a simultaneously determined model (e.g., CED-CC) than a sequentially determined model of similar size (e.g., POLYSIM and CARD-RS). However, the cost differential is not significant
if the simultaneous model is properly structured. Models using the batch process have minimal maintenance costs because of having only the costs of updating exogenous data and equations in the model. None of the model's batch processing costs are significantly smaller by this criterion.

**Empirical validation**

Because various procedures and criteria were employed by each modeler to validate the performance of his model, comparison of the results is not possible. Furthermore, without a standard validation procedure and a standard set of criteria, comparison of the models is meaningless. At this stage, no conclusion can be drawn to indicate which model performs best.

Since AGRIMOD, DRI, CHASE, and WHARTON are proprietary, these models are off the selection list. The remaining candidates are: POLYSIM, NIRAP, CARD-RS, and CED-CC models. Because of POLYSIM's synthetically based nature, it is not suitable for market projections. Also, NIRAP is not designed for short-run projections. Only CARD-RS and CED-CC remain. Between the two, CED-CC has a better structure for making policy analysis and therefore it is logical that this model should be selected as the econometric component in the hybrid model.
CHAPTER VI. NATIONAL AGRICULTURAL PROGRAMMING MODEL

SURVEY OF CARD-LP MODELS

For the past two decades, Heady and his associates at The Center for Agricultural and Rural Development, Iowa State University have developed various programming models and applied them to evaluate impacts of alternative policies on U.S. agricultural production under various settings of production problems. The programming models at CARD include Linear Programming (LP) and Quadratic Programming (QP)\(^1\) models. However, only the LP models will be surveyed in this section. The reason for the exclusion of the QP models is that they are relatively expensive to run and the additional information gained from the QP could be obtained from the econometric component of the hybrid model to be developed.

The objective of this survey is to investigate applications of CARD LP models in the past and to provide the background of specific LP models which could be used as a programming component in the hybrid model. It is believed that considerable expense can be saved by building the hybrid model by using an LP model currently operational at CARD. This chapter starts with a survey of benchmark LP models and their past applications and is followed by a description of the CARD-NRED LP model to be used as a component in the hybrid model.

\(^1\)For information about the CARD QP models, see Olson, Heady, Chen, and Meister (1977).
Descriptive Survey of Benchmark Programming Models

CARD LP models can be classified with respect to different structures and applications into eight benchmark models. Structural differences encompass the number of producing areas, consumption areas, and water regions. The models are:

A. Egbert-Heady Model

This is labeled as the prototype model used for the analysis of U.S. agricultural production. The model includes 104 producing areas (PA) and one consuming region (CR). Crop production activities are defined on the basis of these regions. The model was used to find optimum spatial production and land use patterns for wheat and feed grains under various production settings, to estimate competitive rents for cropland, and to estimate the prices of wheat and feed grains (Egbert and Heady, 1961). The model later was expanded to include 10 consuming regions and applied to find ex-post and ex-ante spatial production-distribution patterns. The model also was extended by including soybeans and cotton and by increasing the number of PAs from 104 to 122, and used to find efficient spatial patterns of crop production.

B. Heady-Whittlesey Model

The model includes 144 PAs and 31 CRs. In addition to production activities, wheat-feed transfer activities and transportation activities were included. Three land classes in each PA were specified. The model was used to investigate production-distribution patterns under alternative assumed production settings. Land withdrawn to eliminate production surpluses and supply controls to remove marginal land were investigated. Whittlesey and Heady (1966) used the model to estimate regional equilibrium production prices and land values under several government program alternatives.

Heady and Skold (1965) also used the model, with inclusion of additional feed grain soybean rotation activities, to project agricultural production capacity under each of the assumed technical changes and population growths. Adding crop acreage quota restraints to the model, Madsen, Heady, and Nicol (1975) used it to evaluate tradeoffs in farm policy.

For detailed description of each of the models mentioned in this section, see Heady and Srivastava (1975) and other CARD reports listed in the reference section.
The farm policy alternatives included long-term, land rental programs with various levels of restrictions on the concentration of land diversions.

Sonka and Heady (1973) used the model to analyze impacts of farm policy on rural income and employment. The model's solutions were summarized into 10 farm production regions. The summary is linked to an input-output (I-O) analysis to evaluate secondary impact on rural income and employment. The basic coefficients in the I-O analysis were derived from Schluter (1971). Four alternatives: free market, land retirement, and two bargaining power situations were analyzed.

C. Brokken-Heady (1968) Model

This is the first published model designed to analyze crop and livestock production simultaneously within the framework of inter-regional competition. The model includes 157 PAs and 20 CRs. Feed grain transfer and cotton lint activities were included in the crop production sector. The livestock production sector included the following activities: milk cows, beef cows-feeder calves, yearling feeders, beef feeding and transportation. The model was used to determine optimal land use patterns of crop and livestock production, and to estimate alterations in the production patterns due to changes in crop and livestock costs, production technology, output requirements and other factors. Equilibrium returns to land and equilibrium prices for commodities were also estimated.

D. Eyvindson-Heady-Srivastava Model

The model is similar to the Brokken-Heady model except for the:

(1) Incorporation of various farm size groups and land quality classes,

(2) Use of crop-producing areas as livestock-producing regions,

(3) Inclusion of cost minimization from producing areas within a consuming region to the region center, and

(4) Inclusion of labor and capital constraints.

The model was used to determine the optimal regional production patterns for major crops and livestock and to determine resource shifts to overcome excess production capacity problems.
E. Madsen-Heady-Nicol-Hargrove Model

Models A through D mentioned earlier used land as the only resource constraint. This is the first model incorporating water constraints. The model has 223 PAs, 51 WRs, and 27 CRs. Major activities include crop production, water use, and transfer and transportation. The model was used to project water and related land uses under alternative futures with respect to agricultural policy, technological advance, export level of farm products, magnitude of the U.S. population, and the pricing and public investment policies for agricultural water uses. Commodities included in the model were wheat, feed grains, soybeans, cotton, sugar beets, hay, pasture, milk, pork, beef, and exogenous commodities.

