

143

April 28, 1972

To the Graduate Council:

I am submitting herewith a dissertation written by Steven Darrell Mundy entitled "Technology and the Farm Labor Market." I recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Agricultural Economics.

Laury L Bauer
Major Professor

We have read this dissertation
and recommend its acceptance:

Luther H. Keller

R. J. Whaley
Karl E. Phillips

Accepted for the Council:

Hilton A. Smith
Vice Chancellor for
Graduate Studies and Research

TECHNOLOGY AND THE FARM LABOR MARKET

A Dissertation
Presented to
the Graduate Council of
The University of Tennessee

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

by
Steven Darrell Mundy

June 1972

ACKNOWLEDGMENTS

The author expresses his gratitude to Dr. L. L. Bauer, Chairman of his Advisory Committee, for his advice and direction in the development and completion of this study. Appreciation is also extended to the other members of his committee, Dr. L. H. Keller, Dr. T. J. Whatley, and Dr. K. E. Phillips. The author acknowledges the willing assistance of Dr. J. G. Snell, Dr. C. B. Sappington, Dr. R. Daniel, and Dr. F. O. Leuthold.

The author is indebted to Judy Mellor for her assistance with the computer programming. The helpfulness of Anna Norwood, Melitta Stout, and Linda Howard in the typing of the rough draft is most appreciated. Gratitude is extended to Ann Lacava who typed the final draft.

A special thanks is extended to the author's parents, Mr. and Mrs. Wade Mundy, and to other members of his family for their aid, love, and understanding during the many years the author has been involved in his graduate studies.

Above all the author expresses his deepest love and appreciation to his wife Sandy for her love, sacrifice, perseverance, and encouragement.

ABSTRACT

This study analyzed the effects of technology on the farm labor market. The basic hypothesis was that agricultural technological advancement has been and still is an important determinant of low returns to farm labor, unemployment, and underemployment. Furthermore, technology can act as (1) an initial stimulus in generating these conditions in the farm labor market and (2) as a major contributor to the persistence of the problems along with other conditions affecting the farm labor market.

Two objectives were established in order to test the hypothesis. The objectives were (1) to estimate the effect of technology on the farm labor market and (2) to estimate the distributed lag function or time path of this effect.

The six-equation conceptual model, consisting of the demand and supply of farm labor, the demand and supply of aggregate farm product, and the equilibrium identities for the two markets, was specified. The major distributed lag techniques used in dynamic analyses of economic phenomena were examined and evaluated. The general polynomial distributed lag technique was selected based upon this evaluation for use in this study.

The statistical model was specified subject to the limitations imposed by the data, the distributed lag technique, and the method of estimation. A time series of cross sections of state data was used. Technology was represented by the proxy variable, research and extension expenditures. Deflated data on research and extension expenditures for

the years 1889 through 1964 were used. Data for all of the other variables in the model included the years 1949 through 1964.

Behavioral theory on distributed lags was limited. Hence, the configuration of the optimum lag distribution showing the distributed lag effect of technology on the demand for farm labor was unknown. As a result, several alternative structures were estimated. They included alternative lag polynomials from zero degree through fifth degree at various lag intervals from 0 through 60 lagged years on technology. Four choice criteria were used in selecting the optimum distribution from these estimated alternatives. They were (1) judgment, (2) the signs of the calculated weights, (3) the close similarity of the lag weights among distributed lags of optimal and longer length, and (4) reasonableness of the magnitude of the weights.

The statistical analysis was performed using a least-squares regression technique. The simultaneous system was overidentified so the two-stage least-squares estimation method was employed.

No optimal lag function was selected because all estimated functions failed to satisfy one or more of the selection criteria. However, one alternative structure was selected as the least unreasonable for the purpose of demonstration. The results of this estimated structure indicated that technology had a positive influence on the demand for farm labor. This result did not agree with the prior expectations hypothesized in the conceptual model. Furthermore, the estimated coefficient on technology was less than its asymptotic standard error.

None of the stated objectives were attained. Consequently, no substantive conclusions regarding the distributed lag effects of

technology on the farm labor input could be drawn. However, some conclusions were drawn regarding the analytical approach taken. The major conclusion was that this specific application of general polynomial, finite distributed lags in a simultaneous equation framework which did not constrain or restrict the signs and magnitudes of the lag weights, permitted too much freedom in the estimated structures by allowing the data to determine the optimum distribution. A logical redirection would be to convert from regression analysis to quadratic programming and to incorporate restrictions on the lag weights. Other redirections were also suggested.

TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION	1
Some Controversies Surrounding Technology and	
Farm Labor	5
Statement of the Problem	9
Justification	11
Some Previous Economic Research on Technology and the	
Farm Labor Market	13
Purpose and Objectives of the Study	19
II. DERIVATION AND DESCRIPTION OF THE ECONOMIC MODEL	21
The Model	21
Some General Underlying Assumptions Regarding Model	
Specification	24
Some Preconceptions Concerning Causes and Effects	
in the System	30
III. CONCEPTS AND ECONOMETRIC ASPECTS OF DISTRIBUTED LAGS . .	42
The Concept of Distributed Lags	42
Methods of Restricting the Parameter Space of the	
General Distributed Lag Function	45
Major Choice Criteria for Using the General Polynomial	
Form of Distributed Lags	63
IV. SOME STATISTICAL AND ECONOMIC CONSIDERATIONS IN	
FORMULATING THE STATISTICAL MODEL	67

CHAPTER	PAGE
The Level of Aggregation of Commodities	67
Choice of Spatial Unit	68
Choice of the Time Unit and Years Included	69
Functional Form	71
Measuring Technology	72
Some Difficulties with Indices and Deflators in the Time Series of Cross Sections	77
Choice of Variables in the Statistical Model	78
V. THE STATISTICAL MODEL AND PROCEDURES FOR ESTIMATION . . .	80
The Structural Model	80
The Reduced Form	87
Estimation of the Reduced-Form Parameters	89
Estimation of the Structural Parameters	91
Derivation of Estimates of the Weights of the Lag Distribution	92
The Average Lag	96
Selection of the Degree of the General Polynomial and Points of Truncation	99
VI. RESULTS	103
The Estimated Lag Distributions	103
An Example of the Estimated Structure	108
Some Final Observation on the Results	118
VII. SUMMARY, CONCLUSIONS, AND SUGGESTIONS FOR FUTURE RESEARCH	121
Summary	121

CHAPTER	PAGE
Conclusions	123
Suggestions for Future Research	125
REFERENCES	129
APPENDICES	134
Appendix A	135
Appendix B	153
Appendix C	154
VITA	155

CHAPTER I

INTRODUCTION

Farming has undergone an almost complete technological metamorphosis since the turn of the century because of the invention, development, and adoption of new technologies. The utilization of energy sources which largely replaced human and animal sources for doing physical work, the use of hybrid seed and chemical fertilizers, and others have enabled farmers to produce a greater quantity of product with increased productivity of inputs. The total output of the economy has increased over time as a result of these new technologies and increased use of inputs in total.

The new technologies, along with social, political, and economic conditions on both the demand and supply sides of the markets for inputs and products, have had and are continuing to have a substantial effect on the nature of the input mix and the quantities employed of the various inputs in farming. For example, the use of mechanical and power machinery in the United States in 1969 represented a 400 percent increase over its utilization in 1934; fertilizer and lime, 1600 percent; nonfarm feed, seed, and livestock purchases, over 600 percent; miscellaneous inputs, over 190 percent; and farm real estate, over 120 percent (U.S. Department of Agriculture, 1949-1970, specifically 1967, Table 665 and 1970, Table 652). The index of farm labor, on the other hand, was the only input index to decline, except for some annual fluctuations, over the same period. The utilization of farm labor in 1969 had declined to

about one-third of the level in 1934 (U.S. Department of Agriculture, 1949-1970, specifically 1967, Table 665 and 1970, Table 652).

These statistics indicate some of the changes that have been taking place in the mix and level of utilization of farm inputs. These changes, their causes, and their effects are important with respect to all farm inputs. However, the changes and their effects on the labor input are especially relevant because the labor input is a human resource. The human characteristic of the labor input forces one to consider the welfare aspects of changing conditions in the farm economy. Increasing the productivity of agricultural resources and transferring these resources from one use to another, including nonagricultural uses, are still relevant economic instruments for the entrepreneur, the laborer, the economist, and the policymaker. But because the human element is a major force, the costs of change are not just simply dollar costs, but involve noneconomic and nonquantifiable human costs as well. Agricultural economists and policymakers are more concerned than ever with the human effects of changing technological conditions among other forces of change in the product and inputs markets in agriculture.

Continually changing product and input market conditions in the agricultural sector lead to continually changing resource levels and mixes at the firm level and at the industry level. These changes lead to different effects on different groups. For example, changes that have taken place in modern United States agriculture have resulted in an abundant supply of economical food and fiber. The consumer has reaped the economic benefits of this abundance. Other groups such as farm laborers have not fared so well as evidenced by unemployment,

underemployment, and the persistence of low labor returns when compared with equivalent skills in nonfarm employment.

An earnings differential exists at the present between the farm and nonfarm working forces. A comparison of hourly wages of farm laborers and production workers in the United States showed that the 1969 farm wage rate of \$1.58 per hour (without room or board) (U.S. Department of Agriculture, 1949-1970, specifically 1970, Table 64) was about 50 percent of the 1969 wage rate of \$3.19 per hour for production workers in manufacturing industries (U.S. Bureau of the Census, 1949-1970, specifically 1970, Table 344). Another indicator of the labor earnings differential was the difference in median earnings of male civilians by occupation of longest job between year-round, full-time farm laborers and foremen and year-round, full-time laborers (except farm and mine) (U.S. Bureau of the Census, 1949-1970, specifically 1970, Table 347). The median earnings of farm laborers and foremen were about 50 percent of the median earnings of laborers (except farm and mine) in 1968.¹

In addition to the low relative returns to farm labor when compared to the returns to equivalent nonfarm labor, the returns to farm labor are low absolutely. For example, 1,466,000 workers (14 years of age and older) with 25 days or more of farm wage work during 1969 earned an average of \$1,886 at farm and nonfarm wage work for an average of 163 days worked. Farm wage work accounted for 131 days with an average

¹This statistic is not wholly representative as an indicator of the differential because females, children under 14, migratory laborers, and part-time laborers, were not included. Also, less than 25 percent of the farm labor force is hired labor. A more complete measure would include the returns to operator and family labor.

income from farm wage work of \$1,406² (U.S. Department of Agriculture, 1949-1970, specifically 1970, Table 639). Both low relative and absolute returns to farm labor indicate a need to develop the economic potential of the farm labor resource.

Other evidence indicates that technological advancement, in conjunction with other structural changes, have taken their toll among various groups in society by either direct or indirect means, even though the total benefits to society outweigh the costs to some individual groups. Problems facing farm laborers are just one example. Similar problems exist for other relatively unskilled laborers in non-farm occupations. Technological advancement generally brings with it the need for higher levels of worker skills and specialization. The relatively unskilled, uneducated groups of workers are the first to be outmoded and thus displaced, because these groups are usually least capable for various social, political, institutional, and economic reasons, of improving their relative positions in the work force. Consequently, many of the individuals in these groups are directly affected by being either underemployed or displaced and unemployed. Other groups of employed workers, who are more fortunate either because their skill levels are adequate and marketable, or they are more able to advance or to modify their skill levels to make them marketable in face of technological advancement, also are affected indirectly. Burdens are placed on all members of society in the form of unused manpower potential and through added public cost of various

²These statistics do not include nonwage work done by farm operators and unpaid family laborers.

welfare programs when government decides to assist these disadvantaged groups.

I. SOME CONTROVERSIES SURROUNDING TECHNOLOGY AND FARM LABOR

Agricultural economists have disagreed on the predominant causes of farm labor problems and their persistence, especially in recent years. For example, Cochrane (1958, Chapter 5) with his treadmill theory hypothesized that technology is the major cause of low returns to farm resources, emphasizing the vicious cycle created by the introduction of new technology and its adoption by farmers. He theorized that as new technology is introduced, some farmers adopt it initially to increase profits. However, as more farmers adopt the new technology, industry output is increased since individual firm supply curves shift to the right, causing optimum firm output to be greater for a given product price. But as total farm output expands, product price falls because of product supply shifting to the right (increasing) relatively more than the slower shifting, inelastic demand for product. Late adopters are forced to adopt the new technology in order to remain in business. Those that do not adopt are forced out. This cycle is repeated over and over so that farmers have to tread fast just to keep up.

The magnitude and direction of the effect of technology on the demand for labor under the above treadmill or other conditions would depend on the nature of the technological change. That is, as Bishop (1967, pp. 3-5) pointed out, the impact of technological change on manpower requirements in farming depends on the direction and relative

magnitudes of the resource substitution, firm size, organizational, and market effects generated by technological improvement. He emphasized that technological innovations in agriculture have not been neutral in their effects on the demand for resources because relative resource productivities have been altered. Changes in relative productivities provided incentives to alter the resource mix employed. Most technological innovations, particularly the early ones (since the mid 1930's), increased the productivity of capital relative to that of labor. This change in relative productivities reduced the amount of capital that was necessary to replace a unit of labor at each level of output. Many of these innovations also made it possible to reduce the amount of labor relative to land to produce a particular quantity of output. Consequently, changes in relative productivities caused by technological innovations have had a negative effect on the use of labor.

Bishop continued by stating that most technological innovations have tended to decrease the cost of producing additional output, which in turn increases the most profitable level of output for the firm. The size effects are positive because to increase size requires more manpower, other things remaining constant.

Changes in firm size generally lead to changes in organizational structure and generally to a greater degree of production specialization, either within the firm or through vertical or horizontal integration among firms. Organizational improvements tend to result in general in a substitution of other resources for manpower, so this effect is generally negative.

Finally, Bishop considered the market effect, which is somewhat indirect. He continued by stating (1967, p. 4),

. . . Because of the slow rate of growth in farm product demand, however, prices fall rapidly as aggregate output expands. Under these conditions the only way for the size effects to be fully realized at the firm level is through a reduction in the number of firms.

He summarized:

. . . In the aggregate, therefore, the substitution effects between capital and manpower exceed the size effects, with a consequent sharp reduction in the employment of manpower in farming. Thus farm people have been forced to choose between (1) massive transfers of manpower to nonfarm occupations in an effort to increase the return for the manpower remaining in farming or (2) occupational immobility, large-scale under-employment, and sharply reduced returns for manpower in farming.

Bishop then concluded (1967, p. 5),

. . . Even so, the adjustments (since the late 1930's) have not been completed, or are they likely to be completed in the near future. They are not once-and-over changes. As economic development occurs, capital is accumulated, and it becomes relatively low priced in comparison to manpower. Under these conditions research tends to be oriented in the direction of creating new technology which increases the rate of substitution of capital for manpower. . . . The decrease in the demand for manpower in farming, therefore, is a normal complement of economic growth and should be expected to continue.

The argument forwarded by Bishop that technology has had a dominant role in promoting a decline in manpower requirements was convincing, especially for the years prior to 1960. His argument seemed sound for conditions after 1960 except for conditions in the product market. For example, the United States aggregate productivity index of all inputs for farms (U.S. Department of Agriculture, 1970b, p. 12) increased only 4 percent from 1960 to 1970. At the same time the index of prices received by U.S. farmers (U.S. Department of Agriculture, 1970b, p. 7) increased in general from 1960 to

1970. The price increase implied that the aggregate demand for farm product increased at a greater rate than aggregate supply over the period. Consequently, the negative effect on the demand for labor caused by falling product price in the aggregate could not be used as a partial explanation of the decline in the demand for farm labor. However, the other effects were still pertinent. Output per man-hour increased from 189 to 323 (1950=100) during the same period (U.S. Department of Agriculture, 1970b, p. 12), while the quantity of labor utilized continued to decline and use of all other major categories of inputs increased (U.S. Department of Agriculture, 1970b, p. 11).

Tweeten (1969, pp. 798-817) placed less emphasis on output-increasing technology as the predominant cause of present problems of low returns to all farm resources, including labor. He stressed that economic growth, inflation, and/or technology can generate initial disequilibrium.³ Other causal factors such as the cost structure of farm firms, the degree of noncompetitiveness of the markets for inputs and farm products, and the existence of fixed resources, which are conditions that are related to technology, also can contribute to the persistence of the problem. Tweeten implied that these factors and conditions are themselves a function of technology among other things. Yet, he relegated technology to a minor role in explaining the persistence of low returns to farm resources. Realistically, the role of technology is a strong one in conjunction with these other conditions, because it, in part, determines them (and vice versa in

³Tweeten defined disequilibrium as resource returns below opportunity costs.

a complete system). For example, the cost structure of farm firms is highly interrelated to organizational changes and improvements brought about by technological change.

Schultz (1967, pp. 53-56) deemphasized the relative importance of structural maladjustments resulting from rapid gains in agricultural productivity as being a predominant cause of the farm labor problems. According to Schultz, farm labor problems are caused by (1) the slack in aggregate demand for all labor especially since 1957, and (2) the low level of marketable skills of the farm labor force. However, Schultz did not stress the importance of possible causal variables underlying his two predominant reasons for farm labor problems. That is, the aggregate demand for all labor is expected to be a function of technology in both the farm and nonfarm sectors along with other causal variables. Also, the level of marketable skills could be, among other things, a function of the level of technology, which was stressed in the earlier discussion. The possible presence of these functional relationships indicated that an important source of farm labor problems could be structural maladjustments resulting from technological change both within and outside the farm sector.

II. STATEMENT OF THE PROBLEM

A hypothesis of this study was that agricultural technological advancement has been and still is an important determinant of low returns to farm labor, underemployment, and unemployment. Furthermore, technology can act as (1) an initial stimulus in generating these conditions in the farm labor market, and (2) as a major contributor

to the persistence of the problems along with other conditions affecting the demand and supply of farm labor. That is, technical change in agricultural is more than just a generator of the problems in the sense of a once-and-for-all effect on the farm labor market. It also contributes to the persistence of the problems because the effects of technology in both the input and product markets are not instantaneous, but are spread or distributed over time. The longer the lag(s) in effect, the greater is the contribution of technology to the persistence (not necessarily the magnitude) of low returns to farm labor, unemployment, and underemployment.

Technological impact on the farm labor market includes a combination or convolution of distributed lag effects, which, in many cases, are highly interrelated. The number, characteristics, and extent of these effects depend on many aspects of both the agricultural sector and the rest of the economy. For example, such aspects include, among others, the nature of the technical change itself, characteristics of the cost structure of farm firms, and characteristics of variables such as government agricultural programs; national monetary and fiscal policies; weather; nonfarm employment conditions; laborer attitudes toward mobility and nonfarm work; and age, educational level, and marketable skills of farm workers.

The above argument does in no way attempt to imply that technology is the only cause of farm labor problems. But technology, through its stimuli on farm labor that could occur directly in the farm labor market and indirectly in the farm product and other input markets and in the nonfarm sector, could be a major causal factor in the

initiation and persistence of the problem. Nevertheless, much controversy remains as to the validity of this hypothesis, so further knowledge of technology and its effects on the farm labor market is needed. This study was directed toward this end because it sought to examine further the controversy surrounding the role of technology in the farm labor market.

