CHAPTER IV

ESTIMATED RELATIONS

REVISED NOTATION

In this chapter the economic model of Chapter II is assumed to have a particular parametric form, and estimates of the parameters are obtained by alternative statistical procedures. To facilitate the application of certain theoretical propositions to the present model, a notation similar to one that has been used in various theoretical discussions is adopted. The value of a current endogenous variable for a particular time period is denoted by a y with two subscripts. The first subscript identifies the particular variable, and the second subscript refers to the time period. Similarly a z with two subscripts represents the value of a predetermined variable for a given time period. The new symbols are defined below. The symbol used to represent a particular variable in the notation of (2.19)-(2.26) is listed inside the parenthesis following each new symbol.

$y_{1t}(l_t)$ = quantity of livestock and livestock products produced

$y_{2t}(d_t)$ = price of feed grains

$y_{3t}(g_t)$ = price of protein feed

$y_{4t}(l_t^*)$ = quantity of livestock and livestock products sold

$y_{5t}(p_t)$ = price of livestock and livestock products

$y_{6t}(a_t)$ = amount of feed grains fed to livestock

$y_{7t}(b_t)$ = amount of protein feed fed to livestock

$y_{8t}(f_t)$ = total feed fed to livestock

$z_{1t}(h_t)$ = beginning inventory of livestock on farms

$z_{2t}(t)$ = time

$z_{3t}(c_t)$ = quantity of roughage fed to livestock

$z_{4t}(l_{t-1}^*)$ = quantity of livestock and livestock products sold in previous year

$z_{5t}(y_t)$ = disposable personal income

$z_{6t}(n_t)$ = population

$z_{7t}(r_r)$ = general price index (excluding livestock and livestock products)

$z_{8t}(k_t)$ = price of farm labor

$z_{9t}(a_t^*)$ = available quantity of feed grains

$z_{10t}(b_t^*)$ = production of protein feeds

Capital letters are used to denote logarithms of the above variables; i.e., $Y_{it} = \log y_{it}$, $Z_{it} = \log z_{it}$. We now consider a statistical model in which the first seven relations of the economic model of Chapter II are represented by equations that are linear in the logarithms of the observed variables. Though there are certain conveniences in using equations of this form (computational simplicity, direct interpretation of coefficients as elasticities) and such usage is usually consistent with our qualitative presumptions about the form of our relations, it should still be recognized that their use involves a highly special and incompletely justified assumption. Statistical results utilizing other algebraic forms are reported briefly in later chapters. The results given in this chapter are based on the model given by (4.1) to (4.7) below, which is expressed in terms of the logarithms of observed variables, except for z_{2t} which represents time.

A Statistical Model

If log z_{2t} had been inserted in the production relation (4.1 below), both the interpretation of the relation and the values obtained for estimates of the parameter of the variable would depend on an arbitrary choice of the origin from which time is measured. Inserting z_{2t} along with logarithms of the other variables implies a constant rate of increase in the quantity of livestock products obtained from given feed and herds. If the secular increase in feeding efficiency is to be represented by a variable representing time, this seems as reasonable as any alternative assumption about the way in which it enters.[1]

(4.1) $$Y_{1t} + \beta_{16}Y_{6t} + \beta_{17}Y_{7t} + \gamma_{11}Z_{1t} + \gamma_{12}z_{2t} + \gamma_{13}Z_{3t} + \gamma_{10} = U_{1t}$$

(production relation)

[1] This agrees with Lorie's treatment of time in the production relation. See James H. Lorie, *op. cit.*, pp. 83–96.

$$(4.2)\quad Y_{2t} + \beta_{23}Y_{3t} + \beta_{25}Y_{5t} + \beta_{26}Y_{6t} + \gamma_{21}Z_{1t} + \gamma_{23}Z_{3t} + \gamma_{20} = U_{2t}$$

(demand for feed grain)

$$(4.3)\quad \beta_{32}Y_{2t} + Y_{3t} + \beta_{35}Y_{5t} + \beta_{37}Y_{7t} + \gamma_{31}Z_{1t} + \gamma_{33}Z_{3t} + \gamma_{30} = U$$

