Agricultural Sector Models: The Colombia Case*

Richard E. Suttor**

In recent years the Agency for International Development and other foreign aid donors have given increasing attention to the analysis of sectors of the economies of less developed countries. This interest arose from the realization that a level of analysis intermediate between traditional macro models, on the one hand, and a series of unrelated project studies on the other, was needed to improve decision making by both the donors and LDC governments.

Sector Analysis

The word "sector" can be loosely defined as a set of activities traditionally grouped together and/or convenient for the purpose of analysis. Although most analyses refer to traditional sectors (agriculture, education, health, etc.), the list of activities included may vary considerably among studies; for example, activities such as livestock slaughter may be included in the agriculture sector in one analysis while in another it may be considered part of a non-agricultural sector.

Given that the objective of sector analysis is to provide information for improved decision making by LDC governments and aid donors, it should be possible to define criteria for designing sector analytical models. This paper proposes five rather general criteria. First, there should be a high level of product disaggregation within the sector. This implies that individual crops and livestock products would be treated separately rather than as parts of a larger aggregate in an agricultural sector analysis. Second, an analysis should be comprehensive in the sense that all activities in the sector are included in the analysis. Third, interrelationships both within the sector and between the sector and the rest of the economy should be included in the model. Thus, an agricultural sector analysis might include intrasectoral relationships linking feed grain production activities with livestock production activities which consume feed grain, as well as intersectoral relationships linking fertilizer supplies to crop production activities which use fertilizer, to name two among many important linkages. Fourth, resource constraints should be analyzed. Available land classified by soil types, rainfall and other ecological characteristics is perhaps the most important constraint for agriculture, although labor and various types of capital may also be crucial. Fifth, models should be capable of handling multiple objectives. Ideally, the impact of proposed policies and projects on production, employment, income distribution and/or other objectives

*The views expressed in this paper are those of the author and are not necessarily coincident with those of the Agency for International Development.
**Economist, Sector Analysis Division, Latin America Bureau, Agency for International Development.
should be analyzed, since decisions are in fact typically based on several considerations rather than a single all purpose goal.

Models designed with these criteria in mind will be rather complex and require large amounts of data. An analyst's appetite for bigger and better models is never satisfied because of limited time and resources, however, considerable progress has been made in recent years.

**Agriculture Sector Analyses**

Agriculture sector models with a high level of disaggregation and incorporating a large number of interrelationships utilize mathematical programming and Leontief input-output techniques for the most part. Some recent examples are briefly reviewed in this section. This is not an exhaustive list, but the models chosen serve as illustrations of alternative approaches.

The first Colombia Agriculture Sector Analysis (CAS I) was undertaken by the Sector Analysis Division of the Latin America Bureau in A.I.D. in the spring of 1971. A Leontief input-output table for the entire economy and a linear programming model of agricultural activities were completed by mid-1972. Since then, the LP model has been enlarged to analyze the relationships between nutrition and agricultural production. At the same time, the second round of analysis (CAS II), which will build on the experience of CAS I, is now being undertaken by a Colombian team in the Ministry of Agriculture with assistance from A.I.D. and the U. S. Department of Agriculture.

The input-output table of CAS I consists of 250 sectors of which 72 are primary agriculture activities. Each major crop and livestock product is represented by at least one row and column. This high degree of disaggregation provides a wealth of information on interdependencies both within the agriculture sector, and between agriculture and other sectors.

The LP model includes restrictions on labor, land, capital, and domestic and export markets; minimum nutritional requirements were added later. Several objective functions have been used. The model was solved to determine those levels of production in 1972 and 1975 of the various agricultural activities which (a) maximize employment, (b) maximize value added in agriculture, (c) maximize return to land and capital, (d) maximize various weighted income functions, and (e) minimize the retail cost of food. The first two solutions were to measure possible increases in employment and production--two major goals of the Colombian Plan. Returns to land and capital was maximized not because it was a Colombian goal, but because it provided a standard of comparison and was thought to be a better representation (or simulation) of the sector in the

---

1Analytical results and a bibliography of the documents produced during the analysis can be found in [1] and [11]. All documents are available upon request from the Sector Analysis Division, Latin America Bureau, A.I.D., Department of State, Washington, D. C.
absence of interventions aimed at achieving the Plan's goals. Alternative
weights were attached to the incomes of low, medium and high income groups and
the resulting weighted income functions maximized to gain some insight into
the income distribution problem, which is another major concern of the Plan.
Finally, the retail cost of food was minimized subject to the restrictions that
minimum amounts of nutrients be provided with currently available land and labor.