F. CARD-RANN Model (Nicol and Heady, 1975)

This is the first model designed for studying relationships of agricultural production to land and water use and the environment. The model provides the framework for several research studies to be mentioned later. The model has 223 PAs, 51 WRs, and 30 CRs. The delineation of these regions, except three additional market regions, can be found in the Heady-Madsen-Nicol-Hargrove Model. Commodities included in the models are barley, corn, corn silage, cotton, legume hay, nonlegume hay, oats, sorghum, sorghum silage, soybeans, sugar beets, wheat, summer-fallow, dairy cows, beef cows, beef feeding, and pork. Environmental quality aspects of the model include soil loss, nitrogen, and animal wastes.

Heady-Nicol-Madsen (1975) used a variation of the model to project land and water use when various limits on soil loss were imposed.

Dvoskin and Heady (1975) used another variation of the model to project U.S. agricultural export capabilities under various price alternatives, regional production variations and fertilizer-use restrictions.

G. CARD-NWA Model (Meister and Nicol, 1975)

The major differences between the CARD-RANN model and the CARD-NWA model are:

(1) Delineation of producing regions, market regions, and water supply regions (105 PAs, 28 CRs, 51 WRs, and 9 land classes).

(2) The technological coefficients used.

Otherwise, the basic structure between these two models is similar. Both models were designed to analyze interactions of agricultural production, resource uses, and environmental quality.
Dvoskin and Heady (1976) used a reduced version of the NWA model to examine agricultural production under various assumptions of energy supplies and prices. The model uses one land class and handles livestock production on an exogenous basis.

Colette, Heady, and Nicol (1976) used the model to analyze the impact of water rights on agricultural production. The legal constraints incorporated were water transfer between states, between basins, and water transfer due to international agreements, and the ownership of water under the existing water rights system.

H. CARD-NRED LP Model

The CARD-NRED LP model is the latest in the series of these national agricultural sector models available at CARD. The model was a result of a joint research effort between CARD and the Natural Resource Economics Division, Economics, Statistics, and Cooperatives Service (ESCS) of the U.S. Department of Agriculture. The model is a scaled-down and modified version of the National Water Assessment (CARD-NWA) model described by Meister and Nicol (1975). This modified version is less costly to run as compared with running the CARD-NWA model, but provides an analytic structure adequate for current ESCS needs.

The differences between the CARD-NWA model and CARD-NRED model are:

1. Data sets used were updated by using more current information. For instance, production costs are directly derived from the FEDs budget system (1974, 1975, 1976). Various data sets in the exogenous sector were revised and updated.

2. Livestock sectors are now exogenized. Livestock feed demand, livestock water demand, and livestock nitrogen supplies are pre-estimated and used to adjust the corresponding RHS values of commodity demand, water supply, and nitrogen demand of the model (Boggess, 1977).

3. The water sector is now restructured. Bounded water supply is used in each region to replace the water supply constraints in the Assessment Model. Conversion of irrigated hay to nonirrigated hay is built in the model to allow the saved water on crop use.

4. The model uses five land classes.

The CARD-NRED model divides the 48 contiguous states into 105 hydro-logical areas (Aggregated Subareas or ASAs) defined during the 1975 National Water Assessment. Each area is called a producing area (PA). The 105 PAs are aggregated into 28 market regions.
Vocke, Heady, Boggess, and Stockdale (1977) used the model to evaluate the impact of alternative policies designed to curb pollution problems created by excessive erosion of soil, persistence of certain organo-chlorine insecticides in the environment, feedlot runoff, and nitrogen fertilizer use. Saygidéger, Vocke, and Heady (1977) used the model to analyze trade-offs between a production efficiency goal and a soil loss control goal. Various weights were systematically assigned to each goal in the objective function. Koo, Boggess, and Heady (1978) employed the model to study the interaction of weather with alternative environmental and grain reserve policy. Yield reductions due to weather variation are estimated and incorporated into the model. Dvoskin, Heady, and English (1978) derived an energy model by including the energy transfer restraints to the CARD-NRED model. The energy model was used to evaluate national and regional impacts from alternative energy policies.

Summary

The CARD-LP national models have evolved from a crop production model (Egbert-Heady, 1955) to the agriculture resource-environment models (CARD-RANN, CARD-NRED). The size of the model varies from 104 PAs, 1 land class and 1 CR to 223 PAs, 9 land classes, 30 CRs and 51 WRs.

Structural (decision) variables used expanded from crop production activities to crop, livestock, transportation, water transfer, and others. The number of constraints varies from 106 to approximately 4000 (CARD-NWA). The CARD-NWA model has resource constraints for land, water, energy, and fertilizer, as contrasted with only land resource constraints used in the Egbert-Heady model.
Past applications of the CARD-LP models can be described based on two categories: estimation and policy impact study.

With respect to estimation, the models have been used for estimation of (a) ex-post and ex-ante spatial production distribution. The most popular application is to project future production-distribution with respect to technological change and population growth, (b) production capacities under limited land and water resources, (c) land values and crop prices under a spatial equilibrium framework, and (d) by-products of agricultural production for meeting a given demand.

By varying the assumptions of a CARD-LP model in each run according to a change of policy, the model becomes a valuable tool for policy impact studies. The model was used for the following impact studies: (a) resource restrictions, land retirement, energy supply and price, (b) supply control, (c) rural income and employment generating, (d) imposition of environmental quality restrictions, and (e) institutional constraints such as water rights.