(demand for protein feed)

$$(4.4)\quad \beta_{41}Y_{1t} + \beta_{42}Y_{2t} + \beta_{43}Y_{3t} + Y_{4t} + \beta_{45}Y_{5t} + \gamma_{41}Z_{1t} + \gamma_{48}Z_{8t} + \gamma_{40} = U_{4t}$$

(supply of livestock products)

$$(4.5)\quad \beta_{45}Y_{4t} + Y_{5t} + \gamma_{54}Z_{4t} + \gamma_{55}Z_{5t} + \gamma_{56}Z_{6t} + {}_{57}Z_{7t} + \gamma_{50} = U_{5t}$$

(demand for livestock products)

$$(4.6)\quad \beta_{62}Y_{2t} + Y_{6t} + \gamma_{69}Z_{9t} + \theta' s_t' = U_{6t}$$

(supply of feed grains)

$$(4.7)\quad \beta_{73}Y_{3t} + Y_{7t} + \gamma_{7,10}Z_{10t} + \theta'' s_t'' = U_{7t}$$

(supply of protein feeds)

The symbols β_{ij}, γ_{ij} in the above model represent unknown constant coefficients to be estimated. The U_{it} represent random disturbances. θ', θ'' stand for coefficients of the unspecified variables s_t', s_t''. The fact that these variables are not known or utilized means that our model is incomplete and makes some of the estimates described below less efficient than the estimates that would be obtained if these variables were specified and observed. Each of the symbols s_t', s_t'' should be thought of as representing an unknown number of important omitted variables with θ', θ'' representing corresponding numbers of coefficients. If none of the unknown variables are endogenous, the model could be completed without additional equations. However, if some of the unknown variables are endogenous, additional relations would have to be specified to complete the model.[2]

The accounting relation given by (2.26) has been omitted since it is of secondary interest. If the disturbance in this relation is assumed to be independent of the other disturbances, then, for statistical purposes, the accounting relation may be regarded as a separate one-equation model.[3] An estimate of the parameter of this equation will be given in

[2] For a discussion on the concept of a complete model, see T. C. Koopmans, When Is an Equation Complete for Statistical Purposes? Cowles Commission Monograph 10, *op. cit.*, and T. C. Koopmans and William C. Hood The Estimation of Simultaneous Linear Economic Relationships, Cowles Commission Monograph 14, *op. cit.*, pp. 113–127.

[3] See pp. 13, 14.

Chapter VI. The unobserved random disturbances U_{it} are assumed to have a multivariate normal[4] distribution with zero means and a finite covariance matrix. For a given time period, the values taken by the disturbances may be correlated with each other, but they are assumed to be independent of the values of disturbances for other time periods. Their distribution is assumed to remain constant over the observation period. The meaning of any equation is not changed if each term is multiplied by a given constant. This makes it possible to choose arbitrarily a convenient value for one coefficient in each equation, provided a coefficient known to be nonzero is chosen. In the above equations, the coefficients β_{ii}, $i = 1, 2, \cdots, 7$, have been set equal to unity.

Identification of Relations

Estimates of the other parameters in (4.1) to (4.5) are desired. A question that logically arises prior to estimation of parameters is that of identification, because, in general, different sets of structural equations can lead to the same probability distribution of the current endogenous variables for given values of the predetermined variables. Our observations generally enable us to draw statistical inferences about the conditional distribution of the endogenous variables, but our ability to draw inferences about the structural relations depends on establishing logical connections between the structural equations and the conditional distribution of the endogenous variables.

An equation is said to be identified (or identifiable) if its parameters (coefficients of observed variables and parameters of the distribution of the disturbance) could be uniquely determined from the specification of the model and the conditional distribution of the endogenous variables. Unless a relation is identified, we cannot expect to obtain point estimates of its parameters, though we could typically estimate certain functions of the parameters. Koopmans and others have derived conditions for the identification of equations in models like the one presented above.[5] It

[4] The estimates to be presented remain consistent if the normality assumption is relaxed. See H. Chernoff and H. Rubin, Asymptotic Properties of Limited-Information Estimates under Relaxed Conditions, Cowles Commission Monograph 14, *op. cit.*, pp. 200–212.