Because of limited data availability and time, CAS I was a national analysis
with no regional breakdown. The first priority was an analysis which met the
five criteria listed in a previous section. Although CAS I met these criteria,
there was a clear need to improve the data base, particularly data for the land
constraints. The models and data collection techniques for this second round
of analysis will be discussed in the next section.

Hall and Heady [4] provide an excellent review of interregional models.
Several models have been developed which assume fixed demands to be met at
minimum cost. The country is divided into several consuming regions and producing
regions. A set of demands is specified for each consuming region. Production
in each producing region is constrained by available land and possibly other
limited resources. Transportation activities are included to allow shipment of
products among regions. The objective is to determine the regional allocation
of production which minimizes the combined cost of production and transportation.
These models have been applied to limited sets of commodities (e.g., feed grains),
rather than attempting to cover all agricultural commodities in one model.

As an alternative to assuming fixed demands, linear programming techniques
have been developed to allow simultaneous determination of prices and quantities
consumed and produced. Two recent examples are models developed by the World
Bank and cooperating agencies in Mexico and Portugal.

The linear programming model of the Mexican agriculture sector [2] incor-
porates national demand functions for agricultural products. Prices, wage
rates, and migration of labor between regions, as well as production levels are
determined endogenously. The production side of the model consists of sub-models
for twenty geographical areas. Demand functions, however, refer to national
markets. Linkage with the non-agricultural sectors is provided by means of the
downward sloping product demand functions and agricultural input supply functions
which are assumed to be perfectly elastic. The production sub-models include
restrictions on land, labor, and in some cases irrigation water. Because of
the relatively high degree of disaggregation on the production side and the
sophisticated treatment of product demand and labor, the model is very large:
1,500 rows, 5,000 columns and 80,000 non-zero elements. A shortcoming of the
model is that it covers only short-cycle crops, which account for only about
one-half of agricultural production. Work is underway on companion models for
tree crops and livestock.

The Portugal model [3] treats both demand and supply on a regional basis.
It includes eleven regions for both production and consumption. Regional demand
functions are used for the 22 final products. Production activities are dis-
aggregated on a regional basis. Interregional transportation, import and export
activities for several products are also included. Restrictions refer to land,
labor, and tractor power and capacity of certain agricultural processing and input industries. An optimum solution includes prices, production, and consumption on a regional basis and the amount of various products transferred between regions and internationally. Although the Portugal model includes regional demand functions and interregional transportation activities, which are not found in the Mexico model, it is smaller (1,000 rows and 5,000 columns) due to less disaggregation on the supply side.

The objective in both the Mexico and Portugal models is maximization of net social payoff [6]. It can be shown that the optimum solution is consistent with a competitive equilibrium [7]. Therefore, if the theoretical concept of perfect competition is appropriate for the agriculture sector, these models should be useful for making predictions. The Colombia model is designed primarily for normative analysis and is not as well suited for making predictions.

The National Model of Agricultural Production Response developed by the U. S. Department of Agriculture [10] uses recursive linear programming models of representative farms. Prices are introduced by means of a cobweb mechanism. National demand functions are formulated in terms of current prices and quantities, that is, price in year t is a function of quantity in year t. Farmers, however, determine how much to produce on the basis of lagged prices; thus, prices in year t are inputs into a series of representative farm linear programs which are solved (using profit maximization objectives) to determine production in year t+1. By "blowing up" the production levels indicated by the farm models and aggregating, estimates are obtained for national production which in turn, determine prices through the demand functions. While the USDA model treats a limited set of commodities, this type of model could be applied to an entire agriculture sector.