III. JUSTIFICATION

Further economic study on the effect of technology on the farm labor market was needed for two major reasons. The first was that further knowledge of the effect of technology on the farm labor market could be useful to policymakers. The other was that hopefully some additional economic knowledge could be advanced concerning technology and the farm labor market.

Policymakers have realized for many years that discrepancies have existed between the lower incomes of farmers and farm laborers and the relatively higher incomes of managers and laborers in nonfarm occupations. The professed goal of many policymakers has been to reduce or remove this discrepancy. They attempted to do this using various policy alternatives ranging from production controls and price supports to manpower programs such as minimum wage legislation and vocational training programs. But the discrepancy has persisted and has even increased.⁴ Regardless of whether the policies were helpful, neutral,

⁴The discrepancy for both males and females can be observed in U.S. Bureau of the Census (1949-1970, specifically 1970, Table 496) for 1947 through 1968.

or perverse in their effects on farm labor problems, the point is that the income discrepancies still exist. A knowledge gap exists concerning the actual effects of various policies and the distribution of the effects over time. A fundamental prerequisite for choosing among policy alternatives and for achieving the intended goals is a reasonably reliable knowledge of the structure and the behavioral nature of the input and product markets with respect to both the relevant variables involved and the behavior of the markets through time. Lack of adequate information about the behavior of the farm labor market and other input markets and of the markets of agricultural products under conditions of changing technology can lead to policy effects that are restrained at less than their potential or even erroneous. Further knowledge of the farm labor market could be beneficial in making the correct choice of policy alternatives based on reliable estimates of the effects of changing technological conditions.

The second justification for this study was the possible gain in knowledge concerning the relationship between technology and the demand for farm labor. The relationships existing in the farm labor market have received less empirical attention when compared with work done in other areas such as in various commodity markets. In addition, incorporation of distributed lags into the relationships has received relatively little empirical attention in past farm labor research. Furthermore, this study attempted to go one step beyond past empirical research on the farm labor market. It attempted to include in the model the supply and demand for aggregate farm product, while incorporating distributed lags for the effects of technology on both the demand for

farm labor and the supply of aggregate farm product. Previous empirical research on the farm labor market, where distributed lags were employed, did not include equations for the product market in the system. The model of this study allowed a simultaneous determination of the parameter estimates in the product market and in the labor market.

IV. SOME PREVIOUS ECONOMIC RESEARCH ON TECHNOLOGY AND THE FARM LABOR MARKET

Several researchers recently have examined empirically the effects of technology on the farm labor market. The following studies will be discussed: Schuh (1962), Gisser (1965), Wallace and Hoover (1966), Bauer (1968), and Tyrchniewicz and Schuh (1969). Econometric analyses were used in each case in a simultaneous-equation framework.

Schuh was concerned with the hired labor market in agriculture. His general model was the two-equation simultaneous system:

$$\begin{array}{ll} \text{Hired labor supplied} & = f(\text{farm wages, nonfarm income, unemployment, size of civilian labor force}) \\ \text{in agriculture} & \\ \text{Hired labor demanded} & = g(\text{farm wages, prices of agricultural products, prices of other inputs, technology}) \\ \text{in agriculture} & \end{array}$$

Technology was a measure of output over input. He used annual time series data for the years 1929-1957 at the national level. His coefficient estimate of the technology variable was negative (but nonsignificant), so technological advancement has a depressing effect on the demand for farm labor and thus on farm labor income (if one ignores the nonsignificance).

Gisser (1965) investigated the effects of education on the farm labor market in general. He was not concerned with technology directly.

But as Bauer (1968, p. 5) pointed out, one can draw some implications from the coefficient of the education variable because, like technology, it effects the quality of the labor input. His model was the following general equations:

Demand: Farm wages = $f(\text{labor input, capital, level of education, race})$

Supply: Farm wages = $g(\text{labor input, alternative wage, level of education, race})$

Gisser, in estimating the coefficients, initially assumed a constant price of product, i.e., an infinitely elastic product demand. He then adjusted the estimated coefficients for an inelastic product demand. He found that with an inelastic product demand, a 10 percent increase in the level of schooling will lead to an equilibrium in the farm labor market consisting of a 6 to 7 percent lower quantity of labor at 5 percent higher wages. Gisser concluded that schooling in rural areas encourages farm outmigration, while also increasing income in rural areas. These results and conclusions imply an opposite positive effect (rather than negative as in other studies) of technology on farm labor income. However, Gisser used education which is not as inclusive as technology in general. Education is at most only a partial measure of the quality of the labor input. Quality of the labor input is determined also by other factors such as health, stamina, attitudes, and so on. In addition, a comprehensive measure of embodied technology also includes the quality inherent in all inputs in production.⁵

⁵ Furthermore, as Schuh (1966, p. 76) pointed out, an adequate measure of technology probably would include a provision for technology that is not embodied in inputs, e.g., improved organization or arrangement of inputs.

The study by Wallace and Hoover was concerned with the farm labor market in general. Their general implicit model was the following two-equation simultaneous system:

Demand for agricultural labor: $f(\text{labor, wages, technology, land, other inputs, age, education})$

Supply of agricultural labor: $g(\text{labor, wages, age, education, alternative wage})$

Their measure of technology was a proxy variable, public expenditures on research and extension. They reasoned that public expenditures on research and extension activities are highly positively related to technology. This measure treated technology as an explicit causal variable rather than incorporating it into the model as a residual. Wallace and Hoover used cross-sectional state data converted to logarithms for the year 1959. However, technology was measured as the geometric mean of the fiscal years 1955 to 1959 of public expenditures by state on experiment station research and cooperative extension service activity.

Results showed all variables to be significant⁶ with expected signs except age and education. Age was nonsignificant in both equations and education was nonsignificant in the demand equation.

A necessary assumption was that the demand for aggregate agricultural product was infinitely price elastic because no explicit provisions were made for the product market in the econometric model. With this assumption of perfectly elastic product demand, their results

⁶Significance was based on asymptotic standard errors in the structural equations.

indicated that increased technology increases the demand for farm labor and hence increases farm labor income.

However, Wallace and Hoover contended that in reality the demand for aggregate agricultural product is not perfectly price elastic, but is price inelastic.⁷ They assumed that the labor demand relationship is a value of marginal product curve derived from a Cobb-Douglas production function of aggregate farm output. A known, specific product demand relationship was assumed also. Armed with these assumptions, they were able to adjust the coefficient of public expenditures on research and extension activities for different product demand price elasticities after making the one additional assumption of equilibrium in the product market.

Adjustment for a price-inelastic demand for aggregate product resulted in an adjusted coefficient relating research and extension to the quantity demanded of farm labor that was negative rather than positive, as it was before adjustment. Bauer (1968, p. 8) pointed out that the adjusted coefficient becomes negative because improved technology increases the marginal productivity of labor, but the supply of product will increase, sharply driving down product prices because of a price-inelastic product demand. The decrease in product prices causes the value of marginal product (their demand for agricultural labor) to decline.

Bauer investigated the distributed lag effect of technology on

⁷This idea is widely accepted by agricultural economists. Brandow (1961, p. 50) estimated a farm-level price elasticity of demand for all farm food products at $-.2278$.

the farm labor market. His model was the following two-equation simultaneous system:

Demand for farm labor:

$$\text{Quantity of farm labor} = f(\text{farm wages, land stock, other expenditures, technology})$$

Supply of farm labor:

$$\text{Farm wages} = g(\text{quantity of farm labor, alternative wages, rate of unemployment})$$

He used a time series of cross sections of state data for 39 states or groups of states for the years 1951-1961. The data were transformed to logarithms. Technology was measured using the same proxy variable that Wallace and Hoover (1966) used, namely public expenditures for research and extension activities by state or group of states. However, Bauer did not use an arbitrary lag function like Wallace and Hoover used. Instead, he calculated the weights of the lag distribution of the effect of technology on the demand for farm labor. The weights were based on lag parameter estimates obtained by using a rational distributed lag function (Jorgenson, 1966). He then used the calculated weights to compute a new technology variable which was a weighted average for 1959 and the 18 previous years. He used this newly computed technology measure in conjunction with 1959 state data on the other variables to estimate the structural parameters of the model. He used cross-sectional data for 1959 so that his results could be compared to the results of Wallace and Hoover.

He selected two lag functions out of five alternatives that he estimated. The mean length of one function was 49.2 years and the mean of the other, about 45.1 years. The median of one function

was 27 years and of the other, about 22 years respectively. The configurations of the estimated lag distributions had a "hump," so that the greatest effect of a change in technology on the labor market imposed within any given time period occurs after the passage of time rather than in the initial period.

Bauer made the same assumptions concerning the product market as Wallace and Hoover. Initially, he assumed that the demand for aggregate farm product was perfectly price elastic. His estimate of the structural parameter for the overall effect of technology on the quantity of farm labor was positive. He later relaxed the assumption of perfect elasticity using the method employed by Wallace and Hoover discussed above. He adjusted the technology parameter for the effects of price elasticity of demand that was less than infinity. The adjusted coefficients of the technology variable became negative when the product demand function became price inelastic.

Tyrchniewicz and Schuh (1969) separated the hired, operator, and family labor components of farm labor to attempt to account for various interdependencies among the different components of the farm labor force. They hypothesized a six-equation simultaneous-equation model consisting of demand and supply equations for each component.⁸ Price and quantity of each component were jointly determined subject to a set of exogenous or predetermined variables.

Time series national data for 1929-1961 were used in fitting

⁸The model was not recorded here because of its length. See Tyrchniewicz and Schuh (1969, pp. 771-776) for the conceptual and statistical models.

after converting to logarithms. A measure of technology was introduced into the demand equation for each component. Ruttan's index of technology⁹ was used to represent the technology variable. They attempted to identify the role of technology in the component markets of farm labor, but were unsuccessful. They tried technology both as an exogenous variable and as an endogenous variable, but none of the experiments proved fruitful. No numerical results were reported with respect to the technology variable.¹⁰

V. PURPOSE AND OBJECTIVES OF THE STUDY

This study specified, in addition to the farm labor market equations, the equations of the market for aggregate farm product. This was an attempt to avoid extraneously adjusting the coefficient of technology for the effect of a product demand price elasticity that was less than infinity. The price elasticity of product demand was estimated directly in the system. In other words, it was a formal attempt to avoid simultaneous equation bias caused by joint dependencies in the product and labor markets when estimating the effect of technology in the labor demand equation. The model, as specified, further extended the work of Bauer (1968) and others by incorporating a distributed lag function for the effect of technology in the aggregate-product supply equation in addition to the distributed lag function for the effect of technology in the labor demand equation.

⁹See Stout and Ruttan (1958) for a discussion of the index.

¹⁰See the original study, Tyrchniewicz (1967), for the results of these experiments.

The objectives were (1) to estimate the effect of technology on the farm labor market and (2) to estimate the distributed lag function or time path of this effect.

CHAPTER II

DERIVATION AND DESCRIPTION OF THE ECONOMIC MODEL

Specification of a conceptual model from a puristic point of view ignores anticipated or unanticipated constraints such as lack of necessary analytical skills, techniques, and data that are expected to arise as the research process proceeds. That is, specification of the conceptual model does not depend on these constraints, but depends primarily on the nature of the problem under study.

Nevertheless, one could carry to an unmanageable extreme the specification of an "ideal" economic model by attempting to account for all seemingly relevant interrelationships. Consequently, even in the early conceptual stage, the model was constrained at a level of conceptualization that perhaps was less than the "ideal." However, the assumption was made that the model was suitable and adequate for the problem being examined. This chapter is an examination of the basic conceptual model, some underlying assumptions concerning it, and some preconceptions regarding interrelationships in the model.

I. THE MODEL

To avoid some undue complexities at the outset and because the primary objective was to examine the effects of technology on the demand for farm labor, the decision was made to limit the conceptual model to a six-equation, simultaneous-equation system. The system included the demand and supply equations for one factor, farm labor,

the farm-level demand and supply equations for aggregate farm product and the respective equilibrium identities. The model¹ was written in general form as:

Demand for farm labor:

$$(1) \quad Q_{it} = a (W_{it}, P_{it}, P_{it}^S, P_{it}^O, R_{it} \dots R_{i,t-s}, M_{it}, E_{it}, A_{it}, D_{mit})$$

Supply of farm labor:

$$(2) \quad Q_{it} = b (W_{it}, \bar{W}_{it}, U_{it}^r, U_{it}^u, N_{it}^r, M_{it}^g, W_{it}^r, W_{it}^u, E_{it}, A_{it}, D_{mit})$$

Demand for Aggregate Product:

$$(3) \quad Y_{it} = c (P_{it}, \bar{P}_{it}, N_{it}, I_{it}, H_{it}, T_{it}, D_{mit})$$

Supply of Aggregate Product:

$$(4) \quad Y_{it} = d (P_{it}, W_{it}, P_{it}^S, P_{it}^O, G_{it}, H_{it}, R_{it} \dots R_{i,t-s}, W_{it}^w, M_{it}, E_{it}^O, A_{it}^O, D_{mit})$$

for the i-th state and the t-th year, the variables were:

Q_{it} = quantity of farm labor

W_{it} = farm wage rate

P_{it} = farm-level price of aggregate farm product

P_{it}^S = price of land

P_{it}^O = price of all other inputs other than land and labor

$R_{it} \dots R_{i,t-s}$ = current (s=0) and lagged (s=1 ... n-1) technology
where s = lag time and n = number of lag weights on R.

M_{it} = level of farm mechanization

¹The identities (with "tildes" indicating equilibrium quantities):

$$\tilde{Q}_{it} \text{ supplied} = \tilde{Q}_{it} \text{ demanded}$$

$$\tilde{Y}_{it} \text{ supplied} = \tilde{Y}_{it} \text{ demanded}$$

which satisfied the equilibrium assumptions of the two markets, were assumed to hold.

E_{it} = average educational level of farm working force

A_{it} = average age of farm working force

D_{mit} = dummy variable for state location ($m=1 \dots 39$ states and groups of states)

$\bar{W}_{it}$ = alternative wage rate

U_{it}^r = rural unemployment rate

U_{it}^u = urban unemployment rate

N_{it}^r = rural population

M_{it}^g = rural to urban migration rate

W_{it}^r = welfare payments to rural poor

W_{it}^u = welfare payments to urban poor

Y_{it} = quantity of aggregate farm product

$\bar{P}_{it}$ = price of nonfarm commodities

N_{it} = population of consumers of farm products

I_{it} = disposable income of consumers of farm products

H_{it} = home consumption of farm products

T_{it} = tastes and preferences of consumers of farm products

G_{it} = government payments to farmers

W_{it}^w = weather

E_{it}^o = average education of farm operators

A_{it}^o = average age of farm operators

The variables Q_{it} , W_{it} , Y_{it} , and P_{it} were assumed to be jointly dependent or endogenous variables. All of the other variables were treated as exogenous variables.

II. SOME GENERAL UNDERLYING ASSUMPTIONS REGARDING MODEL SPECIFICATION

Several underlying assumptions were made regarding the nature of the conceptual model. They were made to facilitate the most direct and simple, yet adequate approach to examining empirically the distributed lag effect of technology on the demand for farm labor at the state level.

Structural Change and Functional Shifts

A conceptual and empirical problem in most economic studies is the difficulty of distinguishing structural change and shifts in a function. Learn and Cochrane (1961, pp. 63-72) considered this problem with respect to supply analyses utilizing regression techniques. However, many of their general points are applicable to both supply and demand. For example, they defined a shifter variable as one that causes a parallel shift in a function when the change in the value of the variable occurs within a fixed environment. The ceteris paribus conditions strictly hold. Structural change results from a change in the environment surrounding and including the system. Structural change occurs when some force exerted on the particular system under examination results in a change in one or more parameters of the function or a change in the form of the function itself. Consequently, structural change occurs from changes in one or more of the factors making up the environment included in the ceteris paribus conditions.

The model of this study contained variables that can be regarded conceptually as shifters such as the price of land and other nonlabor

inputs in the labor demand equation and prices of inputs in the product supply equation. However, the model also contained variables which are structural or quasi-structural in nature which not only can change the value of the constant term parameter, but the parameters associated with other variables as well, or even change the form of the supply or demand function under consideration. Distinguishing between the type of variables was not a simple task since many variables cannot be uniquely classified as causing shifts or structural changes. For example, Learn and Cochrane (1961, p. 69) discussed that some variables treated as structural in nature may give rise to shifts in the function of far greater importance than whatever structural impact they may have. Or conversely, some variables which are not strictly structural in nature affect structural variables and reflect structural change. Consequently, no attempt was made to classify variables into structural or shifter variable classifications. However, some of the pitfalls of analyzing a system undergoing structural change were recognized, along with some of the limitations.

The Dynamic Nature of the System

The model, although incorporating distributed lags on the technology variable, did not include all possible dynamic aspects. Admittedly, an "ideal" specification of an economic system would contain several variables with adjustment processes that are not instantaneous, but are spread over time. For example, contemporaneous wage rate in the labor demand and supply equations probably is conceptually inferior to some other concept of the price of labor such as expected wage rate. Other examples are the use of a measure of

permanent income instead of contemporaneous disposable income in the product demand equation and the use of expected product price rather than actual price in the product supply equation. However, the decision was made, given the objectives of the study, to ignore the dynamic aspects of variables other than technology because of substantially increased complexity.²

Degree of Competition, Specialized Factors, and Other Questions of Specification as Related to Simultaneity

The model as hypothesized was primarily concerned with measuring the distributed lag effect of technology on farm labor demand at the industry level. However, to examine this technological effect on labor required a model which incorporated other variables so that demand and supply curves could be specified. Basic economic theory was helpful in the formulation of the conceptual model, firstly, in selection of relevant variables to consider, and secondly, in establishing relevant theoretical rationale and adjoining assumptions to warrant their inclusion as variables in the demand and supply equations.

Competition was assumed in both the product and factor markets. That is, the individual farmer at the firm level accepts as given the prices he pays for inputs and the price of his product. Prices of inputs at the industry level other than labor were taken as given. The quantity of labor, the wage rate, and the quantity and price of

²Some of the other possible dynamic aspects of technology were ignored also. For example, technology was not specified as an exogenous variable with a distributed lag effect in the labor supply and product demand equation.

aggregate product were determined simultaneously given these prices of inputs other than labor and the remaining exogenous variables as specified in the respective equations of the model.

The model with prices of inputs other than labor, such as land and other nonlabor factors, specified as being exogenous to the system, was to some extent unrealistic. This assumption implied that the variable factors, agricultural land and other nonlabor inputs, are nonspecialized in the farming industry. That is, the farming industry is only one of many industries using the factors, land and other nonlabor inputs. Since the prices of these factors were assumed to be exogenous variables in the model, the actions of the farming industry have no effect on these prices. However, changes in the product and labor markets in the farming industry should affect the price of agricultural land, which is, in actuality, an input highly specialized to agriculture. The same applies to such inputs as farm machinery, fertilizer, etc. Nevertheless, the model ignored the effects of changes in the agricultural product and labor markets on land and other nonlabor inputs because the additional complexity was assumed to outweigh any contribution relative to the stated objectives.

Similar reasoning was applied to the decision to omit an equation in which technology, as an endogenous or dependent variable, was a function of several other variables such as farm product price. The same reasoning held for considering only as exogenous such variables as level of mechanization, rural to urban migration, and government payments to farmers. The increased complexity of adding the additional equation was assumed to outweigh the added contribution to improved model specification, especially for the problem at hand.