By extending the use of the model for the policy impact study, some researchers also attempted to use the model to design possible production practices. For instance, Vocke, Heady, Boggess, and Stockdale (1977) used the model to investigate production activities from several possible choices to curb problems related to environmental quality. Saygideger, Vocke, and Heady (1977) analyzed potential trade-offs between goals of production efficiency and soil loss control.

Although the CARD-LP models have been used for the purpose of estimations and policy impact studies, the main functions of the
models have been the analysis of potential intermediate and long-run impacts of alternative policies on agricultural production patterns. These models were not being used extensively for explanation of past events and were not being validated for their capability in making positive as opposed to normative predictions.

The models such as the CARD-NWA models have demonstrated their flexibility in analyzing problems related to the potential interaction of agriculture, resources, and the environment. Further applications of the models to policy analysis will depend upon the skill of the modeler, data availability, and the computation capability of the computer.
CHAPTER VII. CARD-NRED RECURSIVE ADAPTIVE PROGRAMMING MODEL

The CARD-NRED RAP model consists of two components: An econometric component represented by the CED-CC model and a programming component represented by the CARD-NRED LP model. The formulation of each component is described.

Econometric Component (see Teigen (1977) for details)

The model consists of 127 exogenous variables and 164 endogenous variables represented by 164 regression and identity equations. These equations are divided into 10 groups.

Retail Demand Equations
(a) retail demand equations for the livestock model,
(b) auxiliary retail price equations in the poultry sector, and
(c) retail demand equations for the dairy model.

Retail Product Supply Relations in the Dairy Model

Farm Demand Equations for the Livestock Model

Capital Stock Equations
(a) investment demand equations in the livestock model, and
(b) livestock inventory identities.

Livestock Supply Equations
(a) product supply relationships in the pork model,
(b) product supply relationships in the beef model,
(c) supply relationships in the veal model,
(d) farm supply relationships in the dairy model, and
(e) supply relationships in the poultry model.
Crop Demand Equations
(a) feed demand relationships in the wheat and feed grain models,
(b) food demand relationships in the wheat and feed grain models,
and
(c) commercial export relationships for the wheat and feed grain model.

Product Stock Equations
(a) ending stock relationships in the livestock model,
(b) ending commercial stock relationships in wheat and feed grain models, and
(d) ending government stock in the wheat and feed grain model.

Crop Supply Equations
(a) planted acreage relationships in the feed grain model,
(b) planted acreage relationships in the wheat model, and
(c) wheat production and seed demand relationships.

Supply and Utilization Identities
(a) supply and utilization identities in the wheat and feed grain models, and
(b) supply and utilization identities in the livestock model.

Index Definitions
(a) aggregate price and output index identities in the livestock model,
(b) input cost identities in the livestock model, and
(c) price index identities in the feed grain and wheat models.
The CED econometric model (see Teigen (1977), for explicit mathematical formulation) can be expressed as

$$y_{it} = a_{it} + \sum_{n=1}^{164} b_{in} y_{nt} + \sum_{n=1}^{164} (b_{2n} y_{nt-1} + b_{3n} y_{nt-2})$$

$$+ \sum_{m=1}^{127} (b_{4m} z_{mt}) + e_{it}$$

where $i=1,\ldots,164$.

The Gauss-Seidal algorithm is employed to determine the values of all the endogenous variables simultaneously.

Twelve endogenous variables are selected from the 164 endogenous variables as the linkage variables. These variables include national crop farm prices and production of the six major crops: wheat, corn, barley, oats, sorghum, and soybeans. The national prices are then converted to the regional farm prices by multiplying weighted regional price ratios which are estimated from state and national (1973-1977) price data. Estimated production is fed into the LP as the target production.

Programming Component

The CARD-NRED model (see Huang, Weisz, and Alt, 1979) is a reduced version of the CARD-NWA model (Meister and Nicol, 1975). The model is used as the programming component in the RAP model. It divides the 48 contiguous states into 105 hydrological areas (Aggregated Subareas or ASAs) (Figure 7) defined during the 1975 National Water Assessment.

\footnote{Actually, some of the regressions in the CED-CC model are in non-linear form. The expression (1)' is used for illustration purposes.}
Figure 7. The 105 producing areas
project. Each area is called a producing area (PA). The 105 PAs are aggregated into 28 market regions (Figure 8).

The component has six set of restraints: (a) land restraints for five different land groups which are further disaggregated by: 1) dry cropland, and 2) irrigated cropland, (b) water supply restraints, (c) commodity transfer rows, (d) nitrogen fertilizer restraints, (e) water transfer and canal capacity rows, and (f) crop production constraints.

The activities in the programming component include: (a) **Crop production activities:** A crop production activity is defined as a crop rotation on a specific land class using specific tillage and soil conservation practices. Twelve crops are included in the model. These are: corn, corn silage, sorghum, sorghum silage, nonlegume hay, wheat, oats, barley, soybeans, legume hay, cotton, and summer fallow. (b) **Water use activities:** water-buy, water movement by natural flows, and water transfer between supply regions and inter- and intra-basin transfer. (c) **Other water use activities,** (d) **Commodity transportation activities,** (e) **Nitrogen-buy activities.**

As seen in Figure 9, the model consists of 28 market regions and each market region contains one or more producing areas. The constraints in each PA are the land and water supply constraints, while activities are production, water-buy, water-flow, water-depletion, and water for hay use (conversion of irrigated hay to nonirrigated hay). The PAs in a market region are related by commodity demand and nitrogen fertilizer constraints and water transfer activities. Interdependent relations between MRs are incorporated through the national demand constraints.
Figure 8. The 28 market regions with central cities indicated.
and commodity transportation activities. Other water constraints including canal capacity, water delivery to deficit areas, and water to Mexico also are shown in Figure 9.