Design of CAS II

It is apparent from this brief survey that there is a considerable wealth of analytical techniques which can be applied in an analysis of the agriculture sector. For the second round of the Colombian analysis we have chosen a model which emphasizes interrelationships among subsectors. A new input-output table based on newer and higher quality data and a linear programming model incorporating the Leontief intersectoral relationships are planned. At the same time we are constructing data sets which can be used for further analyses. The sector analysis is conceived as an ongoing process not necessarily limited to one model. I will briefly describe this "first" model (See appendix for mathematical formulation) and its accompanying data base and indicate some additional models that may be implemented by drawing on the data base at a later date.

The input-output table consists of approximately 600 rows and columns. It includes approximately 200 primary agriculture activities disaggregated by commodity and by technology for major products. In addition, there are about 120 agricultural processing activities and 40 household groups classified according to income level, type of income (wages, salaries, rents, interest and profits) and sector (agriculture and non-agriculture). Data is being
obtained from a relatively complete set of censuses conducted in 1970, a national household consumption survey also conducted in 1970, and special surveys of farms and rural households. Although the input-output matrix can be used in the usual way to estimate the direct and indirect impacts of final demand on interindustry outputs, our primary purpose in constructing it is to obtain internally consistent coefficients for the linear programming model.

Agricultural production activities for the linear programming model will be defined in a four step procedure. First, detailed data on inputs from the special survey of 20,000 farms will be analyzed to obtain representative "input configurations" for each crop and livestock product [12]. An input configuration will specify the amount of each input (fertilizer, pesticides, labor, machinery, etc. per hectare of land). The set of input configurations will represent the entire range of technologies observed throughout the country.

At the same time, soil surveys, rainfall and other climatic data will be analyzed to develop a set of "ecological areas." The entire agricultural land area will be mapped. Because of the great diversity of the Colombian landscape, a given type of ecological area is likely to exist in many different geographical areas. Thus, although there may be a fairly small number of ecological types (say 20 or 30), the number of contiguous parcels of land will be numbered in the thousands.

After the input configurations and types of ecological areas have been defined, they must be matched, i.e., it must be decided which input configurations are feasible for each type of ecological area. All input configurations which now exist in a given type of ecological area (as indicated by the farm survey) are obviously feasible, but, in addition, certain input configurations not observed in the given type of ecological area but observed in other types may be deemed feasible.

Finally, the output associated with each feasible input configuration will be estimated for each type of ecological area. This will be based on the farm survey data and, in some cases expert opinion. We will then have a data base consisting of a set of activities defining alternative technologies for each agricultural product that is considered feasible in each type of ecological area. The data will be stored on tapes to facilitate the use of "matrix generator" computer programs which can be written to generate coefficient matrices for a wide range of programming models.

The linear programming model includes restrictions for land, labor, exports, and perhaps other limited resources, along with a set of balance equations. The land restrictions refer to types of ecological areas. For each type there is at least one restriction stating that the amount of land used is less than or equal to the total agricultural land area in that type of ecological area throughout the country. In some cases only one restriction is needed, but in those where multiple cropping is common, perhaps quarterly or even monthly constraints will be needed.

2For a detailed discussion of the input-output construction, see [8].
Monthly labor restrictions refer to the entire agriculture sector. Each restriction states that labor used in agricultural production is less than or equal to the available supply of labor that month.

There is a balance equation for each commodity produced or imported in the country. These are mathematical statements of the accounting equality: production plus imports equals interindustry consumption plus household consumption plus exports plus capital accumulation.

For each household group there is a balance equation stating that income equals expenditures plus savings. Associated with these rows are activities which specify how expenditures are distributed among commodities. In this connection, we have devised a scheme which allows the incorporation of income elasticity estimates. (See appendix for details).

Export activities are constrained by available foreign markets. Another set of restrictions are used to bound the capital accumulation activities. Other restrictions, such as upper limits on agricultural credit, may be added for some runs.