Market Levels

Equations to account for the market levels between the farm, wholesale, and retail levels in the system could have been desirable from a conceptual standpoint. The utilization of such equations could have improved the economic specification by separately accounting for primary and secondary market levels. However, such a specification, using aggregate data on many commodities, which, in many cases, have unique marketing structures, would have been of questionable practicality. Problems of measurement at the different market levels of the variables aggregated across commodities within states would have ruled out empirical analysis. For example, data on the index of prices received by farmers (the farm level) existed by state. Individual commodities making up the aggregate on which the price index is based are changed in form through packaging, preservation, and transformation, so that the products are no longer the same upon reaching the retail level. Evidence of this difference can be observed in the grouping of available national data on price indices at the retail level. The indices no longer measure the aggregate price of farm commodities because the commodities have lost their identity in most cases by the time they reach the retail level. The measurement problem is intensified because state data on prices for categories of commodities at the retail level do not exist for most states.

Although equations for the product market were specified for farm-level prices and quantities, an attempt was made to account for the primary or retail level³ in the demand equation by including

³Breimyer (1962, pp. 679-699) discussed some of the shortcomings

disposable income and population as shifters. The argument could be forwarded that the demand equation as specified did not represent a true demand equation, since the market levels were combined. But under the existing limitations, this specification was taken as a reasonable compromise, rather than adding additional market-level equations to the system. Income and population were assumed to affect indirectly the secondary level (farm) demand through the retail and wholesale levels.

Nature of Technological Change

No assumptions were made regarding the specific nature of the technological change occurring in the system. For example, the model was not restricted to specify an underlying industry production function under a specific condition of changing technology, e.g., output-increasing, labor-using, capital-using, neutral technical change, etc. Controversy exists among agricultural economists as pointed out in Chapter I concerning the causes of low returns to agricultural resources. Different assumptions regarding the nature of technological change, among other things, underlie the alternative viewpoints relating technology as a cause of farm labor problems. For example, Cochrane's "economic treadmill" relied on output-increasing technology combined with an inelastic demand for product which shifts to the right relatively less over time than does product supply. Tweeten's (1969, pp. 798-817) deemphasis of the importance of technology as a cause of

of attempting to link primary production using the supply schedule at the farm level directly to the consumer demand schedule.

low returns relied on the assumptions of labor-saving technical progress, resource fixities, the cost structure of the farming industry, and product demand keeping pace with supply in the 1960's.

As implied by Tweeten, the nature of technological progress in agriculture may have changed from that of output-increasing technology prior to the 1960's to that of labor-saving technology during the 1960's. Nevertheless, no explicit initial assumptions were made regarding the effects of changes in technology on the relative and absolute input productivities in the underlying production function.

III. SOME PRECONCEPTIONS CONCERNING CAUSES AND EFFECTS IN THE SYSTEM

The system, being simultaneous, included four variables--wage rate, price of product, quantity of labor, and quantity of product--which were jointly determined within the system. All remaining variables were determined outside the system. Thus, cause and effect relationships existed among the equations of the system and also within each single equation. Equilibrium prices and quantities in the product and labor markets were jointly determined through the interactions of these jointly dependent variables with each other and the influence of the predetermined variables within the system.

Discussion of cause and effect in the structural model of any simultaneous system, even at a general, conceptual level, becomes quite complex. For example, in a particular structural equation of a multi-equation system, a change in the level of a specified exogenous variable not only leads to a change in the left-hand (normalized) endogenous

variable for the equation, but in the other endogenous variables (on the right side of the equal sign) as well. The other endogenous variables are changed since (1) they may be themselves each a function of the same exogenous variable in other equations of the system and/or (2) they are each functions of all exogenous variables in the equations of the reduced form. Hence, the nature of a simultaneous system leads to a circular process of change in all endogenous variables in all equations of the system as a result of a change in one exogenous variable in one of the structural equations (or more than one if the exogenous variable is in more than one of the structural equations).

Some preconceptions concerning the causes and effects occurring within each equation of the system of this study are developed in this section. The likely direction of the relationship(s) between the jointly dependent variable on the left side of the equation and the jointly dependent variable(s) on the right side of the equation are examined along with the likely direction of influence of each of the predetermined variables.⁴

The Demand for Farm Labor

The labor demand equation hypothesized that the quantity of labor demanded by the farming industry of each state is jointly determined with farm wages and the price of aggregate product and depends upon prices of nonlabor inputs, technology, level of mechanization, education, and age of the farm working force and state location.

⁴The direction of each joint relationship or influence was considered under the condition that all other variables and conditions remain constant in all equations.

Prior expectations regarding the direction of the joint relationship or influence of each of the different variables were the following:

1. The wage rate was expected to vary inversely with the quantity of labor demanded. Farmers were expected to employ less labor as the wage rate increased or more labor if the wage rate decreased.

2. The demand for labor should shift directly with changes in the price of the product. For example, an increase in the price of aggregate farm product should shift the demand for labor to the right.

3. Labor demand should shift inversely with a change in the price of another input if the input is an economic complement with respect to labor. Labor demand should shift in the same direction as a change in the price of another input if the input is an economic substitute for labor. Or more specifically, factors in production of a given product are economic complements or substitutes if the cross price elasticity of demand is negative or positive, respectively. Each of the factors, land and other nonlabor inputs, were expected to be economic substitutes with respect to labor. Consequently, positive or direct relationships between the price of land and the demand for labor and between the price index of other nonlabor inputs and the demand for labor were conceptually reasonable.

4. Technology was expected to have a negative effect on the demand for labor. Furthermore, the effect of technology was not expected to be instantaneous, but distributed over time. The shape and length of the distribution of this effect was undecided a priori, but some

past studies⁵ indicated that it should be bell-shaped, possibly skewed to the right, and quite long in terms of years.

5. The level of mechanization in farming was expected to have a negative influence on the demand for farm labor. For example, the increased use of farm machinery and equipment was expected to lead to an increase in the demand for land while decreasing the demand for labor. The rate of mechanization would be dependent upon technical conditions facing the industry, and upon market conditions in the factor and product markets.⁶

6. Education and age were included as an attempt to quantify labor quality. The demand for labor was expected to increase as labor quality increases. The demand for labor was expected to vary directly with the level of education. However, the relationship between age and quality was not expected to be linear and monotonic, but curvilinear. Quality and employment opportunity probably vary directly with age up to a certain level and then decline as age increases. Consequently, the expected direction of the effect of age on the demand for farm labor would depend on the magnitude of the age variable, i.e., the demand for labor would vary directly with age up to a certain age. Then labor demand would vary inversely at higher age levels.

⁵Bauer (1968), in particular.

⁶The model did not account for the two-way causality between factor markets and the level of mechanization and between the market and the level of mechanization because the level of mechanization was treated only as an exogenous variable.

7. Locational differences in the demand for labor were expected to exist. These differences would arise because of locational differences in levels of and responses to input prices, technology, education levels, etc.

The Supply of Farm Labor

The farm labor supply equation in the model expressed the quantity of farm labor supplied as being jointly dependent with farm wages and dependent upon alternative nonfarm wages, rural and urban unemployment rates, rural in-state population, rural to urban migration, welfare payments to rural poor, welfare payments to urban poor, education, age, and geographic location. Conceptually, prior expectations regarding the direction of the relationship between the individual variable and the quantity of labor supplied when all other variables in the equation and other equations of the system were held constant were the following:

1. The quantity of labor supplied was expected to vary directly with the wage rate, i.e., the supply curve would be positively sloped.

2. The supply of farm labor was expected to be affected inversely by changes in the alternative wage rate. For example, as alternative nonfarm wage rates increase, rural farm laborers would have greater incentives to seek nonfarm employment. Therefore, the supply curve for farm labor would shift to the left.

3. The unemployment rate is one indicator of the potential (or lack of potential) for employment. A high rate of unemployment in the farm and nonfarm sectors would tend to have a positive influence on the supply of farm labor.

4. Farm labor supply was hypothesized to vary directly with rural population. For example, as the rural population increases, the supply of farm labor was expected to increase.

5. Changes in rural to urban migration was expected to affect inversely the supply of farm labor. For example, a decrease in rural to urban migration would shift the supply curve of labor to the right.

6. The direction of influence of welfare payments on the supply of farm labor was expected to depend on many aspects such as the nature of the payment, the requirements necessary to qualify for receiving and continuing to receive payments, the effects of payments on the incentives of workers and the unemployed, the existence or non-existence and magnitudes of payments in the farm and nonfarm sectors, etc. In general, one would expect under existing welfare programs, a negative influence on the supply of labor caused by welfare payments in the farm and nonfarm sectors. The influence was expected to be negative in both cases under existing welfare programs because of net disincentives to work.

7. Education and age of the farm labor force were expected to affect the supply function. Less educated and older farm people tend to seek fewer nonfarm employment opportunities. Consequently, the supply of labor would vary inversely with respect to a change in the level of education and directly with respect to a change in the average age of the farm working force.

8. Locational differences in the supply of labor were expected to exist because of state differences in levels of and responses to alternative employment opportunities, education, age, welfare payments, etc.

Demand for Aggregate Farm Product

The quantity of aggregate product demanded was hypothesized to be jointly dependent with aggregate product price and dependent upon the price of nonfarm commodities, the population of consumers of farm products of a state, total disposable income of these consumers, home consumption of farm products, tastes and preferences of the consuming public, and location. Expected direction of the economic effect of each of these variables were the following:

1. The quantity of farm product demanded was expected to vary inversely with the price of the product. That is, the demand curve was expected to be negatively sloped,
2. The demand for aggregate farm product was expected to vary directly with the price index of aggregate nonfarm commodities. The aggregate of nonfarm commodities contains commodities that are complementary to farm products in addition to commodities which are substitutes. However, the positive substitution effect was expected to outweigh the negative complementary effect of "price" of nonfarm commodities in the aggregate.
3. Changes in the consuming population were hypothesized to cause variations in the same direction in the demand for product.
4. The demand for product was expected to vary directly with and as a result of changes in the disposable incomes of consumers of farm products. This would be a net, direct relationship because aggregate farm product contains some commodities that could qualify as inferior goods. However, in the aggregate the income elasticity was expected to be greater than zero.

5. Changes in the level of home consumption of farm products was expected to cause the demand for aggregate products to shift inversely. For example, as home consumption increases, the demand curve for aggregate product would shift to the left. Farmers are producers, but they are also consumers of farm products.⁷ Farmers demand fewer purchased commodities of farm origin when they consume more of their own products at home without first selling them and then buying the products or their derivatives as consumers in the market place. Consequently, if farmers as consumers demand less on the market, the demand for aggregate farm product would shift to the left.

6. Aside from the measurement problem, the many facets of consumer and human behavior that, in large part, fall outside the realm of economics, prevented any substantive statement regarding direction of effect on demand of changes in tastes and preferences. Furthermore, tastes and preferences over time are conditioned or changed as a result of changes in availability of various products in the market place, incomes, etc. However, the model accounts only for the effect of tastes and preferences on demand. As tastes and preferences for farm products in the aggregate increase, the demand for aggregate farm product would increase. Or, as tastes and preferences for farm products in the aggregate decrease, the demand for aggregate farm product was expected to decrease.

⁷The failure of the model to distinguish the various market levels prevented a thorough explanation of the market behavior of farmers as producers and as consumers. Furthermore, the consumer good in many cases bears little resemblance to the raw farm product.

7. Locational differences in demand for aggregate product were expected to exist. Levels and responses to such variables as population, income, tastes and preferences, alternative prices, etc., vary from location to location.

Supply of Aggregate Farm Product

The quantity of aggregate farm product supplied by farmers was hypothesized to be jointly dependent with product price and the wage rate and influenced by the prices of the factors of production other than labor, government payments, home consumption of farm products, weather, level of mechanization, management, technological conditions, and geographical location. Given that all other variables remain constant, expected direction of the relationships were the following:

1. The quantity of farm product supplied by the farmers of a state was expected to vary directly with changes in product price. Perhaps expected price would have been more appropriate rather than actual price because the production process is not instantaneous and because farmers must plan, in general, on the basis of expected product price rather than actual price for most commodities. Actual price is usually determined in the market following the planning and production phases. Nevertheless, actual price was considered adequate for the purposes of this study.

2. The supply curve for aggregate farm product was expected to shift inversely with changes in factor prices. For example, as factor prices increase, the supply curve for aggregate farm product would shift to the left. This direction of movement applies to the price of each factor--labor, land, and other inputs.

3. Changes in government payments were expected to have a net direct influence on the supply of product in the long run. This expected net influence in the long run is contrary to the original intentions of most policymakers in setting up such programs as price supports, acreage controls, conservation reserve programs, etc. Such programs to some extent probably operate as intended in the short run. However, once farmers are given time to adjust to the presence of a particular program(s) and to plan accordingly, given the government policy and expected market conditions, supply would not be curtailed in general, but increased in the aggregate.

4. The supply of aggregate product was expected to vary inversely with changes in home consumption, e.g., when farmers, as consumers, consume more at home rather than marketing this portion of their product, the supply of aggregate product would shift to the left.

5. Changes in technology were anticipated to influence in the same direction the supply curve of product. Also, like technology in the labor demand equation above, the effect of technology on supply was not expected to be instantaneous, but distributed over time.

6. Weather, measured by an index which takes into account all meteorological and closely associated factors that affect production, was hypothesized to be a major influence on the supply of agricultural product.⁸ Weather generally is considered a major random influence

⁸The effects of weather on crop production generally is considered to be more important than its effect on livestock production. Weather affects crop production directly and livestock production directly to a lesser extent. However, since livestock consume feed originating from current crop production and storage, livestock production is affected both indirectly and directly over time.

in agriculture as contrasted with nonrandom influences such as technology. The direction of the influence of weather on supply would be in the same direction as its influence on production, other things remaining constant. For example, as the index of weather varies from bad to good relative to its effect on agricultural production, then the supply of product would increase accordingly.

7. The supply of product was expected to be influenced in the same direction by changes in the level of mechanization. Farmers in the past have mechanized to increase net revenues along with other reasons. Higher levels of mechanization generally require expansion of their operations to larger farm sizes. The effects of this expansion in the aggregate has been a reduction in the number of farms, an increase in the size of farms, and an overall aggregate increase in output.

8. Education and age of farm operators were included in the supply equation to serve as proxies for entrepreneurial capacity of farmers. Entrepreneurial capacity was expected to vary directly with education. However, entrepreneurial capacity was expected to increase during earlier years of age, reach a maximum, and then decrease as the farmer becomes older. Thus, the supply of agricultural product was expected to increase with increases in the average educational level of farmers. The converse was assumed for decreases. The product supply curve was assumed to shift to the right as the average age of farmers increases and then to start shifting to the left as the average age of farmers advances past some point.

9. Locational differences in the product supply curve by state were expected to exist because of locational differences in levels and responses to such variables as weather, technology, and levels of mechanization.

CHAPTER III

CONCEPTS AND ECONOMETRIC ASPECTS OF DISTRIBUTED LAGS

The statistical model, formulated from the conceptual model of the previous chapter and presented later in Chapter V, employed a distributed lag estimation technique. Some understanding of the basic concepts and econometric aspects of distributed lag analysis was necessary before proceeding to the statistical model. This chapter is concerned with the general concept of distributed lags and some different, available methods of restricting the parameter space of the general distributed lag function. Choice criteria for the selection of the method of restriction used in this study are presented.

I. THE CONCEPT OF DISTRIBUTED LAGS

A causal variable in economics often produces its effect over a period of time rather than instantaneously. That is, there is a lapse in time between a change in a causal variable and its resulting effect. The lapse of time is referred to as a lag. As Nerlove pointed out (1958, p. 1), the lag may be a specific time period such as one week or two months, but in many cases the effects of a change in an economic causal variable are spread over many weeks, months, or years. The spreading of this effect over many time periods is referred to as a distributed lag.

The general distributed lag function¹ can be written as

¹Parts of this discussion rely heavily on Griliches (1967).

$$(5) \quad Y_t = f(X_t, X_{t-1}, X_{t-2}, \dots, \mu_t)$$

where Y_t , the dependent or endogenous variable, is a function of both contemporaneous and past values of X , the exogenous or predetermined variable, and also μ_t , a stationary random variable with zero mean and a fixed covariance structure. The subscript t represents time. That is, the effect of X on Y is not instantaneous. A change in X or ΔX in a given time period t affects Y in the same period t , but also in several following periods. Assuming that the function is linear, one can write the function algebraically as

$$(6) \quad Y_t = \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \dots + \mu_t$$

If the β 's have a finite sum ($\sum_1 \beta_1 < \infty$) and all have the same sign, the effect of a change in X can be said to be distributed or allotted between Y 's of different time periods which allows the following equation to be written:

$$(7) \quad Y_t = [w_0 X_t + w_1 X_{t-1} + w_2 X_{t-2} + \dots] + \mu_t$$

where the $w_1 \geq 0$ are all nonnegative and sum to one, $\sum_{i=0}^{\infty} w_i = 1$. The equation states that current Y (where t can be assumed to be the current period) is a function of current and past values of X , with the weights w_i indicating the relative influence of differently lagged values of the X 's on current Y 's. The equation shows that given a change in X of ΔX , Y still will be changed by $\beta \Delta X$, but the total effect is spread over several time periods. Also, if one assumes for explanatory purposes that the level of X undergoes a once-and-for-all change in time period t (i.e., X changes in period t by ΔX , but did not change in previous time periods and will not change in future

periods), then Y changes by $\beta w_0 \Delta X$ in period t and $\beta w_1 \Delta X$ in period $t+1$ and so on. Thus, the form of the lag, time shape, or time path of the effect of a change in X on Y is described by the sequence of the w_1 's.

The sum of the w_1 's equals one because only the total effect is to be distributed. The restriction $w_1 \geq 0$ indicates that for any given time period only a nonnegative portion of the total effect of X on Y is reasonable. In addition, the restrictions that the w 's are nonnegative and sum to one allow the w 's to be identified formally with probabilities defined over a set of nonnegative integers $(0, 1, 2, 3, \dots, \infty)$. Griliches (1967, pp. 18-19) stated that this identification of the w 's with probabilities permits the use of all the available results about different types of probability distributions and their moments in discussing the form of the lag and various parameters of the lag distribution such as the average or mean lag and the variance of the lag. Griliches (1967, p. 19) used the example of a generating function for the sequence $w_0, w_1, w_2 \dots$ and defined a polynomial function

$$(8) \quad A(z) = a_0 + a_1 z + a_2 z^2 + a_3 z^3 + \dots$$

where z is a dummy variable. The function $A(z)$ is a probability generating function of the sequence (a_1) if the function converges in some interval $-z_0 < z < \bar{z}_0$ and if the a_1 's are nonnegative and $A(1) = 1$, i.e., the a_1 's sum to unity. Then he listed the advantages of using the notion of a probability generating function:

. . . (1) $A(z)$ may reduce itself to a relatively simple algebraic form even though the sequence a_1 is quite complicated; (2) one can get the parameters (moments) of the probability distribution

directly as functions of the derivatives of the generating function; and (3) it facilitates greatly the derivation of distributions of sums of random variables . . . [such as] . . . the cascading or convolution of several lag distributions.

Rationale for the restrictions placed upon the sequence a_1 and the notion of a probability generating function become more apparent in the discussion later in this chapter and in the estimation of the distributed lag functions and the average lag in Chapter V of this paper.