The mathematical formulation of the programming component is expressed as maximize,

\[
6 \sum_{i=1}^{105} \sum_{j=1}^{k_j} \left( XD_{ijkt} + X_i^{ijkt} \right) p_{ijt} - \sum_{k=1}^{k_j} XD_{ijkt} CD_{ijkt} - \sum_{k=1}^{k_j} X_i^{ijkt} CI_{ijkt}
- M_1 \sum_{i=1}^{6} \left( v_i^+ + v_i^- \right)
\]

Subject to

National production balance restraints

\[
\sum_{j=1}^{105} \sum_{k=1}^{k_j} \left( XD_{ijkt} + X_i^{ijkt} \right) + v_i^+ - v_i^- = Q_{it}
\]

i=1, ..., 6; k_j varies from region to region

Regional production response balance restraints

\[
\sum_{k=1}^{k_j} XD_{ijkt} + X_i^{ijkt} \leq [\bar{B}_{ijt}] \left[ \sum_{k=1}^{k_j} XD_{ijkt-1} + X_i^{ijkt-1} \right]
\]

\[
\sum_{k=1}^{k_j} XD_{ijkt} + X_i^{ijkt} \geq [\bar{B}_{ijt}] \left[ \sum_{k=1}^{k_j} XD_{ijkt-1} + X_i^{ijkt-1} \right]
\]

Land restraints

\[
\sum_{i=1}^{13} \sum_{k=1}^{k_j} VD_{ijkt} XD_{ijkt} \leq LD_{jt}
\]

\[
\sum_{i=1}^{13} \sum_{k=1}^{k_j} VI_{ijkt} X_i^{ijkt} \leq LI_{jt}
\]

j=1, ..., 105
Figure 9. A schematic representing the CARD-ESCS model
Other restraints

The formulation of the other constraints, such as water supply, commodity transfer, nitrogen fertilizer, water transfer and capacity, and crop production constraints can be found in Meister and Nicol (1975).

Where $X_{ijkt}$ (or $X_{ijkt}$) is defined as the quantity of production of crop $i$ using rotation and tillage practice $k$ on dry (or irrigated) land in producing area $j$ in time period $t$. $C_{ijkt}$ or $C_{ijkt}^i$ is the cost of producing one unit of $X_{ijkt}$ or $X_{ijkt}^i$, respectively. $V_{ijkt}$ or $V_{ijkt}^i$ is acres of land used to produce one unit of $X_{ijkt}$ or $X_{ijkt}^i$, respectively. $L_{dt}^j$ or $L_{dt}^j$ is total dry or irrigation land available in producing area $j$ in time period $t$. $\bar{\beta}_{ijkt}$ and $\bar{\beta}_{ijkt}$ are respectively the maximum and minimum proportionate increase or decrease of production of crop $i$ in PA $j$ from year $t-1$ to year $t$; the price elasticities are used to determine their values.

Linkage Variables and Feedback Mechanism

Table 2 lists potential linkage variables to be used for information transfer between an econometric and a linear programming model. However, at this stage of model development, only three sets of endogenous variables are selected as linkage variables to transfer information from the econometric component to the programming component. These three sets (expressed as $Y_{it}$ in the econometric component) are regional crop price $P_{ijt}$, cost of production $C_{ijkt}$ (and $C_{ijkt}^i$), and national aggregate crop production $Q_{it}$. At time period $t$ the values of $P_{ijt}$ and $C_{ijkt}$ are used to revise the coefficient in the objective function; the value of $Q_{it}$ is used as the
Table 2. Potential linkage variables between econometric and linear programming components

<table>
<thead>
<tr>
<th>Output from LP</th>
<th>As input to EM to determine</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Crop and livestock regional production</td>
<td>National production supply and commodity national prices</td>
</tr>
<tr>
<td>2. Regional resource uses: crop acreage, water fertilizer and other production inputs</td>
<td>Cost of input factors: land rent, irrigation and chemicals</td>
</tr>
<tr>
<td>3. Shadow price of each commodity, resource and other institutional regulation</td>
<td>Cost of production, and price of commodity and cost of regulation</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Output from EM</th>
<th>As input to LP to determine</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. National and regional commodity prices</td>
<td>Price coefficients and regional production response</td>
</tr>
<tr>
<td>2. National and regional production demands</td>
<td>Values of RHS of production and regional production flexibility coefficients</td>
</tr>
<tr>
<td>3. Cost of production: machinery, labor, chemical and others</td>
<td>Cost coefficient and regional production response</td>
</tr>
<tr>
<td>4. Technological coefficients: regional crop yield, water and fertilizer use per acre</td>
<td>Crop yield, water use and fertilizer use</td>
</tr>
<tr>
<td>5. Resources availability: land, water and others</td>
<td>Values of RHS of land, water and other rows</td>
</tr>
</tbody>
</table>
value of the right-hand side of the national aggregate production balance restraints.

The final production \[ \sum_{j=k}^{105} \sum_{k=1}^{n} (X_{ijkt} + C_{ijkt}) \] determined by the programming component is used as the linkage variable to transfer information from the programming component to the econometric component. When the final production (denoted as \( Q_{it}^* \)) differs from \( Q_{it} \), \( Q_{it}^* \) will be considered a better estimated value of the actual production and will then be used in the econometric component to adjust the values of other endogenous variables in the component. The adjusted values subsequently are used to determine the linkage variables in time period \( t+1 \). When \( Q_{it}^* \) is equal to \( Q_{it} \), no adjustment is performed (see Figure 10).
Note: $Q_{it}$ and $Q^*_it$ denote the $i$th crop production estimated by the econometric component and the LP component, respectively. IADJ is an indicator for controlling the feedback process. When IADJ=1, the solution of the LP component is fed back to the econometric component.