[5] T. C. Koopmans, Identification Problems in Economic Model Construction, *Econometrica*, Vol. 17, 1949. Also Olav Reiersol and T. C. Koopmans, The Identification of Structural Characteristics, *Annals of Mathematical Statistics*, Vol. 21, 1950. For the necessary and sufficient conditions for identification in linear models see the following: T. C. Koopmans and William C. Hood, *op. cit.*, pp. 135–139, and also pp. 185–186; T. C. Koopmans, H. Rubin and R. B. Leipnik, Measuring the Equation Systems of Dynamic Economics in Cowles Commission Monograph 10, *op. cit.*, pp. 69–85. Interesting earlier discussions of identification and related problems in somewhat different language may be found in Henry Schultz, *Theory*

has been shown that, if there are H current endogenous variables that may enter a given equation, then a necessary condition for the equation to be identified is that there be at least $H - 1$ observable predetermined variables in the system that are known not to enter the equation being considered. If the number of excluded predetermined variables is exactly $H - 1$, the equation is said to be just identified; if the number of excluded predetermined variables exceeds $H - 1$, the equation is said to be overidentified. With respect to equation 4.1, for example, $H = 3$, and the number of excluded predetermined variables is 7, so that this equation is overidentified. (4.2) to (4.5) can also be seen to be overidentified.

Estimation of Parameters

Methods that have thus far been developed for obtaining estimates of parameters of simultaneous equations systems are based on the principle of maximum likelihood. If our model were complete, it would be possible to form the likelihood function for our sample of observations and to maximize this function with respect to all the unknown parameters. Estimates obtained in this way would be called full-information maximum-likelihood estimates and would have the usual maximum-likelihood properties.[6] Even if our model were complete, we would find this computation process exceedingly expensive for a model that contains overidentified equations. Partly to provide less expensive (though less efficient) procedures and partly to anticipate circumstances in which the investigator might not feel justified in making all the specifications necessary for the application of full-information maximum likelihood, an alternative procedure called the limited-information method has been developed.[7]

---

and Measurement of Demand, University of Chicago Press, Chicago, 1938, especially Ch. II, Sec. IIIc; A. C. Pigou, A Method of Determining the Numerical Values of Elasticities of Demand, *Economic Journal*, Vol. 20, 1910, pp. 636–640; R. Frisch, Pitfalls in the Statistical Construction of Demand and Supply Curves, *Veroffentlich-ungen der Frankfurter Gesellschaft fur Konjunkturforschung*, Neue Folge, Heft 5, Leipzig, 1933; E. J. Working, What Do Statistical Demand Curves Show?, *Quarterly Journal of Economics*, Feb. 1929, pp. 212–235.

[6] The full-information maximum-likelihood procedure is given in T. C. Koopmans, H. Rubin, and R. B. Leipnik, *op. cit.*, pp. 110–183. See also T. C. Koopmans and William C. Hood, *op. cit.*, pp. 143–162.

[7] The limited-information procedure is developed in T. W. Anderson and H. Rubin, Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations, *Annals of Mathematical Statistics*, Vol. 20, 1950, and The Asymptotic Properties of Estimates of the Parameters of a Single Equation in a Complete System of Stochastic Equations, *Annals of Mathematical Statistics*, Vol. 21, 1950. See also T. C. Koopmans and William C. Hood, *op. cit.*, pp. 162–185, and T. C. Koopmans, H. Rubin, and R. B. Leipnik, *op. cit.*, pp. 110–153. The method was referred to earlier as the reduced-form method. See M. A. Girshick and Trygve Haavelmo, *op. cit.*, pp. 93–95.

This method yields estimates for a selected subset of the equations of a system, neglecting information about which variables appear in particular equations in the remainder of the system. In particular, it can be used to estimate parameters of one equation at a time and can be applied when the equations not being estimated are incompletely specified. The estimates so obtained can be shown to be consistent,[8] but, unfortunately, little is known about their small-sample properties. Since economists often have only small samples to work with, this is a serious limitation.