II. METHODS OF RESTRICTING THE PARAMETER SPACE OF THE GENERAL DISTRIBUTED LAG FUNCTION

Some serious problems arise at the outset in the empirical estimation of distributed lag functions. For example, the distributed lag function

$$(9) \quad Y_t = \beta [w_0 X_t + w_1 X_{t-1} + w_2 X_{t-2} + \dots] + \mu_t$$

$$w_1 \geq 0 \text{ and } \sum_{i=0}^{\infty} w_i = 1$$

cannot be estimated directly because there is an infinite number of regressors which consume an infinite number of degrees of freedom. However, only a finite number of observations are available; hence, only a finite number of degrees of freedom, so the equation cannot be estimated in its present form. Also, multicollinearity most likely would occur among successive regressors. Some method of transforming this equation to one that contains a finite number of variables must be used so that the equation can be estimated. In transforming the equation, it becomes necessary to make some reasonable yet restrictive

assumption(s) about the pattern of the weights. The objective is to choose an assumption which allows the individual coefficients of the lag distribution to depend on a few parameters, which in turn are relatively simple to estimate.

Three different methods of restricting the parameter space of a distributed lag function have been used in economic research. The two methods frequently used are the introduction of a finite number of lagged exogenous variables in the model (e.g., Almon, 1965) and the introduction of lagged endogenous variables by using a reduced equation (e.g., Koyck, 1954). Another method (Jorgenson, 1966), which involves a combination of the first two methods, has been used occasionally. This study used an alternative included in the method which uses a finite number of lagged exogenous variables. Two alternatives (Almon, 1965, and Johnson, 1969) of this method will be explained following the discussion of the other two methods of restricting the parameter space.

Distributed Lag Functions Using Lagged Endogenous Variables

The method which introduces lagged endogenous variables by using a reduction technique has been used quite frequently in economic research.² For example, Koyck (1954) specified that the coefficients of the lagged independent variable decline geometrically over time. He wrote:

$$(10) \quad Y_t = [(1-\lambda)\lambda^i]X_{t-i} \quad i=0 \dots \infty$$

so that

$$(11) \quad Y_t = \beta(1-\lambda)[X_t + \lambda X_{t-1} + \lambda^2 X_{t-2} + \lambda^3 X_{t-3} + \dots]$$

²See Koyck (1954), Cagan (1956), and Nerlove (1956 and 1958) for some pioneering work.

which cannot be estimated in general. At best the original equation can be estimated only with a very large number of lagged variables requiring a large number of degrees of freedom and hence, a large number of observations. Abundance of data over many time periods is not the usual case in economic research. The equation, which contains an infinite series, can be reduced to

$$(12) \quad Y_t = \lambda Y_{t-1} + \beta(1-\lambda)X_t$$

by use of the Koyck transformation. That is, the reduced equation is derived by multiplying the original equation by λ , lagging it one time period, and then subtracting the resulting equation from the original equation. The reduced equation can be estimated with only one lagged endogenous variable.³ The series of weights w_1 of the original equation, showing the time path of the effect of X on Y , is found by substituting the estimates of λ into

$$(13) \quad w_1 = (1-\lambda)\lambda^1$$

The Koyck distribution demonstrated here is a monotonically declining geometric distribution which means that the weights decline as time passes. The reduced equation, as written above, constrains the shape of the time path of the lag to one where the largest portion of the total effect of X on Y occurs in the initial time period and with progressively smaller portions of the total effect attributed

³ This is not to say that the reduced equation in this example is easily estimated because other statistical problems are usually present such as the disturbance of the reduced equation being serially correlated. Problems are compounded even more if the disturbance μ_t of the original equation is also serially correlated. See Goldberger (1964, pp. 276-278) for a discussion of estimation and some problems involved. Note his references to Koyck (1954), Nerlove (1958), and Klein (1958).

to the X of each period as time passes.⁴ The configuration of the lag distribution of this example is imposed a priori on the data. This, in itself, may be a disadvantage for a particular purpose at hand if (1) the theoretically desired configuration is felt to be something else or (2) the form of the lag is unknown. Consequently, another configuration may be desirable for theoretical reasons, but even more computational and statistical problems generally occur as the distribution is less constrained. For example, the Pascal distribution (Solow, 1960) does not restrict the data to a monotonically declining lag distribution as with the Koyck distribution, but can yield time paths of various shapes using the same basic distribution. The weights generated from the Pascal distribution are

$$(14) \quad w_i = (r + i - 1) (1-\lambda) r \lambda^{i-1}$$

where the two parameters to be estimated are r and λ . The parameter r can take only integer values.

A disadvantage of the Pascal distribution when compared with the Koyck distribution is that the Pascal adds two parameters to be estimated whereas the Koyck adds only one. (The Pascal distribution is the same as the Koyck distribution where r equals one.) A major estimation problem with the Pascal distribution in addition to the two added parameters to be estimated is that nonlinear estimation results

⁴ Another possibility is that the coefficients decline after the $k-1$ time periods. The equation

$$Y_t = \alpha + \sum_{i=0}^{k-1} \beta_i X_{t-i} + \beta [(1-\lambda)\lambda^j] X_{t-j} \quad j = k \dots \infty$$

shows that starting with the $(t-j)$ time period (where $j=k$) the coefficients decline geometrically.

when r is greater than one.

Rational Distributed Lag Functions

Another method of restricting the parameter space in order to estimate distributed lags is to use a combination of lagged endogenous and lagged exogenous variables. This method is referred to in this study and some others as the rational distributed lag technique, which was developed by Jorgenson (1966). The major advantage of this method is that one basic distribution can yield a variety of lag configurations, while being easier to estimate than the Pascal distribution when r is greater than one.

To facilitate more efficient explanation of this technique and to relate it to the method discussed above and also to the technique used in this study, it became necessary to resort to some commonly used notation on lag distributions and to the explanation and use of the concept of the lag operator L . The lag operator L indicates the operation of shifting a variable backward through time.⁵

$$(15) \quad LX_t = X_{t-1}$$

$$(16) \quad L^2X_t = X_{t-2}$$

The original form of the distributed lag function is

$$(17) \quad Y_t = \beta(w_0X_t + w_1X_{t-1} + w_2X_{t-2} + \dots)$$

By using the lag operator, it can be written as

$$(18) \quad Y_t = (w_0X_t + w_1LX_t + w_2L^2X_t + \dots)$$

⁵See Griliches (1967, pp. 18-24) for a detailed discussion of the lag operator notation.

or

$$(19) \quad Y_t = \beta(w_0 + w_1L + w_2L^2 + \dots) X_t$$

or

$$(20) \quad Y_t = \beta W(L) X_t$$

The shorthand notation $W(L)$ represents the power series or polynomial in the lag operator L . This infinite series, as written, is expressed in the open form, but certain series can be converted to the closed form containing only one term which will give equal results as with the open form.

The Koyck distributed lag function in open form is

$$(21) \quad Y_t = \beta(1-\lambda) (X_t + \lambda X_{t-1} + \lambda^2 X_{t-2} + \dots)$$

Using the lag operator notation, one can write

$$(22) \quad Y_t = \beta(1-\lambda) (1 + \lambda L + \lambda^2 L^2 + \dots) X_t.$$

The power series

$$W(L) = (1-\lambda) (1 + \lambda L + \lambda^2 L^2 + \dots)$$

can be written in closed form as

$$W(L) = \frac{(1-\lambda)}{(1-\lambda L)}$$

or in shorthand notation, the Koyck lag function is

$$(23) \quad Y_t = \beta W(L) X_t$$

or in closed form the equation becomes

$$(24) \quad Y_t = \beta \frac{(1-\lambda)}{(1-\lambda L)} X_t$$

Upon multiplying both sides of the equation by $(1-\lambda L)$, the equation is:

$$(1-\lambda L) Y_t = \beta(1-\lambda) X_t$$

$$Y_t - \lambda Y_{t-1} = \beta(1-\lambda) X_t$$

which is

$$(25) \quad Y_t = \lambda Y_{t-1} + \beta(1 - \lambda) X_t$$

the reduced equation derived above (equation 12).

The Pascal distribution expressed by using lag operator notation is

$$(26) \quad Y_t = \beta \left(\frac{1 - \lambda}{1 - \lambda L} \right)^r X_t = \beta \frac{(1 - \lambda)^r}{(1 - \lambda L)^r} X_t$$

which is essentially multiples of the Koyck distribution. The Pascal distribution has a restriction that the roots of the denominator of the closed form $W(L)$ must be equal, e.g., a second order Pascal distribution is written as

$$(27) \quad (1 - \lambda L)^2 Y_t = \beta X_t; \lambda_1 = \lambda_2 = \lambda$$

i.e., the roots are all equal. However, the rational distributed lag function, which Jorgenson (1966) developed, can have unequal roots and yet yield different lag configurations from the same basic distribution. The roots λ_1 and λ_2 do not have to be equal, e.g.,

$$(28) \quad (1 - \lambda_1 L) (1 - \lambda_2 L) Y_t = \beta X_t$$

The roots of the rational distributed lag function of any degree in the lag operator L do have certain other restrictions to give desirable properties to the lag distribution.⁶

Jorgenson pointed out that a large class of polynomials can be represented by the rational form

⁶See Bauer (1968, pp. 18-22) for a discussion of the restrictions on the roots. A second degree polynomial was used as an example. Also, note his reference to Griliches (1967, pp. 27-29) for a more detailed discussion of the restrictions.

$$(29) \quad W(L) = \frac{N(L)}{D(L)}$$

so that some complicated polynomials can be expressed as the ratio of two simpler polynomials. For example, one can assume for explanatory purposes that $N(L)$ and $D(L)$ are both second order polynomials in the lag operator L so that

$$(30) \quad W(L) = \frac{N(L)}{D(L)} = \frac{a_0 + aL + bL^2}{1 - cL - dL^2}$$

and

$$W(L) = w_0 + w_1L + w_2L^2 + \dots$$

This lag distribution can be imposed on

$$(31) \quad Y_t = \beta X_t$$

and the function written as

$$(32) \quad Y_t = \beta \frac{a_0 + aL + bL^2}{1 - cL - dL^2} X_t$$

Upon multiplying both sides of the equation by $1 - cL - dL^2$, one obtains

$$(33) \quad (1 - cL - dL^2) Y_t = \beta(a_0 + aL + bL^2) X_t$$

or

$$(34) \quad Y_t = cY_{t-1} + dY_{t-2} + \beta a_0 X_t + \beta a X_{t-1} + \beta b X_{t-2}$$

which is the estimating equation after adding a stochastic disturbance term to the exact form of the equation.

The estimation of the rational distributed lag function introduces or intensifies some problems in estimation in addition to those commonly found when doing economic research. Serial correlation of the disturbance is almost always present when lagged values of the

dependent variable are used as regressors. Also, multicollinearity among the contemporaneous and lagged values of the variables in the estimating equation is likely to be a problem. Another problem which complicates the estimation technique is overidentification. The estimates of the partial regression coefficients in the estimating equation are combinations of more than one parameter. Edwards (1962) demonstrated a technique of nonlinear estimation to get unique estimates of each of the parameters.⁷

The series of weights relating the effects of a change in X through time can be derived from the above estimating equation given the estimates of the lag parameters a_0 , a, b, c, and d. Bauer (1968, pp. 38-40) discussed the derivation of the lag weights for the above example of a rational polynomial consisting of the ratio of two second degree polynomials. Griliches (1967, pp. 23 and 24) also discussed the derivation of the weights.

Distributed Lag Functions Using Only Lagged Exogenous Variables

The third method introduces a number of lagged exogenous variables into a finite distributed lag function. Two alternatives of this method are discussed in this section. They are the Almon alternative (1965), using Lagrangian interpolation polynomials and the Johnson alternative (1969), using a general polynomial form.

Distributed lags using Lagrangian interpolation polynomials. Almon (1965) used only lagged exogenous variables in her study of the

⁷Bauer (1968, pp. 35-38) also explained this technique which was taken from Edwards (1962).

distributed lag between capital appropriations and expenditures in manufacturing industries.⁸ The method is similar to the first two methods discussed above in that it places restrictions on the coefficients of the model

$$(35) \quad Y_t = \sum_{i=0}^{n-1} w(i) X_{t-i}$$

where n is the number of X 's, i is the lag time, and $w(i)$ are the weights with which the several present and past values of X , the independent variable, are combined. However, this method does not rely on the assumption that the data are generated by infinite lag distributions which was assumed in the other two methods. Instead, Almon assumed that the coefficients of the above model are generated by a finite degree polynomial. No assumptions are made about the distribution of the weights, except that successive weights lie on a finite degree polynomial. Hence, the Almon method can generate a variety of lag configurations. The estimating equation can include other unlagged variables or other variables with distributed lags. Generally, the length of the distributed lag is not known prior to estimation, so the distribution is estimated using varying numbers of periods. Certain criteria, which will be discussed later, are used to determine the optimum number of periods to include in the lag distribution and the degree of the polynomial.

The method involves the estimation of a few points on a curve (the polynomial) as regression coefficients while using polynomial

⁸The theoretical discussion of this alternative draws heavily from Almon (1965, pp. 179-181).

interpolation to interpolate between the estimated points to obtain the remaining points. Or, more formally, distributed lags are specified and estimated as linear combinations of Lagrangian interpolation polynomials. Almon (1965, pp. 179-180) presented the following rationale and explanation for use of this technique:

The "interpolation distribution" assumes that the $w(i)$ are values at $x = 0, \dots, n - 1$ of a polynomial $w(x)$ of degree $q + 1$, $q < n$, where n is the number of periods over which the distributed lag extends. Its estimation is based on the fact that once $q + 2$ points on the curve are known-- $w(x_0) = b_0$, $w(x_1) = b_1, \dots, w(x_{q+1}) = b_{q+1}$ --all the $w(i)$ can be calculated as linear combinations of these known values. . . .

The $w(i)$ are calculated by

$$(36) \quad w(i) = \sum_{j=0}^{q+1} \phi_j(i) b_j \quad (i = 0, \dots, n - 1)$$

where the $\phi_j(i)$ are the values of $x = i$ of the Lagrangian interpolation polynomials:

$$\begin{aligned} \phi_0(x) &= \frac{(x - x_1)(x - x_2) \dots (x - x_{q+1})}{(x_0 - x_1)(x_0 - x_2) \dots (x_0 - x_{q+1})} \\ \phi_1(x) &= \frac{(x - x_0)(x - x_2) \dots (x - x_{q+1})}{(x_1 - x_0)(x_1 - x_2) \dots (x_1 - x_{q+1})} \\ &\vdots \\ \phi_{q+1}(x) &= \frac{(x - x_0)(x - x_1) \dots (x - x_q)}{(x_{q+1} - x_0)(x_{q+1} - x_1) \dots (x_{q+1} - x_q)} \end{aligned}$$

These polynomials have the property that:

$$\phi_j(x_j) = 1 \quad (j = 0, \dots, q+1)$$

$$\phi_j(x_k) = 0 \quad (j \neq k; j = 0, \dots, q+1; k = 0, \dots, q+1)$$

so that

$$(37) \quad w(x) = \sum_{j=0}^{q+1} \phi_j(x) b_j$$

which is a polynomial of degree $q + 1$ with the values b_j at points x_j . The model can specify zero weights before time 0 and after time $n - 1$, i.e., $w(-1) = w(n) = 0$ if $x_0 = -1$, $x_{q+1} = n$, and $b_0 = b_{q+1} = 0$ so that (37) becomes

$$(38) \quad w(i) = \sum_{j=1}^q \phi_j(i) b_j$$

The substitution of (38) into (35) yields

$$(39) \quad Y_t = \sum_{i=0}^{n-1} \left[\sum_{j=1}^q \phi_j(i) b_j \right] X_{t-i} = \sum_{j=1}^q b_j \sum_{i=0}^{n-1} \phi_j(i) X_{t-i}$$

Estimation of the b_j is accomplished by regressing the Y_t on the q variables $Z_{tj} = \sum_{i=0}^{n-1} \phi_j(i) X_{t-i}$, $j = 1, \dots, q$. Estimates of the b_j then are used in (38) to compute the weights of the distributed lag.

Prior to calculation of the interpolation polynomials and estimation of the b 's, the range of the magnitudes of q and n must be determined. The selection of q is a problem of model specification, so judgment is important concerning the most theoretically reasonable degree of the polynomial to use. In addition, the number of observations available can determine to some extent the number of parameter points q .

Almon (1965, p. 181) suggested that the magnitude of n be based on judgment relevant to the particular problem being analyzed. She suggested as an additional indicator for the number of lagged exogenous variables to include in the model, the calculation of simple correlation coefficients between the dependent variable and successive lagged values of the exogenous variable. The time period (Almon, p. 181) in which the lagged exogenous variable explains the endogenous variable about as well as the exogenous variable with no lag at all can serve as the

center of a range of n 's, i.e., a number of lags of different lengths then are estimated and the best lag is chosen from among the alternatives. The selection of the best or optimal length of lag is based on judgment, $\bar{R}^2$, the signs of the calculated weights, and the close similarity of the lag weights among distributed lags of optimal and longer length.

The major advantages of the Almon method are flexibility of the lag configuration and ease of estimation. Also, additional serial correlation of the disturbance⁹ is not induced when using this method of reducing the lag parameter space, but such is not the case for the other methods discussed above.¹⁰ Furthermore, as Almon (1965, pp. 187-188) pointed out, serial correlation of the disturbance, even if present initially in the distributed lag function, is not as serious a problem as with the first and second methods discussed above, because serial correlation here does not mean biased estimates. Serial correlation in the initial lag function (before transformation to a smaller parameter space) only means an understatement of the standard errors of the estimated regression coefficients (assuming positive serial correlation of the disturbance which is the usual case in economic data). Because the distributed lag weights are derived using these estimated coefficients, positive serial correlation leads

⁹See Griliches (1967, pp. 33-42) for a discussion of serial correlation of the disturbance in the Koyck-type distributed lag model and its generalizations.

¹⁰Nerlove's partial adjustment model (not specifically discussed above) is an exception. It does not induce additional serial correlation if there is none initially, according to Griliches (1967, p. 17). See Nerlove (1956 and 1958).

indirectly to an understatement of the true standard errors of the lag weights. The Durbin-Watson statistic can be used as an indicator of serial correlation of the disturbance when using the Almon method since this method does not bias the statistic toward unity, whereas the other two methods do bias it toward unity. One other possible advantage comes from the fact that economic data are generated more than likely from a finite lag rather than an infinite lag. The Almon method specifies a finite lag so it avoids the possibility of misleading results caused by fitting an infinite lag to the data.

The main disadvantages of the Almon method include the possibility of misleading results from estimating a finite distribution that is too short relative to the true distribution because of the lack of data over enough periods of time. Misleading results caused by specification error could arise under the following conditions: the researcher is unsure a priori of the length of the lag and the number of time periods available in the data is not sufficiently large enough so that the use of a finite polynomial, constrained by the number of periods of available data, could result in an underestimation of the true length of the lag.