Figure 10. Feedback process in the recursive adaptive programming model.
CHAPTER VIII. MODEL VALIDATIONS

Test Model

The model expressed by equations (1)' to (5)' is validated by the use of test runs. To test the model, the coefficients of the production flexibility restraints $\beta_{ijt}$ and $\beta_{ijt}$ are required. These coefficients are determined by using production elasticities $\varepsilon_{is}$ with respect to price ($P_{ijt-1}$) and the standard deviation (SD$_i$) of elasticities $\varepsilon_{ii}$. The computation procedures (see Figure 11) are:

Let $T_{ijt} (\phi) = \left[ \frac{P_{ijt}}{P_{ijt-1}} \right]^{(\varepsilon_{ii} + \phi(SD_i))} - \frac{6}{\sum_{s \neq i} \left( \frac{P_{ijt}}{P_{ijt-1}} \right)^{(\varepsilon_{is})}}$

(a) If $T_{ijt} (2) < T_{ijt} (0)$

then $\beta_{ijt} = T_{ijt} (0) + (T_{ijt} (0) - T_{ijt} (2))$

and $\beta_{ijt} = T_{ijt} (2)$

(b) If $T_{ijt} (2) > T_{ijt} (0)$

then $\beta_{ijt} = T_{ijt} (2)$

and $\beta_{ijt} = T_{ijt} (0) - (T_{ijt} (2) - T_{ijt} (0))$

$T_{ijt} (\phi)$ is the percentage change of crop production from year $t-1$ to year $t$, estimated from elasticities of production response with respect to price change. If $\phi$ is equal to 2, two standard deviations of $\varepsilon_{ii}$ are added to $\varepsilon_{ii}$. 
Test methods

Two test methods were used to evaluate the performance of the hybrid model in estimating agricultural production, prices, and levels of other agricultural activities. These two methods are: (1) static simulation and (2) dynamic simulation. Each method is applied to the hybrid model and to the CED-CC model, respectively. Estimated values from these two types of models are compared with actual observations. The difference
between the first and second method is: In the first method, for each
time period, actual observed data are used for all predetermined vari-
ables (includes lagged endogenous and exogenous variables), while in
the second method, the lagged endogenous variables are estimated
recursively, and used as input in the next time period. The first method
is an attempt to conduct "ex-post" analysis. Results from this method
provide information indicating how well the model can perform when error
from input data is removed or kept at a minimum. Results from the second
method provide information indicating how well the model can be used
for multi-period simulation: for example, how seriously the error accumu-
lated in previous time periods will affect the performance of the model
in later time periods. This information is extremely important if the
model is to be used to make ex-ante analysis of more than one time period.

The years 1969 and 1972 were arbitrarily selected for the static
test (or ex-post test) of the hybrid model.¹ Years 1969 to 1973 were
selected for the dynamic tests.²

The validations to be made assume the RAP model is to be used in a
predictive or positive sense: It will be used to predict, and compare

¹In conducting an ex-post analysis, it is necessary to use actual
values for all predetermined variables as input data. Although this
requirement poses no difficulty in the econometric component, it does
pose difficulty in the programming component. The LP component uses
extensively synthesized data that do not have observed values. Further-
more, the ex-post analysis also requires forecast values which should
be outside the sampling period in which all the regression coefficients
in the model were estimated. Therefore, it is an approximation of
ex-post analysis.

²Soybean oil price is not exogenized in the 1972-1976 run.
results with, a period which is already known. While this can be one use of a RAP model, the more general use will be otherwise. Generally, evaluation by a RAP-type model is expected to be made of resource and production potentials under policies or proscribed land, water and environmental conditions which have never before been experienced. The potential production or resource use will be determined by the programming component of the model, then the market impact of this potential will be measured through the econometric model. In this case, there would be no reason to validate the production potential and price outcome against known outcomes. While there would be the general intended use of RAP models, we do compare predictions from the model with actual observations already experienced to gauge the model performance.

Data

The regression coefficients of the econometric component (CED-CC model) were established in 1977 by using historical data from years 1950 to 1977. The data set in the programming component (CARD-NRED LP model) was derived from 1975 LP data base residing in CARD. Regional production (1968 and 1971) is used as initial data. The production costs are adjusted according to cost indices for production, interest, taxes, and wage rates (Agricultural Statistics, 1976). In the static simulation projection, production costs were adjusted for test years 1969 and 1972. Stoecker's (1974) yield functions were used to estimate yields for 1969 and 1972.

In the dynamic simulation test, the costs were adjusted by a constant rate and constant yield was assumed during the test period. The derived
regional to national price ratios (1972-74) were assumed unchanged. The values of elasticities $\varepsilon_{ii}$, $\varepsilon_{is}$ and $SD_i$ used in the test are shown in Table 3. Data in the table were assumed to be constant in the dynamic simulation run.

Table 3. Crop production supply elasticities with respect to price

<table>
<thead>
<tr>
<th></th>
<th>Feed grains</th>
<th>Wheat</th>
<th>Soybeans</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feed grains</td>
<td>.20</td>
<td>.03</td>
<td>.06</td>
</tr>
<tr>
<td></td>
<td>(.15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wheat</td>
<td>.03</td>
<td>.15</td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.20)</td>
<td></td>
</tr>
<tr>
<td>Soybeans</td>
<td>.15</td>
<td>.02</td>
<td>.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(.287)</td>
</tr>
</tbody>
</table>

a) These values are derived from the data given in POLYSIM (Ray and Richardson, 1977). Each value is the sum of acreage and yield elasticities.

b) Figures in the parenthesis represent one standard deviation. \((SD_i)\)

Results

Each year's simulation of the econometric component determines the values for 163 endogenous variables, including livestock and crop production, utilization, and marketing activities. The programming component gives the spatial distribution of thousands of crop production activities and land use patterns in 105 producing areas. Because of space limitations, only key portions of the results are presented.