Objective functions in CAS II will be similar to those in CAS I. Specifically, employment, value added, weighted income and net foreign exchange earnings will be maximized. As the work progresses other formulations may evolve.

The exact size of the linear programming matrix depends on the number of alternative technologies and the number of ecological areas. It is expected, however, to contain 600 to 700 rows and 3,000 to 4,000 columns. Although it is relatively large, it is small enough to allow considerable experimentation at a reasonable cost.

The program solution includes the level of production technology utilized, and exports or imports of each commodity and income levels for household groups that optimizes the objective. It is a general equilibrium solution similar to that obtained from a Leontief input/output model with the household sectors endogenous, but different in at least four important respects. First, the proportion of the total supply of each commodity which is domestically produced or imported is not fixed, but rather is a variable determined by the model solution. Second, the proportion of domestic production produced by a given technique is also variable. Third, the distribution of income among household groups is variable. Fourth, while input/output solutions are unconstrained, the linear programming solution must conform to available supplies of land, labor and other limited resources.

The optimum solution specifies a land use pattern for each type of ecological area. This can be used to derive production in contiguous regions. By assuming that each parcel of land of a specified ecological type, regardless of location, has the same land use pattern and knowing the land area of each type in a region, production for each agricultural product can be derived. This regional distribution of production does not take account of the spatial distribution of consumption. Data from the household surveys, and the population census can be used to allocate national consumption implied by the optimum
solution to the various regions. Then simple fixed production-fixed consumption models can be used to determine the pattern of interregional trade which results in the minimum total cost of transfer [6]. The optimum pattern can be determined independently for each commodity and with very low computing costs.

The results of this analysis will indicate the efficiency of the spatial distribution of production with respect to transportation costs. Also, the implied agricultural labor requirements can be calculated and compared with the available supply of labor in each region. If interregional transportation costs are relatively low and labor use is consistent with the spatial distribution of labor, the construction of more complex interregional models may not be worth the cost. However, if there are serious problems with the spatial distribution of production, it may be worthwhile building a fixed demand-variable production interregional model [4]. Agricultural production activities would be drawn from the original data base. Land and labor constraints would be defined for each region. The objective would be the minimization of the combined cost of production and transportation to meet fixed regional demands. Such a model would be excessively large if the same degree of disaggregation in the production activities is maintained.

For example, if only ten production and consumption regions are used, the dimension of the matrix would be about 2,000 rows by 20,000 columns. This would be too unwieldy and costly. To build an interregional model of practical dimensions, considerable aggregation would be necessary over technological levels, commodities and/or ecological types. In any case, much of the richness of the data relative to production alternatives is lost.

At the present time we have no plans to build price endogenous models. However, if suitable data can be obtained for estimating demand curves, models using either national or regional demand functions can be constructed, again drawing on the data base of agricultural production activities. Practical analytical techniques are available, as exemplified by the models for Mexico and Portugal.

The data base can be used for building farm linear programs. We plan to build farm models for areas which are favorably located for exports to estimate supply curves for selected products. We also intend to analyze representative farm situations with respect to nutrition. Although work has been done on minimum cost diets for urban consumers in Colombia [5], little is known about the farm family, which typically produces part of its food and purchases part. Nutritional deficiencies among rural families and programs and policies for improving nutrition, an extremely complex problem complicated by the wide variation in production possibilities and available food supplies in different parts of Colombia, can be investigated by linear programming models of the farm firm with both production and consumption activities. Finally, if and when enough farm models have been constructed, it may be possible to construct a national model along the lines of the USDA National Model of Agriculture Response.
APPENDIX

The first section of this appendix is a list of the restrictions and objective functions of the Colombia agriculture sector model. The second section is a more detailed explanation of the relationships between household income and consumption.