Another disadvantage is the lack of adequate criteria for choice of points of truncation, i.e., the number of periods to lag the exogenous variable. Chetty (1971, pp. 99 and 109) criticized the Almon procedure for making the choice of points of truncation to be included in the model. He indicated that different points of truncation give different distributed lag functions. He claimed that the choice of truncation points based on simple correlations between the dependent and lagged independent variables is not very useful because reliable inferences

about partial regression coefficients cannot be made using simple correlation coefficients.¹¹

Johnson (1969, p. 8) criticized the Almon method of using Lagrangian interpolation polynomials because it lacks conceptual clarity as an econometric tool. He suggested as an equivalent, but more straightforward alternative,¹² the use of general linear restrictions, a more familiar tool in econometrics, as embodied in a general polynomial form instead of Lagrangian polynomials.

The general polynomial form in distributed lag analysis. This study used the concept of general linear restrictions as incorporated in the general polynomial form (Johnson, 1969) to restrict the parameter space of the distributed lag functions hypothesized to exist in the equation system. The technique¹³ is explained first by using a general single equation with a distributed lag on the independent variable. Then in Chapter V the technique is explained and applied to a simultaneous system containing both lagged and nonlagged exogenous variables.

A distributed lag function can be written as

$$(40) \quad Y_t = \sum_{s=0}^{n-1} w(s) X_{t-s}$$

where the dependent variable Y is a function of a finite lag structure

¹¹Chetty (1971) presented an alternative approach using a Bayesian analysis to arrive at various truncation points and to choose the optimum lag distribution.

¹²See Johnson (1969, pp. 5-8) for a demonstration of the equivalency of the two approaches.

¹³The discussion on this method follows closely the explanation given by Johnson (1969).

of the independent variable X . The equation as written contains a lag structure with n lag weights $w(s)$. One can suppose that for some reason (such as discussed above, i.e., multicollinearity among successively lagged values of X , a relatively small number of observations compared to a relatively large number of weights, etc.) that it is desirable to place some restriction on the model by reducing the number of parameters to be estimated by using a polynomial to represent the original lag structure. Thus, the lag structure represented by the polynomial contains fewer parameters to be estimated than the original equation. The lag weights $w(s)$ can be represented by

$$(41) \quad w(s) = a_0 + a_1 s + a_2 s^2 + \dots + a_{p-1} s^{p-1}$$

which is a $p-1$ degree polynomial containing p a 's where $p < n$ and s is the lag time. If one assumes that s takes only integer values, equation (41) in matrix notation¹⁴ becomes

$$(42) \quad \vec{w} = \vec{T} \vec{a}$$

or

$$\begin{bmatrix} w(0) \\ w(1) \\ w(2) \\ . \\ . \\ . \\ w(n-1) \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 4 & \dots & 2^{(p-1)} \\ . & . & . & & . \\ . & . & . & & . \\ . & . & . & & . \\ 1 & (n-1) & (n-1)^2 & \dots & (n-1)^{(p-1)} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ . \\ . \\ . \\ a_{p-1} \end{bmatrix}$$

¹⁴ Note that throughout this paper a lower case letter represents a vector and an upper case, a matrix, when the letter has the symbol ($\rightarrow$) above it.

where $\vec{T}$ is a transformation matrix of the lag time s . Equation (40) in expanded form is

$$(43) \quad Y_t = w(0)X_t + w(1)X_{t-1} + \dots + w(n-1)X_{t-(n-1)}$$

By using the general polynomial form for $w(s)$ (equation 41), equation (43) can be written as

$$(44) \quad \begin{aligned} Y_t = & a_0 X_t + [a_0 + a_1 + a_2 + \dots + a_{p-1}] X_{t-1} \\ & + [a_0 + 2a_1 + 4a_2 + \dots + 2^{(p-1)} a_{p-1}] X_{t-2} \\ & + \dots \\ & + [a_0 + (n-1)a_1 + (n-1)^2 a_2 + \dots + (n-1)^{(p-1)} a_{p-1}] X_{t-(n-1)} \end{aligned}$$

Upon rearranging terms, the equation becomes

$$(45) \quad \begin{aligned} Y_t = & a_0 [X_t + X_{t-1} + X_{t-2} + \dots + X_{t-(n-1)}] \\ & + a_1 [X_{t-1} + 2X_{t-2} + 3X_{t-3} + \dots + (n-1) X_{t-(n-1)}] \\ & + a_2 [X_{t-1} + 4X_{t-2} + 9X_{t-3} + \dots + (n-1)^2 X_{t-(n-1)}] \\ & + \dots \\ & + a_{p-1} [X_{t-1} + 2^{(p-1)} X_{t-2} + 3^{(p-1)} X_{t-3} + \dots + (n-1)^{(p-1)} X_{t-(n-1)}] \end{aligned}$$

or

$$(46) \quad Y_t = a_0 V_{t0} + a_1 V_{t1} + a_2 V_{t2} + \dots + a_{p-1} V_{t,p-1}$$

where

$$\begin{aligned} V_{t0} &= X_t + X_{t-1} + X_{t-2} + \dots + X_{t-(n-1)} \\ V_{t1} &= X_{t-1} + 2X_{t-2} + 3X_{t-3} + \dots + (n-1)X_{t-(n-1)} \\ V_{t2} &= X_{t-1} + 4X_{t-2} + 9X_{t-3} + \dots + (n-1)^2 X_{t-(n-1)} \\ &\vdots \\ &\vdots \\ &\vdots \end{aligned}$$

$$V_{t,p-1} = X_{t-1} + 2^{(p-1)}X_{t-2} + 3^{(p-1)}X_{t-3} + \dots + (n-1)^{(p-1)}X_{t-(n-1)}$$

Using matrix notation, one gets

$$(47) \quad \vec{V} = \vec{X} \vec{T}$$

where $\vec{V} = [V_{t0}, V_{t1}, V_{t2}, \dots, V_{t,(p-1)}]$

$$\vec{X} = [X_t, X_{t-1}, X_{t-2}, \dots, X_{t-(n-1)}]$$

and $\vec{T}$ is as defined above.

With

$$(48) \quad \vec{y} = \vec{X} \vec{w}$$

being the matrix form of equation (40) and given the matrix form of equation (41), then

$$(49) \quad \vec{y} = \vec{X} \vec{T} \vec{a}$$

or

$$(50) \quad \vec{y} = \vec{V} \vec{a} \text{ when } \vec{V} = \vec{X} \vec{T}$$

Equation (50) becomes the estimating equation upon addition of a disturbance term.

Johnson (1969, p. 4) emphasized some of the advantages of using a general, finite polynomial form. They include flexibility of the lag configuration over the finite period and conceptual and computational clarity of the transformation of the original lag function from the n period lag to the $p-1$ degree polynomial. A major advantage is ease of estimation when compared to the other methods. Other advantages include those discussed above for the Almon method.

Disadvantages of the Johnson method include the same ones discussed above concerning the Almon approach. One other disadvantage when compared with the Almon technique is the increased difficulty

of incorporating zero restrictions on some of the lag weights. The general polynomial form with zero restrictions on some of the weights must be estimated using restricted least squares, whereas with the Almon method, the zero restrictions are incorporated directly into the calculation of the Lagrangian interpolation polynomials prior to estimation.¹⁵

The choice of the range of n and p to estimate initially can be based in general on the same criteria as those used by Almon (1965). Selection of optimum n and p also can depend on the criteria used by Almon in her selection of the optimum distribution. The selection of the initial range of n and p and the selection criteria for the optimum distribution will be discussed in Chapter V.

III. MAJOR CHOICE CRITERIA FOR USING THE GENERAL POLYNOMIAL FORM OF DISTRIBUTED LAGS

The choice of the general polynomial form for restricting the parameter space of the distributed lag functions in this study was based to a large extent on considerations relating to the relative advantages and disadvantages of the major methods used in economic research as reviewed above. In essence, the choice was made based on the evaluation of the relative aspects of each method reviewed above and particularly on the following criteria:

1. Conceptual clarity. The use of the general polynomial form to transform the original distributed lag and thereby reduce the parameter space resulted in an estimating equation that was relatively

¹⁵See Johnson (1969, p. 8) and Almon (1965, p. 180).

straightforward as demonstrated above. Although the concept of linear restrictions appears familiar in economics, whether or not the concept was well related to behavioral theory in this problem setting was another question. Bauer (1968, pp. 13-15) emphasized the lack of behavioral theory in economics that relates to distributed lag analysis when applied to particular economic problem settings. Nevertheless, the assumption was made that the general polynomial form was as well related to economic behavioral theory of the problem under investigation as any other available distributed lag model.

2. Relative ease of estimation. No problems of nonlinearity, overidentification, and induced serial correlation of the disturbance arise as a result of using the general polynomial method of reducing the parameter space. Multicollinearity of successively lagged values of X could be a problem, but it also could be a problem with some of the other methods. Therefore, if the assumptions of least squares are met for a single equation system, then the ordinary least-squares estimation method is appropriate when no restrictions on the signs of the lag weights are imposed. The two-stage least-squares method is an appropriate estimation method for an overidentified simultaneous system assuming that no initial restrictions on the signs of the lag weights are imposed. Imposition of restrictions such that each lag weight has to be either zero or positive ($w_1 \geq 0$) would be an application of quadratic programming.¹⁶

3. Flexibility or variety of lag patterns. A monotonically

¹⁶Johnson (1969, p. 4).

declining geometric lag was ruled out because the patterns of the distributions of the distributed lag effects of technology on both the demand for farm labor and the supply of aggregate farm product were not necessarily expected to decline monotonically from the first. Intuitively, one would think that the maximum effect of a change in technology would not occur initially, but would occur sometime after the initial time period. Results obtained by Bauer (1968, pp. 46-51), where a rational polynomial form was used, showed that the distribution of the effect of technology on the demand for farm labor had a "hump" after the initial time period, i.e., the maximum effect occurred after the initial time period. The configuration was not monotonically declining throughout. Thus, a method for allowing a variety of lag patterns was needed. The Pascal distribution was ruled out initially because of nonlinear estimation problems. An attempt was made to use the rational polynomial method in the simultaneous equation framework of this study with a rational polynomial form of distributed lag on technology in both the farm labor demand equation and the aggregate product supply equation. Problems quickly developed because no algebraic solution for each individual lag parameter estimator could be obtained given the information available in the reduced form estimating equations, i.e., the estimators for the reduced form parameters could not be algebraically manipulated to arrive at derived estimators of the individual lag parameters. Therefore, the rational polynomial form was eliminated, given the specified model. Another specified model such as a recursive or bloc recursive model (with one rational polynomial distributed lag in each bloc) could have been

estimated and algebraically derived estimates of the individual lag parameters could have been obtained using rational distributed lags. Also, rational distributed lags could have been used if the specified structural model had assumed the same invariant lagged effect of technology on both the demand for labor and the supply of aggregate product, i.e., corresponding estimated lag parameters would be equal in both distributions. However, such specifications could not be justified on the basis of economic theory.

4. State data on the proxy variable, research and extension expenditures, were available over several years (1889-1964). Therefore, use of the general polynomial form for restricting the parameter space was assumed not to be constrained by lack of data over a period sufficiently long enough to represent the true lag.

CHAPTER IV

SOME STATISTICAL AND ECONOMIC CONSIDERATIONS IN FORMULATING THE STATISTICAL MODEL

Once the economic model was specified, the task of locating or developing the data was undertaken. This phase of the research process required many decisions regarding the specific nature of the variables and relations to be employed in addition to those discussed in Chapters II and III. The end result of this phase of the research was the statistical model listed in Chapter V.¹ In deciding upon this particular statistical model, many decisions had to be made, some of which are discussed in this chapter.²

I. THE LEVEL OF AGGREGATION OF COMMODITIES

The level of aggregation of farm commodities was at the all-commodity level which included all crop and livestock products in one measure for each state and each year. This level was chosen partly out of convenience, but mostly out of necessity because data for such variables as price indices and the value aggregate, cash receipts of farm marketings, were classified by states³ over a period

¹A description of the variables in the statistical model is presented in Appendix A, along with some specific conceptual and statistical considerations for each variable.

²Some of these considerations were discussed by Foote (1958, pp. 19-52). Part of the discussion in this chapter includes some of his ideas.

³Some states were grouped as explained in Appendix A.

of several years. Also, data on farm wage rates, labor quantity, etc., were classified by state and not by commodity or groups of commodities within state. Finally, the use of a level of aggregation over individual commodities or commodity subgroups would have required the addition of more equations to the system.

II. CHOICE OF SPATIAL UNIT

To examine the distributed lag effect of technology on the farm labor market at the industry level required choosing a spatial unit that would represent the industry level. The industry level for farming in the United States probably is, in a strict conceptual sense, the national level for each commodity. However, differences in labor and product market behavior was expected to exist between various farming regions and states of the country. Data used to represent the variables in the model were for the most part collected at a substate level or state level. Regional and national measures were obtained by aggregating state data.

National data were derived for the most part from state data, where the states were not homogenous, so considerable aggregation error was expected to be present in the national data. Use of national data containing aggregation error could have adversely affected the reliability of the results of the statistical analysis. Furthermore, accounting for the possible locational differences was judged to be very important conceptually. Aside from being conceptually interesting, accounting for these expected differences could remove an important source of variation in the residuals of the statistical model.

Furthermore, reliable national data on the variables of the economic model, like state data, did not cover a large number of years. The number of observations would have been quite limited with an annual time series using national data. Therefore, to account for locational differences and to obtain a much greater number of observations while still retaining the level of observation at the "industry" level, the decision was made to use state data in a time series of cross sections.

The level of analysis was judged to be at the industry⁴ level for farming in a state. However, the level was a subindustry classification in a strict, functional sense, because the economic behavior of farmers as suppliers or consumers as demanders for the most part ignores the geographical-political boundaries of the state, i.e., state boundaries, in general, do not delineate functional economic areas with respect to the "farming industry." Nevertheless, state data were used because data were available by state and not by functional economic areas.

III. CHOICE OF THE TIME UNIT AND YEARS INCLUDED

Measurement of time units on a calendar year basis was chosen as the most appropriate for the purposes of this study, given the level of aggregation across commodities, including both seasonal and non-seasonal commodities. Utilization of calendar year units required the assumption that economic conditions on the average within the year

⁴The output of the "industry" was assumed to be aggregate farm product.

were sufficiently homogenous for all farm commodities as a group.⁵ Monthly data did not exist for most variables. Furthermore, because the model attempted to account for the effect of changes in technology (a structural change), annual data probably were more appropriate given the planning horizon of farmers and their manner of adoption of new technology. The existing level of technology for an individual farmer does not fluctuate greatly from month to month, but from year to year (not necessarily calendar year), or even longer periods. The farmer adopts new technology either involuntarily or voluntarily in a discontinuous manner over a planning horizon that usually includes many years. Consequently, using time units smaller than annual units in examining the effects of changes in technology on the demand for labor seemed irrelevant for the purpose at hand.

The possibility of the presence of discontinuities in adoption of new technology at the firm and industry levels served as a reminder that in using annual continuous data on a proxy variable for technology tends to smooth out what may be a discontinuous function. However, a continuous function was assumed, given the time period covered and the levels of aggregation with respect to commodities and spatial units.

Data were available from 1949 through 1964 for all variables in the statistical model except the proxy variable for technology, public research and extension expenditures. Data on this proxy were available from 1889 to 1964. Additional recent years would have been used, but data on one or more of the variables were not available after 1964.

⁵ Individually, this may not have been the case, especially for seasonal commodities like truck crops.

Data on years prior to 1949 for some of the variables other than the technology proxy were unreliable.

Data on the proxy variable for technology covered many years in which the structure of agriculture could have changed for other than technological reasons. However, to examine the distributed lag effect of technology on the demand for farm labor using the method which this study utilized, required observations on research and extension expenditures for a large number of years prior to 1949. The proxy for technology, which included observations during certain years of significant economic upheaval such as the Great Depression, the New Deal, and World War II, included variation caused by structural changes other than those considered in the model. Nevertheless, the decision was made to include observations on these years for the variable. The structures of the farm labor and product markets were judged to be fairly constant from 1949 to 1964 except for the changes which the model attempted to measure.

IV. FUNCTIONAL FORM

The choice of the functional form was based primarily on judgment and to a lesser extent on conventions of past studies and statistical considerations. The relationships between the variables of the demand and supply equations of the model were felt to be curvilinear rather than linear. Because the relationships in the model were felt to be multiplicative rather than additive, each of the four equations of the model was transformed to a logarithmic form.

Footnote (1958, p. 37) pointed out some statistical considerations

in utilizing logarithmic equations. He stated:

From a statistical viewpoint, logarithmic equations should be used when (1) the relationships between the variables are believed to be multiplicative rather than additive, (2) the relationships are believed to be more stable in percentage than in absolute terms, and (3) the unexplained residuals are believed to be more uniform over the range of the independent variables when expressed in percentage rather than absolute terms. . . .

The three conditions were assumed to hold for the model of this study.

Using a logarithmic form in the model which utilized data in a time series of cross sections was felt to be superior over other functional forms from a statistical standpoint.

Past studies on technology and the farm labor market such as Bauer (1968) and Wallace and Hoover (1966) employed logarithmic relationships. Utilization of the same form in this study allowed possible comparison of results with those of Bauer (1968) and Wallace and Hoover (1966).

V. MEASURING TECHNOLOGY

A major difficulty that was central to the statistical and economic analyses of the effects of technology on the demand for farm labor was the problem of empirically measuring technology. The measurement problem originated primarily because technology at the prevailing level of knowledge is not easily quantified. Various attempts have been made to quantify and measure it. At the same time controversies have arisen over the advantages and disadvantages of the different measures. Schuh (1966, pp. 74-83) discussed some of the conceptual difficulties underlying some of the various empirical approaches used to measure technology. The various approaches and their

corresponding advantages and disadvantages were not elaborated upon here.

Expenditure on research and extension was employed as a proxy variable for technology in this study as an attempt to estimate the effect of technology directly as a causal variable in the labor demand and product supply equations rather than ascribing the unexplained residual to technical change, or by representing technical change with a trend variable. Initially, an attempt was made to account, at least partially, for embodied technical change in the labor input by accounting for changes in labor quality using education and age along with research and extension expenditures. However, as demonstrated by the statistical model in the following chapter, education and age of the farm working force were eliminated because of lack of data. Yet, to the extent that changes in research and extension expenditures contribute to changes in labor quality through expenditures on extension education and to changes in quality of all inputs through research, embodied technology was allowed for partially. The use of this proxy also could have provided for the possibility of a portion of technological change that is not embedded in the quality of the inputs, such as improved organization or arrangement of inputs that increase productivity of the same inputs even after quality changes are explained. Also, public expenditures on research and extension were used to measure technology because these expenditures are generally assumed to be related to the productivity of labor according to Wallace and Hoover (1966, p. 326).

Some problems arose in using this measure of technology. For example, use of only public expenditures fails to account for private

expenditures on research and extension activities. Such activities by farm machinery and equipment manufacturers and dealers; chemical, seed, and fertilizer producers; farmers; and others are becoming increasingly important, especially for commercial farms.

Another problem with the measure was related to the concept of distributed lags. The use of research and extension expenditures as a proxy for technology raised questions of distinguishing between potential and available technology. That is, at the time funds are expended for research and extension, new technology logically does not become available instantaneously. Expenditure of funds for research and extension in the current time period provides only the potential for new technology in the current time period, while technology actually becomes available in future time periods. Hence, the use of expenditures on research and extension as a proxy for available technology includes the time lag between the time that a potential exists as a result of expending funds and the point in time when technology becomes operationally available to farmers. In this case, the length of the lag of the effect of available technology could be overstated because the interval of time when potential technology is becoming available technology is also included.