a) Static simulation tests

Static simulation test results of national and Iowa production of corn, soybeans, oats, and wheat are shown in Table 4. In general, the
model performed well as a positive predictive tool in estimating corn and soybean production at both national and state levels. At the national level, both cases show less than a 5 percent error in estimation. For 1969, on the state level, the error was less than 2 percent and was less than 5 percent in the 1972 simulation. Oats and wheat are minor crops in Iowa. The model performed poorly in estimating production at the state level as well as the national level (more than 13 percent error). Although the model performed well (less than 4 percent error) in estimating wheat production on the national level, it did poorly on the state level estimations (32.9 percent error in 1969 and 9.9 percent error in 1972). These simulation results indicate that the hybrid model might do well estimating major crop production at either state or national levels while performing poorly in estimation of minor crops. However, the poor performance can be improved significantly if a more accurate regional crop production response is available and included in the LP component of the hybrid model. Generally, however, it should be remembered that the model would be used mainly as a normative model in estimating future production or resource use potentials and the possible market impact of attaining these potentials.

b) Dynamic simulation tests

Two dynamic simulation test runs are conducted. One has crop production flexibility restraints in each PA. The other does not have flexibility restraints but does have four Iowa regional production restraints for corn, soybeans, oats, and wheat. Dynamic simulation test results from the first simulation run indicate most of the national
Table 4. Ex-post simulation results

<table>
<thead>
<tr>
<th>Crop</th>
<th>1969 National Production (million bushels)</th>
<th>1969 Iowa Production (thousand bushels)</th>
<th>1972 National Production (million bushels)</th>
<th>1972 Iowa Production (thousand bushels)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Actual</td>
<td>Estimated</td>
<td>Error (%)</td>
<td>Actual</td>
</tr>
<tr>
<td>Corn</td>
<td>4687</td>
<td>4487</td>
<td>.27</td>
<td>1,012,563</td>
</tr>
<tr>
<td>Soybeans</td>
<td>1133</td>
<td>1116</td>
<td>1.50</td>
<td>179,850</td>
</tr>
<tr>
<td>Oats</td>
<td>965</td>
<td>959</td>
<td>.62</td>
<td>93,840</td>
</tr>
<tr>
<td>Wheat</td>
<td>1442</td>
<td>1453</td>
<td>.76</td>
<td>1,320</td>
</tr>
</tbody>
</table>

Note: The data in the table represents the actual production, estimated production, and the error percentages for corn, soybeans, oats, and wheat in 1969 and 1972. The 'National Production' and 'Iowa Production' are given in both million bushels and thousand bushels.
crop production generated by the econometric component was adjusted by the programming component. This causes a significant discrepancy in the estimates of the national crop production and prices between the estimates of the national crop production and prices between the hybrid model and the CED-CC model. The following information can be drawn from figures 12 through 15. (a) The hybrid model using the regional restraints fails to give a better estimation on aggregate national production and price as compared with the estimates generated by using the CED-CC model alone. The failure is due to the fact that the restraints constructed by using national price elasticities are not adequate to represent regional response. (b) The adjustment mechanism in the hybrid model assumes that national aggregated production can be better estimated by summing the individually regional production, than by using national aggregated data as is done by the CED-CC econometric model. This assumption is true only if a set of accurate regional response functions is formulated. To improve the performance of the hybrid model, considerable effort is needed to develop regional restraints.

(c) The time recursive structure, as used by the hybrid model, will accumulate error and pass it on to the next time period. This was found in the results (i.e., estimates of the corn and soybean prices in the

---

3 Crop production adjustments by the LP are indicated by "*" below:

<table>
<thead>
<tr>
<th>Simulation year</th>
<th>Corn</th>
<th>Sorghum</th>
<th>Oats</th>
<th>Barley</th>
<th>Wheat</th>
<th>Soybeans</th>
</tr>
</thead>
<tbody>
<tr>
<td>1969</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>1970</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>1971</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>1972</td>
<td>*</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>1973</td>
<td>*</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>
Figure 12. Corn: comparison of performance of the hybrid and CED-CC models to actual observations.
Figure 13. Soybeans: comparison of performance of the hybrid and CED-CC models to actual observations
Figure 14. Oats; comparison of performance of the hybrid and CED-CC models to actual observations.

- Actual Observations
- Simulation Run #1
- Simulation Run #2
- CED-CC Model

Actual Observations
Simulation Run #1
Simulation Run #2
CED-CC Model

RMSE = 0.157
RMSE = 0.153

RMSE = 214.792

RMSE = 90.98

RMSE = 17874
RMSE = 10201

Figure 14. Oats; comparison of performance of the hybrid and CED-CC models to actual observations.
Figure 15. Wheat; comparison of performance of the hybrid and CED-CC models to actual observations.
figures). To reduce this error, perhaps the regional restraints should be formulated as a function of the endogenous variable in the econometric component rather than by depending heavily on the previous year's production, as formulated in equation (4)'. From these findings, it is suggested that whenever accurate regional response restraints are not available, probably the one-way communication model would perform better between time periods than any model with a recursive structure.

In the second simulation run, the regional restraints specified by (4)' were not included; instead, four regression equations representing corn, soybeans, oats, and wheat production responses were used to construct regional restraints for Iowa. These equations were included in the econometric component and used to generate the RHS values of the regional

\[ \begin{align*}
(1) \quad Y_c^t &= 746326 - 13119.56(P_c^t - P_{c,t-1}) + 35297T - 81412 P_{c,t-1} \\
&\quad \text{(59289)} \quad \text{(177184)} \quad \text{(7373)} \quad \text{(59485)} \quad \text{(t-1)} \\
R^2 &= 0.747
\end{align*} \]

\[ \begin{align*}
(2) \quad Y_s^t &= 53528 + 5.27 A_s^t \\
&\quad \text{(14329)} \quad \text{(0.35)} \\
R^2 &= 0.9345
\end{align*} \]

\[ \begin{align*}
(3) \quad Y_o^t &= -27826 + 5.6 A_o^t \\
&\quad \text{(13694)} \quad \text{(0.55)} \text{t} \\
R^2 &= 0.8658
\end{align*} \]

\[ \begin{align*}
(4) \quad Y_w^t &= 1286.5 - 509(P_w^t - P_{w,t-1}) + 390.04 P_{w,t-1} \\
&\quad \text{(401)} \quad \text{(278)} \quad \text{(183)} \quad \text{t-1} \\
R^2 &= 0.4432
\end{align*} \]