However, pending further developments in the relevant statistical theory, it seems useful to try this procedure and to make such judgments as are possible of the results. When the limited-information method is applied to a single equation, the variables are, in effect, divided into four classes—current endogenous variables that enter the equation (sometimes denoted by y_*), current endogenous variables excluded from the equation (y_{**}), predetermined variables that enter the equation (z_*), and predetermined variables excluded from the equation (z_{**}). Except for the normalizing assumption (arbitrarily giving one of the nonzero coefficients a chosen value), the variables within a class are treated symmetrically in the calculation of estimates. The y_{**} do not enter the computations. Estimation of coefficients of the y_* may be interpreted as estimating a linear combination of the y_*; for example, in estimating coefficients of the y_* in (4.1), we are estimating the linear combination, $Y_{1t} + \beta_{16}Y_{6t} + \beta_{17}Y_{7t}$.

Let the linear combination of the y_* be represented by y^0. Consider the linear regression of y^0 on the z_* and the z_{**}. The limited-information estimates of the coefficients in y^0 are those values that minimize the relative contribution of the z_{**} to the explanation of the variance of y^0. Since this contribution would be zero in the population regression of the true y^0 on the z_* and z_{**}, the estimation procedure does not seem unreasonable under this interpretation. Estimates of the coefficients of the z_* (γ_{11}, γ_{12}, γ_{13}, γ_{10} in 4.1) are the least-squares estimates of coefficients of the regression of the estimated y^0 on the z_*.[9]

If some of the z_{**} are omitted in the calculation, the procedure still gives consistent (though less efficient) estimates of coefficients, provided at least $H - 1$ of the z_{**} are utilized. This fact, together with the fact that the y_{**} do not enter the calculations at all, helps explain why the

[8] An estimate is consistent if it approaches, in a probability sense, the true parameter value, as the number of observations increases. See A. M. Mood, *Introduction to the Theory of Statistics*, p. 149, McGraw-Hill Book Co., New York, 1950. Also S. S. Wilks, *Mathematical Statistics*, Princeton University Press, Princeton, 1947, p. 133.

[9] For a more complete account of this interpretation, see T. C. Koopmans and William C. Hood, *op. cit.*, pp. 166–177.

method can be applied to incomplete systems. In an application to an incomplete system, certain y_{**} and z_{**} are ignored. Omitting the former does not affect the computations; omitting the latter sacrifices some efficiency but does not destroy consistency of the estimates. In view of this, one may sometimes consider using only part of the available z_{**} in order to simplify computations. In such a case it seems reasonable to use those z_{**} that might be expected to minimize the variances of residuals in the regressions of the y_* on the z_* and z_{**}.[10]

Parameters of Production Relation

Limited-information estimates for the coefficients in equation 4.1 are given in the first row of Table VI. These were computed, using Z_{4t}, Z_{5t}, Z_{7t}, Z_{8t}, Z_{9t}, Z_{10t} as the z_{**}. Z_{6t}, population, was omitted because

Table VI
RESULTS FOR PRODUCTION RELATION

Method	Estimates of Coefficients					
	β_{16}	β_{17}	γ_{11}	γ_{12}	$\gamma_{1}\ ^{2}$	γ_{10}
Limited information	−0.222(0.078)	−0.349(0.093)	−1362(0.119)	−0.0009(0.001)	+0.221(0.122)	−3.911(0.931)
Least squares	−0.350(0.061)	−0.222(0.075)	−0.252(0.103)	−0.0023(0.0009)	+0.229(0.110)	−4.885(0.927)

$R^2 = 0.982$

it was believed to be too closely related to z_{2t} and other predetermined variables having a significant trend to contribute much to the accuracy of the estimates. The second row of Table VI contains least-squares estimates of the coefficients in (4.1). The computations relating to this equation are given in detail in the appendix. The numbers in parentheses beside the estimates are the calculated standard errors of the estimates. R^2 is the coefficient of determination for the least-squares computation.

To obtain least-squares estimates of parameters of a given equation, it is necessary to designate one of the observed variables as dependent in that equation. The other observed variables are called independent. Least-squares procedures are derived under the assumption that the random disturbance in the equation is distributed independently of the independent variables. If an equation contains one current endogenous variable and several predetermined variables and if the current endoge-

[10] For a fuller account of the properties of limited-information estimates, see T. W. Anderson and H. Rubin, The Asymptotic Properties of Estimates of the Parameters of a Single Equation in a Complete System of Stochastic Equations, *op. cit.* See also H. Chernoff and H. Rubin, *op. cit.*

nous is regarded as dependent in applying least squares, then least-squares and limited-information procedures are identical, as are the assumptions under which they are derived. If an equation contains several current endogenous variables, then all but one of these have to be regarded as independent in the application of least squares.