The problem of distinguishing between potential and available technology was, in actuality, a subset of another more general problem. The time lag between expenditure of funds for research and extension activities and actual effect of new technologies in the farm labor and product markets is the result of several lags, each of which could have its own lag distribution. Consequently, the lag distribution

imposed on technology in the labor demand equation was a composite or convolution of several lag distributions. The same applied to the lag distribution used to represent the total lag between expending of funds and their effects in the product supply equation.

The most serious drawback to using research and extension expenditures to represent technology was not of a conceptual origin, but a statistical problem. For example, those states which have a large aggregate output of agricultural products also have large expenditures on research and extension, while states with small outputs have relatively small expenditures on research and extension.⁶ This high correlation between output and expenditures could imply on the surface that output is highly affected by expenditures on research and extension (or technology). An ordinary least-squares regression of output on research and extension expenditures, among other things, probably would indicate a significant effect of "technology" on output. However, the level of research and extension expenditures in a state is, in turn, a function of the agricultural output of the state. The allocation of funds to research and extension in a state, regardless of the origin of the funds, is influenced by the size of the agricultural industry and hence its output over time. Thus, the problem of simultaneous equation bias occurred in this study as a result of using expenditures on research and extension as a proxy variable for

⁶ Furthermore, the formulas for allocating federal research funds to the states are rather inflexible because allocation is based on such state characteristics as number of farms and rural population as a percentage of the total population. These characteristics may or may not be related to technology in the state.

technology, while not adding an equation that specified current research and extension expenditures as a function of current and lagged output of the state, among other things.

The additional equation was not utilized because of additional formidable equation problems. Adding the equation would have made the technology variable an endogenous or jointly determined variable in the system. Conceptually, this is not unreasonable, but from an econometric standpoint, it presented a serious problem since distributed lags on technology was estimated. The method of reducing the parameter space of the general distributed lag function to arrive at estimated values of the jointly determined variables in the reduced form, prevented the specification and estimation of distributed lags on jointly determined variables in the system. Appendix B explains the problem. Consequently, the decision was made to abandon the specification of a seventh equation in the system which would have accounted for the effect of output (or deflated cash receipts), among other things, on research and extension expenditures.

Another measure of technology was considered, which was total output divided by total input. However, lack of data forced the abandonment of this measure. Productivity indices did not exist by state such as the index calculated for United States agriculture (U.S. Department of Agriculture, 1949-1970, specifically 1970, Table 649). To calculate a similar index by state would have been prohibitively costly from a computational standpoint.

VI. SOME DIFFICULTIES WITH INDICES AND DEFLATORS IN THE TIME SERIES OF CROSS SECTIONS

Outside of the usual cautions and shortcomings surrounding the use of index numbers,⁷ the utilization of price indices in their currently available form in the time series of cross sections presented an addition difficulty. For example, the price index of aggregate farm product (Appendix A) either was computed with a base period of 1957-59 = 100 or was converted to that base period. Having the prices in all states in the cross sections on an equal base at 1957-59 = 100 arbitrarily removed some of the cross-sectional variation in prices. The indices accounted for differences in prices over time, but failed to account for the total spatial differences in prices among states, since the 1957-59 average price for all states was 100.

A closely related problem occurred from utilizing various national price indices to deflate state data. National indices were utilized out of necessity since state indices did not exist for the most part. Consequently, using the national indices could have reduced arbitrarily the cross-sectional variation with respect to the particular deflator. Thus, the deflated variables incorporated measurement errors where the national deflators failed to adequately reflect price conditions in the individual state. Use of the national deflators arbitrarily set state differences with respect to the deflator at zero. Realistically, this would not be the case for comparison between most states. Nevertheless,

⁷ Additional specific problems for each variable are discussed in Appendix A.

lack of data on deflators at the state level prevented the use of more suitable state deflators.

VII. CHOICE OF VARIABLES IN THE STATISTICAL MODEL

The statistical model in Chapter V was developed from the conceptual model of Chapter II. The statistical model was somewhat less inclusive than the conceptual model because, in arriving at a working model, some compromises had to be made. The compromises primarily involved the elimination of some of the variables of the conceptual model. Some variables that were originally in the conceptual model were eliminated for one or more of the following reasons:

1. Lack of data in a suitable form over the required number of years prevented the inclusion of some continuous variables in the model, e.g., average age and education of the farm working force and of farmers, urban and rural unemployment rates, rural population, rural to urban migration, rural and urban welfare payments, prices of non-farm commodities, level of mechanization, tastes and preferences, and weather.

2. Zero-one dummy variables for estimating intercept differences between state locations for each equation in the model were not included because of limitation of computer capacity when using the available two-stage least-squares program (Stroud, Zellner, and Chau, 1963, revised by Thornber and Zellner, 1965). Furthermore, a sufficient number of annual observations for each state did not exist, so that the parameters for each state could be estimated separately and then compared. Consequently, possible state differences in both the

intercept and slope coefficients were ignored empirically as causal variables.

Even though state intercept differences were not measured, regional intercept dummies were included so that regional intercept differences could be estimated. Use of regional dummies did not exceed the computer capacity of a maximum of 30 exogenous (or "independent") variables in each reduced-form equation using the two-stage least-squares program. Appendix C lists the regions and the states in each region.

3. Multicollinearity was expected to be a serious problem in the specified statistical model, but even more serious if all the variables of the conceptual model had been included. However, this expectation was not empirically verified prior to deletion of the variables. It was merely a judgment of the researcher.

4. Compromises were also necessary regarding the form of certain variables. Appendix A elaborates on this problem for some of the variables. An important example was the prices of the nonlabor inputs, land and other nonlabor inputs, in the conceptual model. Instead of utilizing price indices by state for land and other nonlabor inputs, deflated value aggregate data were used as shown in Appendix A. Land stock in thousands of dollars of value was the value aggregate used for land, while other expenditures in millions of dollars of value was used for all nonlabor inputs other than land.

CHAPTER V

THE STATISTICAL MODEL AND PROCEDURES FOR ESTIMATION

The processes of formulating the problem, reviewing the literature, and formulating the conceptual model, along with the making of some necessary decisions to formulate the statistical model were elaborated upon in previous chapters. This chapter presents the statistical model, the application of the general polynomial form to reduce the parameter space of the distributed lag functions, the estimation method, and selection of the degrees of the general polynomials and points of truncation.

I. THE STRUCTURAL MODEL

The general form of the structural model consisted of four¹ structural equations, two of them representing the supply and demand of farm labor and the other two representing the supply and demand of aggregate farm product. They were the following as expressed in the general functional form:

Demand for farm labor:

$$(51) \quad Q_{it} = a (W_{it}, P_{it}, S_{it}, O_{it}, R_{it}, \dots, R_{i,t-s}, D_{lit}, \dots, D_{6it})$$

Supply of farm labor:

$$(52) \quad Q_{it} = b (W_{it}, \bar{W}_{it}, U_{it}, D_{lit}, \dots, D_{6it})$$

¹Not including the equilibrium identities, which were assumed to be satisfied. The model including the identities was a six-equation model.

Demand for aggregate farm product:

$$(53) \quad Y_{it} = c(P_{it}, N_{it}, I_{it}, H_{it}, D_{1it}, \dots, D_{6it})$$

Supply of aggregate farm product:

$$(54) \quad Y_{it} = d(P_{it}, W_{it}, S_{it}, O_{it}, G_{it}, H_{it}, R_{it}, \dots, R_{1,t-s}, \\ D_{1it}, \dots, D_{6it})$$

where:²

Q_{it} = quantity of farm labor in thousands of man-days

W_{it} = farm wage rate in 1957-1959 dollars per hour

P_{it} = farm-level price index of aggregate product (1957-59 = 100)

S_{it} = land stock in thousands of 1957-1959 dollars

O_{it} = other expenditures in millions of 1957-1959 dollars

$R_{it} \dots R_{1,t-s}$ = current ($s = 0$) and lagged ($s = 1 \dots n - 1$ or $n' - 1$) technology in 1957-1959 dollars, where $s =$ lag time and n (or n') = number of lag weights on R .³

D_{mit} = dummy variable for the m -th regional location ($m = 1 \dots 6$ regions)⁴

² Appendix A gives a complete description of the data sources and construction of variables for all the continuous variables along with some specific problems encountered. Appendix C identifies the regional locations and states contained in each region.

³ Primes (') were used to indicate the number of lag weights on R in the product supply equation.

⁴ Estimation of slope coefficient differences by state was not attempted because of the increased complexity of such a model specification in conjunction with distributed lags. Also, state intercept differences were not estimated because the required number of exogenous variables would have exceeded the computer program limitation of a maximum of 30 "independent" variables (Stroud, Zellner, and Chau, 1963, revised by Thornber and Zellner, 1965, p. 9). However the states were

$\bar{W}_{it}$ = alternative wage rate in 1957-1959 dollars per hour

U_{it} = rate of unemployment (a percentage)

Y_{it} = quantity of aggregate farm product in millions of 1957-1959 dollars

N_{it} = population in thousands of persons

I_{it} = disposable income in millions of 1957-1959 dollars

H_{it} = home consumption in millions of 1957-1959 dollars

G_{it} = government payments in millions of 1957-1959 dollars

Observations in the time series of cross sections were represented by the subscripts:

$i = 1 \dots 39$ states or groups of states

$t =$ time, for the years 1949 ... 1964 for all variables in the model except R_{it} where the years covered were 1889 ... 1964.

The variables Q_{it} , W_{it} , P_{it} , and Y_{it} were jointly determined and all others were assumed to be exogenous or predetermined.

More specifically, the stochastic structural model, after transformation of the continuous variables⁵ to logarithmic form and addition of disturbance terms, was the following:

grouped by regions which were equal to those regions designated by the United States Department of Agriculture (1970a, Table 9) except for the Western Region. (Alaska and Hawaii were excluded from this region.) The states included in each region were specified in Appendix C. Utilization of regional grouping of the states allowed the estimation of regional intercept differences while staying within the limitations of the computer program.

⁵The zero-one dummy variables were not transformed to logarithms. The South Central regional dummy ($m = 6$) was eliminated to avoid singularity in the matrix of sums of squares and cross products of the "independent" variables.

$$(55) Q_{it} = \alpha_0 + \sum_{m=1}^5 \alpha_m D_{mit} + \alpha_6 W_{it} + \alpha_7 P_{it} + \alpha_8 S_{it} + \alpha_9 O_{it} \\ + \sum_{s=0}^{n-1} w_1(s) R_{i,t-s} + \epsilon_{1it}$$

$$(56) Q_{it} = \beta_0 + \sum_{m=1}^5 \beta_m D_{mit} + \beta_6 W_{it} + \beta_7 \bar{W}_{it} + \beta_8 U_{it} + \epsilon_{2it}$$

$$(57) Y_{it} = \gamma_0 + \sum_{m=1}^5 \gamma_m D_{mit} + \gamma_6 P_{it} + \gamma_7 N_{it} + \gamma_8 I_{it} + \gamma_9 H_{it} + \epsilon_{3it}$$

$$(58) Y_{it} = \delta_0 + \sum_{m=1}^5 \delta_m D_{mit} + \delta_6 P_{it} + \delta_7 W_{it} + \delta_8 S_{it} + \delta_9 O_{it} + \delta_{10} G_{it} \\ + \delta_{11} H_{it} + \sum_{s=0}^{n'-1} w_2(s) R_{i,t-s} + \epsilon_{4it}$$

where ϵ_{jit} represented the error or disturbance term for the j -th equation ($j = 1 \dots 4$), variables were defined as previously, and

$$(59) \sum_{s=0}^{n-1} w_1(s) R_{i,t-s} = w_1(0)R_{it} + w_1(1)R_{i,t-1} + \dots + w_1(n-1)R_{i,t-(n-1)}$$

$$(60) \sum_{s=0}^{n'-1} w_2(s) R_{i,t-s} = w_2(0)R_{it} + \dots + w_2(n'-1)R_{i,t-(n'-1)}$$

where the w 's represented the lag weights and

s = the lag time

$n-1$ (or $n'-1$) = selected number of lagged weights on R relative to the t -th observation

$i = 1 \dots 39$ states or groups of states

$t = 1 \dots T$ years (1889 ... 1964 for the technology variable).

The lag weights of w_1 and w_2 were represented by a $p-1$ and $p'-1$ degree polynomial with only p and p' parameters, respectively, i.e.,

$$(61) w_1(s) = a_0 + a_1 s + a_2 s^2 + \dots + a_{p-1} s^{p-1} \text{ or } p \text{ a's and } p < n-1$$

$$(62)^6 w_2(s) = b_0 + b_1 s + b_2 s^2 + \dots + b_{p'-1} s^{p'-1} \text{ or } p' \text{ } b\text{'s and } p' < n'-1$$

Then,

$$w_1(0) = a_0$$

$$w_1(1) = a_0 + a_1 + a_2 + \dots + a_{p-1}$$

$$w_1(2) = a_0 + 2a_1 + 4a_2 + \dots + (2)^{p-1} a_{p-1}$$

⋮

$$w_1(n-1) = a_0 + (n-1) a_1 + (n-1)^2 a_2 + \dots + (n-1)^{p-1} a_{p-1}$$

and

$$w_2(0) = b_0$$

$$w_2(1) = b_0 + b_1 + b_2 + \dots + b_{p'-1}$$

$$w_2(2) = b_0 + 2b_1 + 4b_2 + \dots + (2)^{p'-1} b_{p'-1}$$

⋮

$$w_2(n'-1) = b_0 + (n'-1) b_1 + (n'-1)^2 b_2 + \dots + (n'-1)^{p'-1} b_{p'-1}$$

In matrix notation,

$$(63) \begin{bmatrix} w_1(0) \\ w_1(1) \\ w_1(2) \\ \vdots \\ w_1(n-1) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 4 & \dots & (2)^{p-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & (n-1) & (n-1)^2 & \dots & (n-1)^{p-1} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_{p-1} \end{bmatrix}$$

⁶The magnitude of p in (61) did not necessarily have to equal the magnitude of p' in (62). The same applied to n and n' . Primes (') were used over p and n in the general polynomial for the distributed lag function in the product supply equation.

or

$$(64) \quad \vec{w}_1 = \vec{T} \vec{a}$$

and similarly for $\vec{w}_2(s)$.

Upon multiplying $w_1(s)$ and $w_2(s)$ each by $R_{i,t-s}$ and summing each product, i.e., $\sum_{s=0}^{n-1} w_1(s)R_{i,t-s}$ and $\sum_{s=0}^{n'-1} w_2(s)R_{i,t-s}$, the model in

expanded form was:

$$(65) \quad Q_{it} = \alpha_0 + \sum_{m=1}^5 \alpha_m D_{mit} + \alpha_6 W_{it} + \alpha_7 P_{it} + \alpha_8 S_{it} + \alpha_9 O_{it} \\ + a_0 [R_{it} + R_{i,t-1} + R_{i,t-2} + \dots + R_{i,t-(n-1)}] \\ + a_1 [R_{i,t-1} + 2R_{i,t-2} + 3R_{i,t-3} + \dots + (n-1)R_{i,t-(n-1)}] \\ + a_2 [R_{i,t-1} + 4R_{i,t-2} + 9R_{i,t-3} + \dots + (n-1)^2 R_{i,t-(n-1)}] \\ \vdots \\ + a_{p-1} [R_{i,t-1} + (2)^{p-1} R_{i,t-2} + (3)^{p-1} R_{i,t-3} + \dots \\ + (n-1)^{p-1} R_{i,t-(n-1)}] + \varepsilon'_{1it}$$

$$(66) \quad Q_{it} = \beta_0 + \sum_{m=1}^5 \beta_m D_{mit} + \beta_6 W_{it} + \beta_7 \bar{W}_{it} + \beta_8 U_{it} + \varepsilon'_{2it}$$

$$(67) \quad Y_{it} = \gamma_0 + \sum_{m=1}^5 \gamma_m D_{mit} + \gamma_6 P_{it} + \gamma_7 N_{it} + \gamma_8 I_{it} + \gamma_9 H_{it} + \varepsilon'_{3it}$$

$$(68) \quad Y_{it} = \delta_0 + \sum_{m=1}^5 \delta_m D_{mit} + \delta_6 P_{it} + \delta_7 W_{it} + \delta_8 S_{it} + \delta_9 O_{it} + \delta_{10} G_{it} \\ + \delta_{11} H_{it} + b_0 [R_{it} + R_{i,t-1} + R_{i,t-2} + \dots + R_{i,t-(n'-1)}] \\ + b_1 [R_{i,t-1} + 2R_{i,t-2} + 3R_{i,t-3} + \dots + (n'-1)R_{i,t-(n'-1)}] \\ + b_2 [R_{i,t-1} + 4R_{i,t-2} + 9R_{i,t-3} + \dots + (n'-1)^2 R_{i,t-(n'-1)}] \\ \vdots$$

$$+ b_{p'-1} [R_{1,t-1} + (2)^{p'-1} R_{1,t-2} + (3)^{p'-1} R_{1,t-3} + \dots \\ + (n'-1)^{p'-1} R_{1,t-(n'-1)}] + \epsilon'_{4it}$$

By allowing

$$V_{10} = R_{1t} + R_{1,t-1} + R_{1,t-2} + \dots + R_{1,t-(n'-1)}$$

$$V_{11} = R_{1,t-1} + 2R_{1,t-2} + 3R_{1,t-3} + \dots + (n'-1)R_{1,t-(n'-1)}$$

$$V_{12} = R_{1,t-1} + 4R_{1,t-2} + 9R_{1,t-3} + \dots + (n'-1)^2 R_{1,t-(n'-1)}$$

$$\vdots \\ V_{1,p-1} = R_{1,t-1} + (2)^{p-1} R_{1,t-2} + (3)^{p-1} R_{1,t-3} + \dots \\ + (n'-1)^{p-1} R_{1,t-(n'-1)}$$

and

$$V_{20} = R_{1t} + R_{1,t-1} + R_{1,t-2} + \dots + R_{1,t-(n'-1)}$$

$$V_{21} = R_{1,t-1} + 2R_{1,t-2} + 3R_{1,t-3} + \dots + (n'-1)R_{1,t-(n'-1)}$$

$$V_{22} = R_{1,t-1} + 4R_{1,t-2} + 9R_{1,t-3} + \dots + (n'-1)^2 R_{1,t-(n'-1)}$$

$$\vdots \\ V_{2,p'-1} = R_{1,t-1} + (2)^{p'-1} R_{1,t-2} + (3)^{p'-1} R_{1,t-3} + \dots \\ + (n'-1)^{p'-1} R_{1,t-(n'-1)}.$$

The model was rewritten as:

$$(69) \quad Q_{it} = \alpha_0 + \sum_{m=1}^5 \alpha_m D_{mit} + \alpha_6 W_{it} + \alpha_7 P_{it} + \alpha_8 S_{it} + \alpha_9 O_{it} + a_0 V_{10} \\ + a_1 V_{11} + a_2 V_{12} + \dots + a_{p-1} V_{1,p-1} + \epsilon'_{1it}$$

$$(70) \quad Q_{it} = \beta_0 + \sum_{m=1}^5 \beta_m D_{mit} + \beta_6 W_{it} + \beta_7 \bar{W}_{it} + \beta_8 U_{it} + \epsilon'_{2it}$$

$$(71) \quad Y_{it} = \gamma_0 + \sum_{m=1}^5 \gamma_m D_{mit} + \gamma_6 P_{it} + \gamma_7 N_{it} + \gamma_8 I_{it} + \gamma_9 H_{it} + \epsilon'_{3it}$$

$$\begin{aligned}
 (72) \quad Y_{it} = & \delta_o + \sum_{m=1}^5 \delta_m D_{mit} + \delta_6 P_{it} + \delta_7 W_{it} + \delta_8 S_{it} + \delta_9 O_{it} + \delta_{10} G_{it} \\
 & + \delta_{11} H_{it} + b_o V_{20} + b_1 V_{21} + b_2 V_{22} + \dots + b_{p'-1} V_{2,p'-1} \\
 & + \varepsilon'_{4it}
 \end{aligned}$$

These equations were the stochastic structural model after reducing the parameter spaces of the distributed lag functions using the general polynomial form. The system was overidentified, so a simultaneous equation estimation method was required. The two-stage least-squares method was used. Consequently, the reduced form of the statistical model was derived.