The four regression equations are:

Where \( Y_c^t, Y_s^t, Y_o^t, \) and \( Y_w^t \) = crop production of the four crops: corn, soybeans; oats, and wheat. \( P_c^t, P_w^t \) = national prices of corn and wheat. \( A_s^t, A_o^t \) = planted acres of soybeans and oats. Values for these variables are generated from the econometric components in the hybrid model.
restraints for Iowa. The hybrid model gave the same estimate of national production as the estimates generated by the CED-CC model. Meanwhile, a significant improvement in simulation of the Iowa crop production is achieved as judged by the values of root mean square error (RMSE). This outcome demonstrates that if a better econometrically estimated regional response function is used, the hybrid model can give a better estimate of national production as well as regional production and price.

Figures 16 and 17 give pictures of the forecast accuracy. In these figures, the 45-degree line is the line of perfect forecasts, for which the actual and forecasted percentage changes are equal. The first quadrant contains points for which an increase was forecasted and the increase actually occurred, and the third quadrant contains points for which a decrease was forecasted and the decrease actually occurred. The second and fourth quadrants contain the points representing turning point error, i.e., incorrect forecasts of the direction of change. The figures indicate that results for both simulations generally show turning-point error generally is not a significant problem. Furthermore, there is no significant underestimation or overestimation of change, because most points are quite uniformly scattered around the 45-degree line in the first and third quadrants.

5 The four regression equations in footnote 4 estimate regional production responses better than the use of flexibility constraints constructed mainly by the elasticities.
Figure 16. Some results of dynamic simulation run 1

Note: Each number (i,j) denotes the ith estimate in the jth simulation year. i=1,...,6 refers to prices of corn, soybeans, barley, oats, wheat and soybeans, while i=7,...,12 refers to their productions. j=1,...,4 refers to the simulation years 1970 to 1973, respectively.
FORECASTED PERCENTAGE CHANGE

ACTUAL PERCENTAGE CHANGE

Note: each number (i,j) denotes the ith estimate in the jth simulation year. i=1,...,6 refers to prices of corn, soybeans, barley, oats, wheat and soybeans, while i=7,...,12 refers to their productions. j=1, 2, 3, 4 refers to the simulation years 1970, 1971, 1972 and 1973, respectively.

Figure 17. Some results of dynamic simulation run 2
CHAPTER IX. SUMMARY AND SUGGESTIONS

Because government programs have differential effects on agricultural production, farm income and food prices over time and space, there is a need for an analytic model that incorporates temporal and spatial characteristics in detail. A hybrid model which combines an econometric component and a programming component is one of the potential models that possesses these characteristics. A recursive adaptive programming hybrid model subsequently is developed. The following is a summary of the development of the RAP hybrid model and suggestions for further improvement.

Examining the structural and functional differences between econometric and programming models suggests that for short-run policy analysis, an econometric model is the natural tool. Its shortcomings in analysis of interaction between regional production and technological change, resource supply, and handling unprecedent government regulation can be removed by integrating a programming component with an econometric component.

Methods of combining an econometric component with a programming component are investigated. The one-way communication method may either encounter nonfeasible solutions or over-estimate production. Simultaneous solution may have computational difficulty if the econometric component contains a nonlinear regression equation. The recursive interactive programming model may encounter infeasible solution or over-estimate production. To remedy these shortcomings, a recursive
adaptive programming (RAP) approach is suggested. The RAP model developed in this study uses CARD-NRED LP to validate the projection made by the CED-CC econometric model, and adjusts the projection when the values are outside the feasible region defined by the LP. Furthermore, the LP also provides structure and other policy variables to enhance the analytic capability of the econometric component.

Price, production, and production cost are the main linkage variables between the two components. These linkage variables allow communication between the two components, between the periods and within each time period.

In contrasting the RAP model with some other hybrid models, a simple experiment is conducted to demonstrate the characteristic of some hybrid models. A mathematical formulation is employed to illustrate the essence of the RAP model as well as its applications as a national, regional, and regional and national model.

Since efforts to build a RAP model for U.S. agriculture can be minimized by using existing national agricultural econometric and programming models, a survey of currently available econometric models and benchmark programming models at CARD is conducted. The survey results give a basis for using the CED-CC econometric model and the CARD-NRED LP model to build the CARD-NRED RAP hybrid model.

A general use of hybrid models such as RAP would be to estimate future production potentials and resource use possibilities under policies and other conditions never realized or attained in the past. It gives what "could exist" under these conditions of the future. The econometric
component of the model would estimate the market impacts if these future potentials were realized. In the validation made, we departed from this general use of RAP models and compared the model results with observations actually realized in past years. These validations are summarized below.

A simple one land class RAP model is constructed and validated. The test methods include static and dynamic simulation tests. The static tests indicate the model performs well in examining corn and soybean production at both national and regional (state of Iowa) levels, but they show inconsistencies in estimating the production of oats and wheat. The dynamic simulation tests show both national and regional estimates follow the general movement of the observed data, but results also indicate cumulative errors. These errors mainly are from inadequate production of the regional restraints. Currently, the model could be used for national policy analyses. Considerable effort to improve the regional restraints is required in order to use the model for regional or joint national and regional positive policy analyses.