Restrictions and Objective Functions

The model refers to an economy with I commodities. Each commodity can be produced by J techniques. There are K land types or ecological areas. Finally, the population is divided into M household groups. In the implementation of the model, most if not all the non-agricultural goods will be produced by only one technique, i.e., J = 1 for that commodity, however, to simplify the already cumbersome notation the symbol J refers to all commodities with the understanding that J is different for different commodities. Similarly, the land restrictions will probably refer only to agricultural use, and therefore, the index k is irrelevant for non-agricultural activities.

The following six sets of restrictions will appear in all formulations of the model. For some applications, additional restrictions will be needed. If, for example, capital for a set of crops is to be constrained, a restriction similar to the labor or land restrictions could be added.

1. Land restrictions: Use of agricultural land cannot exceed the available supply.

\[ \sum_{i=1}^{I} \sum_{j=1}^{J} b_{ijk} X_{ijk} \leq R_{kl} \quad k = 1, 2, \ldots, K \]

where, \( X_{ijk} \) is production of a unit (expressed in money terms) of product \( i \) with technique \( j \) on land type \( k \).

\( b_{ijk} \) is the amount of land type \( k \) required in month \( l \) to produce a unit of product \( i \) with technique \( j \).

\( R_{kl} \) is the available supply of land type \( k \) in month \( l \)

2. Labor restrictions: Employment is less than or equal to the available supply of labor.

\[ \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} c_{ijk} X_{ijk} \leq S_{l} \quad l = 1, 2, \ldots, 12 \]

where, \( c_{ijk} \) is the amount of labor required in month \( l \) to produce a unit of product \( i \) with technique \( j \) on land type \( k \).

\( S_{l} \) is available supply of labor in month \( l \)
3. Commodity balance equations: Production plus imports equals inter-industry consumption plus household consumption plus exports plus net capital accumulation.

\[ \sum_{j=1}^{J} \sum_{k=1}^{K} X_{ijk} - \sum_{n=1}^{N} \sum_{j=1}^{J} \sum_{k=1}^{K} a_{njk} X_{njk} + (M_i - E_i) - K_i - \sum_{m=1}^{M} e_{im} d_{im} (U_m - V_m) = \sum_{m=1}^{M} d_{im} Y_m \]

\[ i = 1, 2, \ldots, I \]

where, \( a_{njk} \) is the input of commodity \( i \) required to produce one unit of commodity \( n \) by technique \( j \) on land type \( k \)

\( M_i \) is imports of commodity \( i \)

\( E_i \) is exports of commodity \( i \)

\( K_i \) is net capital accumulation of commodity \( i \)

\( d_{im} \) is average propensity to consume commodity \( i \) by household group \( m \)

\( e_{im} \) is the income elasticity of demand for commodity \( i \), household group \( m \)

\( Y_m \) is income of household group \( m \) which is allocated among commodities according to average propensities to consume in the base year

\( U_m \) is income of household group \( m \) in excess of \( Y_m \)

\( V_m \) is the amount by which total income of household group \( m \) falls below \( Y_m \)

4. Income balance equations: Household income equals factor payments

\[ \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} f_{ijkm} X_{ijk} - U_m + V_m = Y_m \quad m = 1, 2, \ldots, M \]

where, \( f_{ijkm} \) is payment received by household sector \( m \) from the production of one unit of product \( i \) with technique \( j \) on land type \( k \).

5. Export markets: Exports are less than or equal to available foreign market demand.

\[ E_i \leq E_i \]
where, $\bar{E}_i$ is the estimated maximum quantity of commodity $i$ that can be sold in foreign markets.

6. Capital accumulation targets

$$K_i \leq \bar{K}_i$$

where, $\bar{K}_i$ is target capital accumulation for commodity $i$. For some commodities this could be a minimum level, while for others it could be a maximum. These could also be fixed targets which are to be met exactly, in which case the constraint would be an equality.

7. Objective functions

(a) Maximize value added or employment

$$\text{Max } \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} v_{ijk} X_{ijk}$$

where, $v_{ijk}$ is value added or employment per unit of production of product $i$ with technique $j$ on land type $k$.

(b) Maximize weighted income

$$\text{Max } \sum_{m=1}^{M} w_m (u_m - v_m)$$

where, $w_m$ is the weight assigned to income of household group $m$.