This involves an assumption in addition to those underlying the limited-information calculations. To the extent that this assumption is not justified, the least-squares estimates of coefficients and their calculated standard errors contain additional biases. However, this could be outweighed by the greater efficiency of least-squares estimates if the least-squares assumptions were approximately true. As was indicated in Chapter II, it has often been contended that, for many relations involving agricultural commodities, least-squares assumptions are sufficiently realistic. For this reason and because little is known about the small-sample properties of limited-information estimates, it has seemed desirable in this study to estimate coefficients using both procedures.

Intuitive considerations, which are in a general way supported by Bronfenbrenner's results,[11] suggest that least-squares bias might be minimized by treating as independent those current endogenous variables that are most strongly influenced by predetermined variables not appearing in the equation being estimated (i.e. the z_{**}). In equation 4.1, we would expect Y_6 and Y_7 to depend very strongly on Z_9 and Z_{10}, respectively, so Y_1 was treated as dependent in obtaining the least-squares estimates given in Table VI. This is also consistent with the reasoning underlying our first-approximation model of Chapter II, in which feed fed to livestock was regarded as predetermined.

Results for Farm Decision Relations

The choice of dependent variables for equations 4.2, 4.3, 4.4 is somewhat more arbitrary than for equation 4.1. The least-squares results given in Tables VII, VIII, IX were obtained, using Y_2 as dependent in (4.2), Y_3 in (4.3), and Y_4 in (4.4). In computing limited-information estimates, the z_{**} used for (4.2) and (4.3) were Z_4, Z_5, Z_7, Z_8, Z_9, Z_{10}. The z_{**} used for (4.4) were the same, except that Z_3 was included and Z_8 excluded.

The choice of a dependent variable in (4.5) is fairly clear from the fact that we expect Y_4 to be strongly influenced by Z_1, Z_3, Z_9, Z_{10}

[11] Jean Bronfenbrenner, Asymptotic Bias in Least-Squares Estimates of the Parameters of a Single Linear Stochastic Equation in a Complete System, unpublished Ph.D. thesis, University of Chicago, 1950. See also Jean Bronfenbrenner, Source and Size of Least-Squares Bias in a Two-Equation Model, Cowles Commission Monograph 14, *op. cit.*

TABLE VII

DEMAND FOR FEED GRAINS

Method	Estimates of Coefficients					
	β_{23}	β_{25}	β_{26}	γ_{21}	γ_{23}	γ_{20}
Limited information	0.259(0.229)	−1.582(0.215)	1.472(0.295)	−2.760(0.572)	0.708(0.620)	7.369(7.359)
Least squares	−0.040(0.191)	−1.295(0.179)	1.258(0.261)	−2.430(0.528)	0.441(0.567)	8.619(3.541)

$R^2 = 0.948$

TABLE VIII

DEMAND FOR PROTEIN FEEDS

Method	Estimates of Coefficients					
	β_{32}	β_{35}	β_{37}	γ_{31}	γ_{33}	γ_{30}
Limited information	0.049(0.108)	−0.982(0.146)	0.543(0.067)	−0.644(0.313)	0.278(0.356)	−1.584(2.380)
Least squares	−0.122(0.073)	−0.739(0.101)	0.533(0.059)	−0.342(0.262)	−0.083(0.298)	−1.213(1.604)

$R^2 = 0.965$

TABLE IX

SUPPLY OF LIVESTOCK PRODUCTS

Method	Estimates of Coefficients						
	β_{41}	β_{42}	β_{43}	β_{45}	γ_{41}	γ_{48}	γ_{40}
Limited information	−0.908 (0.143)	−0.149 (0.032)	0.061 (0.111)	0.191 (0.121)	−0.0013 (0.136)	−0.099 (0.945)	−0.661 (2.839)
Least squares	−0.801 (0.066)	−0.135 (0.929)	0.133 (0.050)	0.136 (0.052)	−0.083 (0.074)	−0.120 (0.035)	−0.993 (0.568)

$R^2 = 0.992$

and that Y_4 was treated as predetermined in our first model. Y_5 is therefore treated as dependent in the application of least squares to (4.5).