Considerable effort is required to use the RAP model for any policy analysis where interest is in predicting in a positive framework. The effort includes (a) modification of the econometric and programming components according to the nature of the policy problem to be analyzed. Some regression equations may have to be added to the econometric component and/or some constraints may have to be built into or removed from the programming component to reflect the implementation of different policies, (b) maintaining both econometric and programming models to
reflect the current situation. When new data become available, the regression coefficients in the econometric component or the technological coefficients in the programming model may have to be updated, (c) estimation of needed input data for the projection year. The work includes preparing values of exogenous variables used in the econometric model and values of the RHS in the programming component.

Thus, to use the RAP model efficiently as a predictive tool in terms of getting results in a short time, a team of researchers consisting of an econometrician, a programming specialist, and a system analyst is needed. The econometrician would be in charge of updating and modifying the econometric component, and preparing input data for that component. The programming specialist would have responsibility in updating the coefficients of RHS values and modifying the component. The system analyst would have responsibility for linking the econometric and programming components, and for arranging consistent data used in both components. Communication among these three specialists is essential if the model is to be used to deliver meaningful answers to specified problems.

At this stage of development, there are numerous areas to be improved when the model might be used for predictive purposes. Some of them requiring immediate attention are: (a) A procedure to estimate production costs should be built into the econometric component to provide a more realistic estimation of the production costs. (b) The procedure to generate input data for the programming component needs to be streamlined. A new matrix generator geared to the RAP model should be developed.
Some modifications are needed to update crop rotation in some production areas to reflect current production practices. (c) The trade-offs between using the RAP model and using an econometric or a programming model directly should be further investigated. Although the RAP model could become a "universal" policy model, some policy problems may be solved more economically by using either a pure econometric or a pure programming model. Identification of problems that can best be solved by the RAP model should be conducted. (d) Development of regional flexibility restraints if the RAP model is to be used to project regional production activities should include economic explanatory variables and regional structural variables. Work by Sahi and Craddock [1974] can be used as a basis for their development.
REFERENCES

Agrawal, R. C. and E. O. Heady

Agricultural Statistics

Baum, Kenneth H.

Baumol, W. J.

Boggess, W. G.

Boss, G., H. D. Bouland, P. Doering, W. E. Gahr, and P. L. Valentine

Boussard, J. M. and M. Petit

Brokken, R. F. and E. O. Heady

Chase Econometric Associates, Inc.
1977 Chase Agricultural Model, Bala-Cynwyd, Pennsylvania.

Chen, Dean T.
Chow, G. C.

Colette, W. Arden, E. O. Heady, and K. J. Nicol

Day, Richard H.

D'Azaro, John and Constantine Houpis

DRI Agricultural Service

DRI Agricultural Service

DRI Agricultural Service

Ducot, Elizabeth R. and Alexander H. Levis

Dvoskin, D., E. O. Heady, and B. C. English

Dvoskin, Dan and E. O. Heady
Egbert, A. C. and E. O. Heady  

Farm Enterprise Data System  

Farm Enterprise Data System  

Farm Enterprise Data System  

Friedman, Milton  

Harrison, Michael J.  

Heady, E. O. and M. Skold  

Heady, E. O.  

Heady, E. O., K. J. Nicol, and H. C. Madsen  

Heady, E. O. and Uma K. Srivastava  

Hiller, A. S. and D. G. J. Lieberman  
Huang, W. Y., R. Weisz, and K. Alt  

Intriligator, M. D.  

Jaske, Michael R.  

Johnston, J.  

Just, Richard E. and Gordon C. Rausser  

Kelso, M. M.  

Koo, W., W. G. Boggess, and E. O. Heady  

Lee, T. C. and S. K. Seaver  


Meister, A. D., E. O. Heady, K. J. Nicol, and R. W. Strohbehn  
Meister, Anton D. and Kenneth J. Nicol
1975 A Documentation of the National Water Assessment Model of Regional Agricultural Water Use and Environmental Interaction. Center for Agricultural and Rural Development, Iowa State University, Ames, Iowa.

Meister, Anton, D., Carl C. Chen, and Earl O. Heady

Meister, Anton D.

Nicol, K. J. and E. O. Heady

Nicol, K. J., E. O. Heady, and H. C. Madsen

Olson, K. E., E. O. Heady, C. C. Chen, and A. D. Meister

Penn, J. B., B. A. McCarl, L. Brink, and G. D. Irwin

Plessner, Yakir

Quance, C. L. and L. G. Tweeten
Quance, C. L.  

Ray, Darryll E. and E. O. Heady  

Ray, Darryll E. and James W. Richardson  
1977  Technical Description of POLYSIM. Department of Agricultural Economics, Oklahoma State University, September.

Reynolds, T. M., E. O. Heady, and D. O. Mitchell  

Sahi, R. K. and W. J. Craddock  

Saygideger, O., G. F. Vocke, and E. O. Heady  

Schaller, W. Neill  

Schluter, G. E.  

Shapiro, H. T.  

Sharples, Jerry, A. W. Neill Schaller  
Shumway, Richard and Anne A. Chang  

Sonka, S. T. and E. O. Heady  
1973 Income and Employment Generation in Rural Area in Relation to Alternative Farm Programs. North Central Regional Center for Rural Development, Iowa State University, December.

Sposito, V. A.  

Stoecker, A.  

Takayama, T. and G. G. Judge  

Teigen, Llogy D.  

U.S. Department of Commerce, Interindustry Economic Division  

U.S. Water Resources Council  

U.S. Water Resources Council  

Vocke, G. H., E. O. Heady, W. G. Boggess, and H. J. Stockdale  

Whittlesey, N. K. and E. O. Heady  
Yeh, C. J.

ADDITIONAL COPIES of this publication can be obtained by writing the Center for Agricultural and Rural Development, 578 East Hall, Iowa State University, Ames, Iowa, 50011. Price is $3 each.

ALL PROGRAMS and publications of the Center are available to everyone regardless of race, color, national origin, religion, or sex.