II. THE REDUCED FORM

The reduced form was derived by writing:

$$\begin{aligned}
 (73) \quad Q_{it} - \alpha_6 W_{it} - \alpha_7 P_{it} = & \alpha_o + \sum_{m=1}^5 \alpha_m D_{mit} + \alpha_8 S_{it} + \alpha_9 O_{it} + \alpha_o V_{10} \\
 & + \alpha_1 V_{11} + \alpha_2 V_{12} + \dots + \alpha_{p-1} V_{1,p-1} + \varepsilon'_{1it}
 \end{aligned}$$

$$(74) \quad Q_{it} - \beta_6 W_{it} = \beta_o + \sum_{m=1}^5 \beta_m D_{mit} + \beta_7 \bar{W}_{it} + \beta_8 U_{it} + \varepsilon'_{2it}$$

$$(75) \quad Y_{it} - \gamma_6 P_{it} = \gamma_o + \sum_{m=1}^5 \gamma_m D_{mit} + \gamma_7 N_{it} + \gamma_8 I_{it} + \gamma_9 H_{it} + \varepsilon'_{3it}$$

$$\begin{aligned}
 (76) \quad Y_{it} - \delta_6 P_{it} - \delta_7 W_{it} = & \delta_o + \sum_{m=1}^5 \delta_m D_{mit} + \delta_8 S_{it} + \delta_9 O_{it} + \delta_{10} G_{it} \\
 & + \delta_{11} H_{it} + b_o V_{20} + b_1 V_{21} + b_2 V_{22} + \dots + b_{p'-1} V_{2,p'-1} + \varepsilon'_{4it}
 \end{aligned}$$

or

$$(77) \quad \begin{bmatrix} 1 & -\alpha_6 & 0 & -\alpha_7 \\ 1 & -\beta_6 & 0 & 0 \\ 0 & 0 & 1 & -\gamma_6 \\ 0 & \delta_7 & 1 & -\delta_6 \end{bmatrix} \begin{bmatrix} Q_{it} \\ W_{it} \\ Y_{it} \\ P_{it} \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_0 & \alpha_1 & \dots & \alpha_5 & \alpha_8 & \alpha_9 & 0 & 0 & 0 & 0 & 0 & 0 & a_0 & \dots & a_{p-1} & 0 & \dots & 0 \\ \beta_0 & \beta_1 & \dots & \beta_5 & 0 & 0 & \beta_7 & \beta_8 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \gamma_0 & \gamma_1 & \dots & \gamma_5 & 0 & 0 & 0 & 0 & \gamma_7 & \gamma_8 & 0 & \gamma_9 & 0 & \dots & 0 & 0 & \dots & 0 \\ \delta_0 & \delta_1 & \dots & \delta_5 & \delta_8 & \delta_9 & 0 & 0 & 0 & 0 & \delta_{10} & \delta_{11} & 0 & \dots & 0 & b_0 & \dots & b_{p'-1} \end{bmatrix}$$

times

$$\begin{bmatrix} 1 \\ D_{11t} \\ D_{21t} \\ D_{31t} \\ D_{41t} \\ D_{51t} \\ S_{1t} \\ O_{1t} \\ \bar{W}_{1t} \\ U_{1t} \\ N_{1t} \\ I_{1t} \\ G_{1t} \\ H_{1t} \\ V_{10} \\ . \\ . \\ . \\ V_{1,p-1} \\ V_{20} \\ . \\ . \\ . \\ V_{2,p'-1} \end{bmatrix}$$

+

$$\begin{bmatrix} \epsilon'_{11t} \\ \epsilon'_{21t} \\ \epsilon'_{31t} \\ \epsilon'_{41t} \end{bmatrix}$$

or

$$(78) \quad \vec{C} \vec{e} = \vec{D} \vec{z} + \vec{\epsilon}$$

III. ESTIMATION OF THE REDUCED-FORM PARAMETERS

The estimating equations for the reduced form were found by

$$(79) \quad \vec{e} = \vec{C}^{-1} \vec{D} \vec{z} + \vec{C}^{-1} \vec{\epsilon}$$

or

$$(80) \quad \vec{e} = \vec{\pi} \vec{z} + \vec{u}$$

where

$$\vec{\pi} = \vec{C}^{-1} \vec{D}$$

a $4 \times (14 + p + p')$ matrix of reduced-form coefficients and

$$\vec{u} = \vec{C}^{-1} \vec{\epsilon}$$

a 4×1 vector of reduced-form disturbances.

The reduced form in expanded form was⁷

$$(81) \quad Q_{it} = \pi_{10} + \sum_{m=1}^5 \pi_{1m} D_{mit} + \pi_{16} S_{it} + \pi_{17} O_{it} + \pi_{18} \bar{W}_{it} + \pi_{19} U_{it} \\ + \pi_{1,10} N_{it} + \pi_{1,11} I_{it} + \pi_{1,12} G_{it} + \pi_{1,13} H_{it} + \sum_{q=0}^{p-1} \pi_{1q} V_{1q} \\ + \sum_{v=0}^{p'-1} \pi_{1v} V_{2v} + \mu_{1it}$$

$$(82) \quad W_{it} = \pi_{20} + \sum_{m=1}^5 \pi_{2m} D_{mit} + \pi_{26} S_{it} + \pi_{27} O_{it} + \pi_{28} \bar{W}_{it} + \pi_{29} U_{it} \\ + \pi_{2,10} N_{it} + \pi_{2,11} I_{it} + \pi_{2,12} G_{it} + \pi_{2,13} H_{it} + \sum_{q=0}^{p-1} \pi_{2q} V_{1q} \\ + \sum_{v=0}^{p'-1} \pi_{2v} V_{2v} + \mu_{2it}$$

⁷ Only one set of V's was utilized to avoid singularity when $n = n'$ and $p = p'$ because both sets of V's were equal.

$$\begin{aligned}
 (83) \quad Y_{it} = & \pi_{30} + \sum_{m=1}^5 \pi_{3m} D_{mit} + \pi_{36} S_{it} + \pi_{37} O_{it} + \pi_{38} \bar{W}_{it} + \pi_{39} U_{it} \\
 & + \pi_{3,10} N_{it} + \pi_{3,11} I_{it} + \pi_{3,12} G_{it} + \pi_{3,13} H_{it} + \sum_{q=0}^{p-1} \pi_{3q} V_{1q} \\
 & + \sum_{v=0}^{p'-1} \pi_{3v} V_{2v} + \mu_{3it}
 \end{aligned}$$

$$\begin{aligned}
 (84) \quad P_{it} = & \pi_{40} + \sum_{m=1}^5 \pi_{4m} D_{mit} + \pi_{46} S_{it} + \pi_{47} O_{it} + \pi_{48} \bar{W}_{it} + \pi_{49} U_{it} \\
 & + \pi_{4,10} N_{it} + \pi_{4,11} I_{it} + \pi_{4,12} G_{it} + \pi_{4,13} H_{it} + \sum_{q=0}^{p-1} \pi_{4q} V_{1q} \\
 & + \sum_{v=0}^{p'-1} \pi_{4v} V_{2v} + \mu_{4it}
 \end{aligned}$$

These four equations were the reduced-form estimating equations. Estimates of the reduced-form parameters ($\hat{\pi}_{re}$, where $r = 1 \dots 4$ reduced-form equations, and the "hat" indicated an estimated value), were obtained using classical least squares in the first stage of the two-stage least-squares estimation method, e.g.,

$$\begin{aligned}
 (85) \quad Q_{it} = & \hat{\pi}_{10} + \sum_{m=1}^5 \hat{\pi}_{1m} D_{mit} + \hat{\pi}_{16} S_{it} + \hat{\pi}_{17} O_{it} + \hat{\pi}_{18} \bar{W}_{it} + \hat{\pi}_{19} U_{it} \\
 & + \hat{\pi}_{1,10} N_{it} + \hat{\pi}_{1,11} I_{it} + \hat{\pi}_{1,12} G_{it} + \hat{\pi}_{1,13} H_{it} + \sum_{q=0}^{p-1} \hat{\pi}_{1q} V_{1q} \\
 & + \sum_{v=0}^{p'-1} \hat{\pi}_{1v} V_{2v} + \hat{\mu}_{1it}
 \end{aligned}$$

and similarly for the other three equations. The estimates of the reduced-form parameters obtained in the first stage were used to arrive at predicted values of the right-hand endogenous variables, W_{it} and P_{it} , in the original structural model (equations 69-72), e.g., $\bar{W}_{it}^*$ (or predicted W_{it}) was obtained by

$$\begin{aligned}
 (86) \quad \bar{W}_{it}^* = & \hat{\pi}_{20} + \sum_{m=1}^5 \hat{\pi}_{2m} D_{mit} + \hat{\pi}_{26} S_{it} + \hat{\pi}_{27} O_{it} + \hat{\pi}_{28} \bar{W}_{it} + \hat{\pi}_{29} U_{it} \\
 & + \hat{\pi}_{2,10} N_{it} + \hat{\pi}_{2,11} I_{it} + \hat{\pi}_{2,12} G_{it} + \hat{\pi}_{2,13} H_{it} + \sum_{q=0}^{p-1} \hat{\pi}_{2q} V_{1q} \\
 & + \sum_{v=0}^{p'-1} \hat{\pi}_{2v}
 \end{aligned}$$

and similarly for $\bar{P}_{it}^*$ (or predicted P_{it}).⁸

IV. ESTIMATION OF THE STRUCTURAL PARAMETERS

The second stage of two-stage least squares was used to obtain estimates of the structural parameters in the structural model. The estimates of the structural parameters of each structural equation were obtained by using classical least squares to regress the endogenous variable appearing on the left-hand side of each structural equation on the predicted values (obtained in the first stage) of the endogenous variables, $\bar{W}_{it}$ and $\bar{P}_{it}$, appearing on the right-hand side and the exogenous variables of each respective structural equation as follows:

$$\begin{aligned}
 (87) \quad Q_{it} = & \alpha_0 + \sum_{m=1}^5 \alpha_m D_{mit} + \alpha_6 \bar{W}_{it}^* + \alpha_7 \bar{P}_{it}^* + \alpha_8 S_{it} + \alpha_9 O_{it} + \alpha_{10} V_{10} \\
 & + \alpha_{11} V_{11} + \dots + \alpha_{p-1} V_{1,p-1} + \phi_{1it}
 \end{aligned}$$

$$(88) \quad Q_{it} = \beta_0 + \sum_{m=1}^5 \beta_m D_{mit} + \beta_6 \bar{W}_{it}^* + \beta_7 \bar{W}_{it} + \beta_8 U_{it} + \phi_{2it}$$

$$(89) \quad Y_{it} = \gamma_0 + \sum_{m=1}^5 \gamma_m D_{mit} + \gamma_6 \bar{P}_{it}^* + \gamma_7 N_{it} + \gamma_8 I_{it} + \gamma_9 H_{it} + \phi_{3it}$$

$$\begin{aligned}
 (90) \quad Y_{it} = & \delta_0 + \sum_{m=1}^5 \delta_m D_{mit} + \delta_6 \bar{P}_{it}^* + \delta_7 \bar{W}_{it}^* + \delta_8 S_{it} + \delta_9 O_{it} + \delta_{10} G_{it} \\
 & + \delta_{11} H_{it} + \delta_{20} V_{20} + \delta_{21} V_{21} + \dots + \delta_{p'-1} V_{2,p'-1} + \phi_{4it}
 \end{aligned}$$

⁸Predicted values of Q_{it} and Y_{it} (or $\bar{Q}_{it}$ and $\bar{Y}_{it}$) were not required because they did not appear as right-hand endogenous variables in the structural model.

where ϕ_{rit} were disturbances of the structural equations ($r = 1 \dots 4$) and all other variables were, as previously defined. Classical least squares performed on each of the above estimating equations resulted in the following equations containing the estimates of the structural parameters:

$$(91) \quad Q_{it} = \alpha_o + \sum_{m=1}^5 \alpha_m D_{mit} + \alpha_6^* \bar{W}_{it} + \alpha_7^* \bar{P}_{it} + \alpha_8 S_{it} + \alpha_9 O_{it} \\ + \hat{\alpha}_o V_{10} + \hat{\alpha}_1 V_{11} + \dots + \hat{\alpha}_{p-1} V_{1, p-1} + \hat{\phi}_{1it}$$

$$(92) \quad Q_{it} = \beta_o + \sum_{m=1}^5 \beta_m D_{mit} + \beta_6^* \bar{W}_{it} + \beta_7^* \bar{W}_{it} + \beta_8 U_{it} + \hat{\phi}_{2it}$$

$$(93) \quad Y_{it} = \hat{\gamma}_o + \sum_{m=1}^5 \hat{\gamma}_m D_{mit} + \hat{\gamma}_6^* \bar{P}_{it} + \hat{\gamma}_7^* N_{it} + \hat{\gamma}_8 I_{it} + \hat{\gamma}_9 H_{it} + \hat{\phi}_{3it}$$

$$(94) \quad Y_{it} = \delta_o + \sum_{m=1}^5 \delta_m D_{mit} + \delta_6^* \bar{P}_{it} + \delta_7^* \bar{W}_{it} + \delta_8 S_{it} + \delta_9 O_{it} + \delta_{10} G_{it} \\ + \delta_{11} H_{it} + \hat{b}_o V_{20} + \hat{b}_1 V_{21} + \dots + \hat{b}_{p'-1} V_{2, p'-1} + \hat{\phi}_{4it}$$

with "hats" indicating estimated values of the structural parameters.

V. DERIVATION OF ESTIMATES OF THE WEIGHTS OF THE LAG DISTRIBUTION

The estimates of the lag distribution parameters (a 's and b 's), describing the lagged effects of technology on the labor demand and product supply equations respectively, were obtained directly in the second stage of two-stage least squares as shown by equation (91) to (94) above. The estimates of the lag parameters were used to derive estimates of the annual weights representing the effects of a change in technology through time on both the labor demand and product supply

equations. As shown previously,

$$(95) \ w_1(s) = a_0 + a_1s + a_2s^2 + \dots + a_{p-1}s^{p-1}$$

and

$$(96) \ w_2(s) = b_0 + b_1s + b_2s^2 + \dots + b_{p'-1}s^{p'-1}$$

Consequently, estimated values of $w_1(s)$, the lag weight for the lagged effect of technology on the demand for farm labor, and $w_2(s)$, the lag weight for the lagged effect of technology on the supply of aggregate farm product were determined by the following computations:

$$(97) \ \hat{w}_1(s) = \hat{a}_0 + \hat{a}_1s + \hat{a}_2s^2 + \dots + \hat{a}_{p-1}s^{p-1}$$

and

$$(98) \ \hat{w}_2(s) = \hat{b}_0 + \hat{b}_1s + \hat{b}_2s^2 + \dots + \hat{b}_{p'-1}s^{p'-1}$$

where

$$\hat{w}_1(0) = \hat{a}_0$$

$$\hat{w}_1(1) = \hat{a}_0 + \hat{a}_1 + \hat{a}_2 + \dots + \hat{a}_{p-1}$$

$$\hat{w}_1(2) = \hat{a}_0 + 2\hat{a}_1 + 4\hat{a}_2 + \dots + (2)^{p-1}\hat{a}_{p-1}$$

⋮

$$\hat{w}_1(n-1) = \hat{a}_0 + (n-1)\hat{a}_1 + (n-1)^2\hat{a}_2 + \dots + (n-1)^{p-1}\hat{a}_{p-1}$$

and similarly for $\hat{w}_2(s)$ using the $\hat{b}$'s.

The $\hat{w}_k(s)$ ($k = 1$ or 2) computed from the estimates of the a 's or b 's were only approximations of the desired weights for two reasons. First, the sum of the desired lag weights (parameters) were restricted to unity, but using estimated values of the a 's or b 's did not give computed values of the weights, $\hat{w}_k(s)$, that summed to

unity. Consequently, the two series of weights were forced to sum to unity by dividing each element in each series by $\sum_{s=0}^{n-1} \hat{w}_k(s)$, $k = 1$ or 2 , and $n = n$ or n' , respectively.

Second, using the general polynomial form to reduce the parameter space of the function resulted in estimates of the weights which not only included the lag weights (where the weight was nonnegative and less than or equal to unity), but also the coefficient estimate of the overall lag effect of current and lagged technology. Consequently, the lag weights as defined by Johnson (1969, p. 3) were slightly different from those described by Griliches (1967, p. 18). For example, the general function

$$(99) Y_t = \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \dots + u_t$$

was specified by Griliches (1967, p. 18). He rewrote the function as

$$(100) Y_t = \beta [w_0 X_t + w_1 X_{t-1} + w_2 X_{t-2} + \dots] + u_t$$

where the w 's indicated the relative influence of differently lagged values of X and $w_j = \beta_j / \sum_1 \beta_1$ while β represented the overall effect. The w 's were all nonnegative and summed to unity.

Johnson (1969, p. 3) on the other hand specified a finite function similar to

$$(101) Y_t = \sum_{s=0}^{n-1} w(s) X_{t-s}$$

where $w(s) = a_0 + a_1 s + a_2 s^2 + \dots + a_{p-1} s^{p-1}$

The $\hat{w}(s)$, obtained with the Johnson method, included both the $\hat{\beta}$ and w_1 as specified by Griliches,⁹ i.e., the estimate of the overall

⁹The subscript "i" used by Griliches was equivalent to "s" used by Johnson.

effect of X (or β) and the estimate of the relative influence of differently lagged values of X (or the $\hat{w}_1$'s). However, by dividing each $\hat{w}(s)$ by the sum of the $w(s)$, obtained by using the Johnson method, to force the lag weights to sum to unity, resulted in a series of lag weights that were equivalent to those defined by Griliches, i.e.,

$$(102) \quad \bar{w}(s) = \frac{\hat{w}(s)}{\sum_{s=0}^{n-1} \hat{w}(s)} \sim \frac{\beta \hat{w}_1}{\beta \sum_{s=0}^{n-1} \hat{w}_1} = \frac{\hat{w}_1}{\sum_{s=0}^{n-1} \hat{w}_1} = \bar{w}_1$$

where "hats" indicated estimated values and $\bar{w}(s)$ and $\bar{w}_1$ were the estimates of desired lag weights or relative influence of differently lagged values of X in the Johnson and Griliches discussions, respectively.