DEFLATION OF VARIABLES IN DEMAND RELATION

Since the demand relation expresses the aggregate result of the behavior of a larger number of individual consumers, it seems reasonable to express the demand relation in per-capita terms. This involves deflating the original quantity variables (y_4, z_4, z_5) by population. If we retain the assumption that the relation is linear in the logarithms of the observed variables, then our per-capita demand relation is given by

$$(4.5') \quad \beta'_{54}Y'_{4t} + Y_{5t} + \gamma'_{54}Z'_{4t} + \gamma'_{55}Z'_{5t} + \gamma'_{57}Z_{7t} + \gamma'_{50} = U'_{5t}$$

TABLE X

DEMAND FOR LIVESTOCK PRODUCTS

Method	Estimates of Coefficients				
	β'_{54}	γ'_{54}	γ'_{55}	γ'_{57}	γ'_{50}
Limited information	1.319(0.348)	0.041(0.344)	−0.986(0.100)	−0.492(0.110)	−6.744(4.022)
Least squares	1.307(0.295)	0.045(0.306)	−0.984(0.089)	−0.494(0.098)	−6.690(2.469)

$R^2 = 0.981$

where

$$(4.8)\qquad Y'_{4t} = Y_{4t} - Z_{6t} = \log y_{4t}/z_{6t}$$

$$(4.9)\qquad Z'_{4t} = Z_{4t} - Z_{6,t-1} = \log z_{4t}/z_{6,t-1} = Y'_{4,t-1}$$

$$(4.10)\qquad Z'_{5t} = Z_{5t} - Z_{6t} = \log z_{5t}/z_{6t}$$

The least-squares results reported in Table X were obtained by using Y_5 as the dependent variable and the other observed variables in (4.5′) as independent.

To obtain limited-information estimates of (4.5′), we visualize a model in which Y'_{4t}, Z'_{4t}, and Z'_{5t} have been substituted for Y_{4t}, Z_{4t}, and Z_{5t}, respectively, in the system given by (4.1) to (4.7). The effect of this is to replace (4.5) by (4.5′) and to replace (4.4) by

$$(4.4')\qquad \beta_{41}Y_{1t} + \beta_{42}Y_{2t} + \beta_{43}Y_{3t} + Y'_{4t} + \beta_{45}Y_{5t}\gamma_{41}Z_{1t} + Z_{6t} + \gamma_{48}Z_{8t} + \gamma_{40} = U_{4t}$$

The significance of (4.4′) is that the substitution of Y'_4 for Y_4 has introduced Z_6 into the equation. In the estimation of (4.5′), therefore, Z_6 may be used as a z_{**}. The z_{**} used in obtaining the limited-information estimates given below were Z_1, Z_3, Z_6, Z_8, Z_9, Z_{10}.

In estimating coefficients of (4.5′), observations relating to the years 1943 through 1946 were omitted because price control and rationing of livestock products were effective during this period.[12]

In all of the above tables, the calculated standard errors of the limited-

[12] There is a question of whether these observations should also have been omitted in the limited-information estimation of other relations. If we assume that the livestock demand relation is the only one appreciably affected and that the change in this relation can be adequately represented by the intervention of new exogenous factors, then the variables representing these exogenous factors would, if observed, enter as z_{**} in the estimation of (4.1) to (4.4). However, if they are unobserved and omitted from the calculations, the resulting estimates of (4.1) to (4.4) remain consistent.

information estimates are based on asymptotic formulas[13] and may be expected to be biased for small samples. In the absence of a small-sample theory, they are perhaps as good indicators as we have of the order of accuracy of the limited-information estimates. The choice of units of measurement for the variables affects only the constant term. In the calculation of the estimates presented here, y_1, y_4, z_1, z_4, z_5 are in dollars; y_6, y_7, z_3, z_9, z_{10} are in thousands of pounds; the unit for y_2 and y_3 is dollars per thousand pounds. y_5 is an index whose average level, 1920–49, is approximately 1; z_7 is an index whose average value, 1935–39, is 100. z_2 is in years, z_8 in cents per hour, and z_6 in millions of persons.