(c) Maximize net foreign exchange earnings

$$\text{Max } \sum_{i=1}^{I} (E_i - M_i)$$

Some objective function rows could be used as restrictions, e.g., limited foreign exchange reserves could serve as a restriction when the objective is maximization of value added.

**Household Income-Consumption Relationships**

For household group $m$, let $p_m$ be population, $y_m$ per capita income, and $c_{im}$ per capita consumption of good $i$. The basic assumption is that per capita consumption of any good is a function of per capita income:

$$c_{im} = f(y_m)$$
We define \( d_{im} \) as the average propensity to consume good \( i \) in the base year

\[
d_{im} = \overline{c_{im}} / \overline{y}_m
\]

where the bar denotes base year levels.

If per capita income remains at the base year level, total income in the year to which the model refers is:

\[
Y_m = P_m \overline{y}_m
\]

This is the right hand side of the income balance equation for household group \( m \).

Noting that the income elasticity of demand -- proportionate change in consumption divided by proportionate change in income -- is the derivative of consumption with respect to income divided by the average propensity to consume, it follows that the marginal propensity to consume (the derivative) equals the average propensity times the elasticity. Thus, income elasticity estimates from various sources can be used to "adjust" the average propensities to consume, which can be obtained from the input/output table. In practice this may be a convenient way of estimating the marginal propensities, although they can be estimated directly if income-consumption curves are fitted.

Defining \( e_{im} \) as the income elasticity in the base year, per capita consumption associated with per capita income levels in the neighborhood of the base year level is approximated as follows:

\[
c_{im} = \overline{c_{im}} + e_{im} d_{im} (y_m - \overline{y}_m)
\]

Multiplying by population,

\[
P_m c_{im} = P_m \overline{c_{im}} + e_{im} d_{im} (P_m y_m - P_m \overline{y}_m)
\]

The left hand side of the equation is total consumption of good \( i \) by household group \( m \); the elements on the right hand side are defined in the commodity balance equations of the linear programming model as follows:

\[
d_{im} y_m = P_m \overline{c_{im}}
\]

\[
U_m = P_m y_m - P_m \overline{y}_m, \text{ for } y_m > \overline{y}_m
\]

\[
V_m = P_m \overline{y}_m - P_m y_m, \text{ for } y_m < \overline{y}_m
\]

122
The linear programming formulation of this relationship can be illustrated with the aid of the following graph. The slope of the line OA is the average propensity to consume in the base year, and the line BD is the tangent of the income-consumption curve.

If per capita income in the program solution happens to be equal to per capita income in the base year, total income for household group \( m \) is \( Y_o \) and consumption of good \( i \) by this household group is \( C_o \). Relating the previous equation to the graph,

\[
Y_o = p_m \bar{Y}_m \\
C_o = p_m \bar{c}_{im} = d_{im} Y_m
\]

The linear programming variables \( U_m \) and \( V_m \) equal zero.

A second possibility is that per capita income in the optimum solution is higher than base year per capita income. In the graph, total income for group \( m \) is \( Y_2 \) and consumption is \( C_2 \). In this case the linear programming variable \( U_m \) will be at a positive level in the optimum solution, and,

\[
U_m = Y_2 - Y_o \\
C_2 = C_o + e_{im} d_{im} U_m
\]

Similarly, if per capita income in the optimum solution is less than per
capita income in the base year, the linear programming variable $V_m$ will be at a positive level in the optimum solution, and in terms of the graph

$$V_m = Y_o - Y_1$$

$$C_1 = C_0 - e_{im} d_{im} V_m$$

Note that two variables, $U_m$ and $V_m$, are needed if we allow the possibility of either an increase or a decrease in per capita income levels from that of the base year. If only $U_m$ were used, per capita income would be constrained to being equal to or greater than the base year level.
REFERENCES


5. Instituto de Investigaciones Tecnológicas, Diseno de Dietas de Costo Minimo, Bogota, Colombia, September, 1972.


125