Some Aspects of the Initial Results

For the most part, differences between estimates obtained by the two methods are not striking when compared with the indicated magnitudes of sampling fluctuations. Except for equation 4.5, however, they are large enough to have important practical consequences if their reliability were firmly established. An interesting example concerns the differences in signs of the estimated coefficients of β_{23} in (4.2) and β_{32} in (4.3). The positive coefficients obtained by limited information imply that feed grains and protein feeds are technical complements, whereas the negative coefficients obtained by least squares imply that they are technical substitutes. Our a priori knowledge in this case is probably insufficient for us to regard either outcome as implausible. For the least-squares treatment, the outcomes of ordinary t tests of the hypotheses that particular coefficients are equal to zero can be told fairly well by inspection. Tests of coefficients in (4.1), (4.2), and (4.3) involve the distribution of t with 24 degrees of freedom; for coefficients in (4.4) and (4.5) t has 23 and 21 degrees of freedom, respectively. The value of t in a given case is simply the ratio of the estimated coefficient (in the least-squares row) to the estimated standard error. For 21 or more degrees of freedom, the distribution of t is almost normal, the 5% point being 1.72 for t with 21 degrees of freedom and 1.64 for the normal distribution.[14] Analogous tests based on limited-information estimates are not available, but the ratios

[13] Limited-information estimates of coefficients have been shown to have an asymptotic normal distribution, whose covariance matrix may be obtained from the expected values of the second partial derivatives of the logarithm of the likelihood function. The variances of the estimates are functions of the true values of the coefficients of an equation. The calculated standard errors given in the tables are obtained by replacing the true coefficients with their limited-information estimates. See T. W. Anderson and H. Rubin, Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations, *op. cit.*, pp. 53–56. Also T. C. Koopmans, H. Rubin, and R. B. Leipnik, *op. cit.*, pp. 133–153, and T. C. Koopmans and William C. Hood, *op. cit.*, pp. 177–178.

of estimates to calculated standard errors should still give a rough idea of the reliability attached to the signs of the estimates.

The failure of one of the tests indicated above to reject significantly the null hypothesis could arise because the variable whose coefficient is tested is not important in the relation considered, or for a variety of other reasons. Variations in the observed values of the variable in our particular sample might be small or might be closely related to variations in some of the other variables. Use of an inappropriate algebraic form for the relation or other defects in the statistical specification could also bias the estimates and the associated tests. Therefore, we do not necessarily drop a variable from the analysis if our estimate of its coefficient does not appear significant. The relation between the estimates and their calculated standard errors is one type of evidence taken into account in considering possible revisions of the statistical model.

Another type of evidence taken into consideration is the comparison of the estimates with what we would regard as plausible from our a priori knowledge of the underlying relations of our model. The estimates of γ_{13} in Table VI appear quite implausible in that a negative coefficient for Z_{13} implies that increasing the amount of roughage fed to livestock, while holding other factors constant, would decrease the production of livestock products. The negative sign for γ_{33} in Table VIII obtained by applying least squares to equation 4.3 would also lead to a rather implausible interpretation: namely, that increased quantities of roughage fed to given herds of animals would lead producers to increase the quantities of protein feeds fed. It is worth noting, in this connection, that errors of observation in z_3 may be unusually large, since the dominant component is pasture consumption, and this is particularly difficult to measure.[15] Other results that seem inconsistent with our a priori information are the positive estimates of β_{43} and the small absolute values of the estimates for γ_{41} in Table IX, and the positive estimates of γ'_{54} in Table X. These will be considered again in later chapters along with some results using alternative statistical specifications.

---

[14] In considering these tests we are mainly interested in whether or not the data are consistent with the hypothesis that the true value of the coefficient is zero or of opposite sign to that of the estimate. Consequently the 5% points indicated are for one-tailed tests. See A. M. Mood, *op. cit.*, p. 425.

[15] For the procedure adopted in obtaining a measure of pasture consumption by livestock see pp. 47 to 49 in Chapter III.