The estimates of the desired weights of each of the two distributed lag functions of this study were denoted by $\bar{w}_k(s)$, $k = 1$ or 2 . For example, for $k = 1$

$$\begin{aligned} \bar{w}_1(t-0) &= \frac{\hat{w}_1(0)}{\sum_{s=0}^{n-1} \hat{w}_1(s)} \\ \bar{w}_1(t-1) &= \frac{\hat{w}_1(1)}{\sum_{s=0}^{n-1} \hat{w}_1(s)} \\ &\vdots \\ \bar{w}_1(t-n+1) &= \frac{\hat{w}_1(n-1)}{\sum_{s=0}^{n-1} \hat{w}_1(s)} \end{aligned}$$

and similarly for the $\bar{w}_2$'s.

VI. THE AVERAGE LAG

Griliches (1967, p. 18) stated that the lag weights can be identified formally with probabilities defined over the set of non-negative integers (0, 1, 2, 3, ...) if the weights are nonnegative and sum to one. Identification of weights of the infinite lag distribution with probabilities allowed Griliches to discuss and derive the formula for the average or mean lag of the distribution.

Derivation of the mean or average lag for each of the distributions in the model followed from the discussion by Griliches. For example, the average lag in years for the lag distribution for the effect of technology in the labor demand equation was found by taking the first derivative of the finite general polynomial function with respect to the lag operator L at $L = 1$. The labor demand equation in exact form was

$$(103) \quad Q_{it} = \alpha_0 + \sum_{m=1}^5 \alpha_m D_{mit} + \alpha_6 W_{it} + \alpha_7 P_{it} + \alpha_8 S_{it} + \alpha_9 O_{it} \\ + \sum_{s=0}^{n-1} w_1(s) R_{i,t-s}$$

or in expanded form with respect to the R 's

$$(104) \quad Q_{it} = \alpha_0 + \sum_{m=1}^5 \alpha_m D_{mit} + \alpha_6 W_{it} + \alpha_7 P_{it} + \alpha_8 S_{it} + \alpha_9 O_{it} \\ + w_1(0)R_{it} + w_1(1)R_{i,t-1} + \dots + w_1(n-1)R_{i,t-(n-1)}$$

Upon employing lag operator notation and defining the w 's in terms of the a 's (or lag parameters) the equation became

$$\begin{aligned}
 (105) \quad Q_{it} = & \alpha_0 + \sum_{m=1}^5 \alpha_m D_{mit} + \alpha_6 W_{it} + \alpha_7 P_{it} + \alpha_8 S_{it} + \alpha_9 O_{it} \\
 & + a_0 R_{it} + [a_0 + a_1 + \dots + a_{p-1}] R_{i,t-1} + \dots \\
 & + [a_0 + (n-1)a_1 + \dots + (n-1)^{p-1} a_{p-1}] R_{i,t-(n+1)}
 \end{aligned}$$

The equation, using lag operator notation, was

$$\begin{aligned}
 (106) \quad Q_{it} = & \alpha_0 + \sum_{m=1}^5 \alpha_m D_{mit} + \alpha_6 W_{it} + \alpha_7 P_{it} + \alpha_8 S_{it} + \alpha_9 O_{it} \\
 & + \left\{ a_0 + [a_0 + a_1 + \dots + a_{p-1}] L + \dots \right. \\
 & \left. + [a_0 + (n-1)a_1 + \dots + (n-1)^{p-1} a_{p-1}] L^{n-1} \right\} R_{it}
 \end{aligned}$$

or

$$(107) \quad Q_{it} = \alpha_0 + \sum_{m=1}^5 \alpha_m D_{mit} + \alpha_6 W_{it} + \alpha_7 P_{it} + \alpha_8 S_{it} + \alpha_9 O_{it}$$

$$+ \rho_1 A(L) R_{it}$$

$$(108) \quad \rho_1 A(L) = \left\{ a_0 + [a_0 + a_1 + \dots + a_{p-1}] L + \dots + [a_0 + (n-1)a_1 + \dots + (n-1)^{p-1} a_{p-1}] L^{n-1} \right\}$$

and ρ_1 = a measure of the overall influence of current and lagged technology

$A(L)$ = shorthand notation using the lag operator to show relative influence of different lagged values of R

so that

$$(109)^{10} \quad \rho_1 A(L) = w_1(0) + w_1(1)L + w_1(2)L^2 + \dots + w_1(n-1)L^{n-1}$$

¹⁰ The numbers in parentheses were for identification of weights only.

Estimates of the $w_1(s)$, or $\hat{w}_1(s)$ were used and they were forced to sum to unity so that

$$(110) \quad B(L) = \frac{\rho_1 \hat{A}(L)}{\sum_{s=0}^{n-1} \hat{w}_1(s)}$$

or

$$(111) \quad B(L) = \bar{w}_1(0) + \bar{w}_1(1)L + \bar{w}_1(2)L^2 + \dots + \bar{w}_1(n-1)L^{n-1}$$

which became the desired distribution.

Because the $\bar{w}_1$'s were nonnegative and summed to unity, they were identified as probabilities for the purpose of obtaining the average lag. Consequently, the general polynomial function $B(L)$ was assumed to be a probability generating function. Taking the first derivative of $B(L)$ with respect to L at $L = 1$ resulted in the mean lag (the first raw moment or expected value) of the lag time s at $L = 1$.

$$(112)^{11} \quad \left. \frac{dB(L)}{dL} \right|_{L=1} = \bar{w}_1(1) + 2\bar{w}_1(2) + \dots + [n-1] \bar{w}_1(n-1)$$

$$= \sum_{s=1}^{n-1} s \bar{w}_1(s) = \bar{s}_1$$

Similarly, the average lag for the distributed lag effect of technology in the product supply equation was found to be

$$(113) \quad \bar{w}_2(1) + 2\bar{w}_2(2) + \dots + [n'-1] \bar{w}_2(n'-1) = \sum_{s=0}^{n'-1} s \bar{w}_2(s) = \bar{s}_2$$

¹¹Numbers in parentheses in the equation were for identification of weights only.

VII. SELECTION OF THE DEGREE OF THE GENERAL POLYNOMIAL AND POINTS OF TRUNCATION

Prior to the actual estimation of the reduced-form and structural coefficients, alternative magnitudes of the degree of polynomial ($p-1$ and $p'-1$) and the number of lag weights (n and n'), had to be determined for the distributed lag functions of the model. The decision was one concerned with model specification, so judgment of the researcher, conditioned by past economic research of others, was important. Only a small amount of behavioral theory was available to aid in making the decisions, so several alternative combinations of general polynomials of various degrees and numbers of lag weights were estimated in both the labor demand equation and product supply equation.

Choices of the alternative combinations of p , p' , n , and n' to estimate were based primarily on the limitations of the data and judgment. General polynomials of various degrees ($p-1$ and $p'-1 = 0$ through 5) with n and n' , the numbers of lag weights, ranging from 1 to 61 were considered. A 60-year lag (observations on technology for the current year plus 60 back years) was the maximum length allowed by the data.

Not all possible combinations of p , p' , n , and n' in the labor demand and product supply equations were tried. The estimation of every possible set of distributed lag functions for both the labor demand and product supply equation where the set included every possible n and n' (from 1 through 61) over six different degrees (zero to fifth degree) of the polynomial would have resulted in a total

possible number of combinations of 133,947; many of which were conceptually unreasonable initially.¹² Consequently, the decision was made to limit the possible combinations of specifications.

This was accomplished by using (initially) 10-year intervals on n and n' (11, 21, 31, 41, 51, and 61 years) while using six different degree specifications in each of the two structural equations (zero through fifth degree).¹³

Prior to actual estimation, all possible combinations in the labor demand equation were not expected to be compatible with all combinations in the product supply equation from a conceptual standpoint. Both the supply of product and demand for labor were theoretically functions of, or jointly dependent with, the underlying production function and many other variables in the structural environment. Consequently, the optimal lag functions for both the supply of product and the demand for labor were expected to be similar with respect to the degree of the polynomial and the number of lag weights. So extreme differences in lag specifications of the two equations of the model were eliminated from consideration. Initially, no differences

¹²Each of the two structural equations incorporating distributed lags would have had 61 possible lag weights with 6 possible degrees (with a fifth degree polynomial assumed to be the maximum degree used), so that $61 \times 6 = 366$ possible specifications for each of the two structural equations. Then, in combining all possible specifications of the labor demand and product supply the total number of possible combinations would have been 366×366 or 133,956. Nine of the 133,956 combinations at $n = n' = 1$ for the first through fifth degrees and at $n = n' = 2$ for the second through the fifth degree either were not defined mathematically or led to perfect multicollinearity.

¹³The combination with zero lagged years of technology ($n = n' = 1$) was estimated using only a zero degree polynomial because the higher degree polynomials were equal to the zero degree polynomial after transformation.

in the degree of polynomial of the two distributed lag functions were allowed in any one specification (combination) in the first phase of estimation.

After the estimation of the initial alternative specifications (the first phase of estimation), the ranges with respect to p , p' , n , and n' of the most reasonable alternatives were to have been selected based upon the criteria discussed below for selection of the optimal specification. The specifications closely surrounding and including the most reasonable ranges of combinations of p , p' , n , and n' were to have been estimated in the second phase of estimation using annual intervals on n and n' while varying n (or n') and p (or p') both within each function and between each of the two distributed lag functions.

Selection of the optimal alternative combination of p , p' , n , and n' was to have been based on criteria similar to those used by Almon (1965). The criteria were judgment, the signs of the calculated weights, the close similarity of the lag weights among distributed lags of optimal and longer length, and reasonableness of the magnitude of the weights.

Positive signs on the calculated weights were desired. The sequence of the signs was important in determining the number of lag weights to include and also the degree of the polynomial. The optimal number of lag weights for a polynomial of a given degree also was considered to be in the time period previous to the period that lag weights among the distributed lags of optimal length and longer length became similar.

The final criterion for selection of the optimal lag functions

was reasonableness of the estimated lag weights and resulting shape of the plotted function. Reasonableness was based upon judgment and comparison of results of other studies.

Unfortunately, as Chapter VI demonstrates, results of the first phase of estimation were unreasonable with respect to one or more of the criteria so that the second phase of estimation was never fully initiated. Some exceptions were that the estimation of annual intervals for 9 through 1 lagged years on technology where $n = n' = 10, 9, \dots, 3$ at $p-1 = p'-1 = 0, \dots, 5$ and $n = n' = 2$ at $p-1 = p'-1 = 0$ and 1 was tried because of the sign patterns on weights obtained in the specifications of the first phase. This is discussed in the following chapter.

CHAPTER VI

RESULTS

This chapter presents a discussion of results obtained by using the estimation procedure in the previous chapter. No optimal lag function was selected because all estimated functions failed to satisfy one or more of the four criteria of selection listed in the preceding chapter. However, one function was selected as the least unreasonable for purposes of demonstration to allow discussion of parameter estimates in the model, especially as they related to the conceptual model of Chapter II.

I. THE ESTIMATED LAG DISTRIBUTIONS

First Phase of Estimation

None of the distributions estimated in the first phase of estimation were reasonable according to the selection criteria used. The most obviously violated criterion was the presence of negative lag weights distributed usually in series or runs of negative weights alternating with runs of positive lag weights over the range of back years of research and extension expenditures (technology) included in the specifications of lag polynomials of degrees greater than zero. A typical example was the alternative where $n = n' = 61$ and $p-1 = p'-1 = 2$. The lag weights for the distributed lag effect of technology on the demand for farm labor were positive for contemporaneous ($t = 0$) technology and from $t - 1$ through $t - 19$ back years of

technology. Then the weights became negative from $t - 20$ through $t - 54$ back years and positive again from $t - 55$ through $t - 60$ back years. Similarly, the lag weights for the distributed lag effect of technology in the product supply equation were positive for contemporaneous technology and also for $t - 1$ through $t - 28$ back years. The weights were negative for $t - 29$ through $t - 60$ back years. Other estimated alternatives were similar, i.e., runs of oppositely signed weights occurred in all alternatives for the distributed lag functions in the labor demand equation and also the product supply equation for all lag polynomials with degrees higher than zero. Consequently, these alternatives were eliminated because of the presence of negative weights.

The relatively naive zero-degree lag polynomials all had positive weights for the 10-year intervals from 0 to 60 back years on technology. However, the zero-degree lag polynomials hypothesized that all back years on technology had an equal effect on the dependent variable. For example, the zero-degree distributed lag including 60 back years on technology hypothesized that the distributed lag effect of technology was distributed equally among contemporaneous technology and 60 back years of technology. Consequently, each of these years of technology accounted for $1/61$ of the total effect of technology on the dependent variable. The lag weight for each of the years of technology included in the model were equal to each other and equalled $1/61$ or .0163934.

The estimated alternatives employing the zero-degree lag polynomials each satisfied the necessary conditions to qualify as

mathematically reasonable lag distributions. Namely, the desired lag weights $\bar{w}_k(s)$, $s = 0 \dots 60$, $k = 1$ or 2 , were all positive and summed to unity. However, they failed the criterion of conceptual reasonableness, and on this standpoint, were rejected. The effect of a change in technology was not expected to be allotted equally, but allotted unequally among technologies of different annual periods. The alternatives using the zero-degree lag polynomials, being restricted to an equal allotment of the effect, did not result in estimated lag distributions which would have allowed unequal allotment of the overall effect of technology on the demand for labor or supply of product. Higher degree polynomials would have resulted in unequal allotments, but they failed the positive-sign criterion. But rather than accept the alternative estimates of the zero degree polynomial as being reasonable and rejecting the higher order polynomials as being unreasonable, the decision was made to reject all alternatives.

The estimated lag weights of the alternatives with the higher ordered polynomials ($p-1 = p'-1 > 0$) had other notable characteristics other than the general presence of positive and negative weights that occurred in runs throughout the range of back years of technology. First, none of the estimated alternatives ever behaved in the manner implied by Almon. That is, Almon (1965, pp. 184-185) implied that the range of reasonable distributions included lag weights that were either all positive if the distribution was too short, or if the distribution was the optimum distribution. On the other hand, the weights should be positive except for a few weights alternating in sign in the earliest time periods (chronologically) when the specified

distributions were longer than optimum. The presence of only positive weights in the initial phase of estimation was no guarantee of selection of the optimum lag function because a function with too few time periods on the lagged exogenous variable theoretically could result in all positive weights. However, the optimum distribution had to have all positive weights, so the presence of all positive weights was a necessary, but not sufficient, condition for the selection of the optimum distribution based on the sign criterion. Consequently, sufficient data over a larger than optimum number of lagged time periods on the exogenous variable were desirable. Once the distributions that were longer than optimum were estimated, the earlier data on the lagged exogenous variable (having the lag weights that alternate in sign) could be removed and thus arrive at the specification of the optimum distribution.

The pattern of the weights in the estimated alternatives of the first phase of estimation in this study were not this well behaved, i.e., the weights were not mostly positive with minor and/or damped oscillations and changes in sign in the earlier years (chronologically) of lagged technology. The pattern was in general one of change in sign starting either with changes from negative to positive or positive to negative and continuing throughout the range of lagged technology.

All of the nonzero-degree lag polynomials that were estimated in the first phase had positive and negative signs which implied that the distributions were longer than optimum. To follow the Almon procedure of eliminating lag distributions that were longer than

optimum required estimation of lag distributions with fewer than 10 back years on technology. This was the second phase of estimation.

Second Phase of Estimation

Various alternative specifications of the model were estimated again with $p-1 = p'-1 = 0, 1, 2, 3, 4$, and 5 and $n = n' = 3, \dots, 10$ and $p-1 = p'-1 = 0$ and 1 at $n = n' = 2$. The second phase was an attempt at obtaining a lag distribution with positive weights at $p-1 = p'-1 > 0$ with lag parameter estimates that were greater than their asymptotic standard errors.¹ No such combination of desirable characteristics was obtained. Multicollinearity induced by the method of reducing the parameter space became a very serious problem in the second phase. Simple correlations between the V's (equations 87 and 90) were very high, which implied the presence of multicollinearity among the V's. Consequently, estimates of coefficients (the lag parameter estimates) on the V's in many cases were smaller than their standard errors and/or their signs and magnitudes were such that negative annual lag weights were obtained.

¹The reliability of an estimated coefficient in a regression equation is usually tested with the t-test after assuming that the error terms are normally distributed. However, the variance estimators for the parameters of a simultaneous system are asymptotic variances with their small sample properties being unknown (Klein, 1953, pp. 133-143). Hence, the utilization of the t-test to test whether or not a variable exerts a significant influence on the normalized (left-hand) endogenous variable is subject to question. Myers (1968, p. 55) discussed this problem in testing and followed a less restrictive "test" or "rule of thumb." He assumed that any estimated coefficient that was larger than its standard error (with both expressed as absolute values), had an important influence on the "dependent" variable. The same "rule of thumb" was employed in this study, although in most cases the coefficients tested significantly different from zero at the .05 level of significance using the more restrictive t-test.

All of the lag distributions in the second phase of estimation were found to be unreasonable because of failure to meet one or more of the selection criteria. Most of the estimated distributions for $p-1 = p'-1 > 0$ failed to satisfy the sign criterion. The one for the labor demand function that satisfied the sign criterion ($p-1 = p'-1 = 1$ at $n = n' = 6$)² failed to satisfy the criterion of conceptual reasonableness, besides having lag weights smaller than their standard errors and the wrong sign (positive) on the overall coefficient for technology. It failed the criterion of conceptual reasonableness because agriculture technology was expected to introduce changes in the labor and product markets that are spread over much longer periods of time than five years.

II. AN EXAMPLE OF THE ESTIMATED STRUCTURE

The estimated alternative at $n = n' = 61$ and $p = p' = 0$ was chosen for the purposes of demonstrating the structural parameters and their asymptotic standard errors and contrasting and comparing the signs of the coefficients with the expected signs hypothesized in Chapter II.

The example structure was presented only for these purposes because the researcher judged that it was pointless to attempt to make any substantive economic statements regarding the parameter estimates in the absence of a reasonable lag distribution. Such discussion would have been very relevant with respect to the stated objectives of the

²Some of the lag weights for the product supply relationship were negative in this combination. Hence, they were unreasonable.

study if the example structure had been the specification of the optimum lag structure. The estimated structure is presented in Table I. The estimated equations of the structure are discussed individually and then some general observations are made.

The Demand for Farm Labor

Signs of the estimated coefficients of farm wage and aggregate agricultural product price were in agreement with theoretical expectations, while the signs for the coefficients on land stock, other expenditures, and technology did not meet preconceived expectations. All regional dummy variables showing estimated intercept differences (in logarithms) were positive with respect to the South Central Region, which was the omitted region in the estimating equation. All estimated coefficients were greater than their asymptotic standard errors with the one exception being the coefficient on technology, which was less than its asymptotic standard error.

The positive signs of the estimated coefficients for land stock and other expenditures agreed with the positive values hypothesized in Chapter II. However, land stock and other expenditures were measured as quantity variables rather than the more desirable price measures. Consequently, the signs were not in agreement with theoretical expectations given that land stock and other expenditures represent true quantity measures.

The estimate of the lag parameter ($\hat{a}_0$) on technology for the zero degree polynomial and 60 back years of technology was 0.0002066. The overall effect (ρ_1) of all years of technology on the demand for

TABLE 1

TWO-STAGE LEAST-SQUARES ESTIMATES OF THE STRUCTURAL PARAMETERS
AND THEIR ASYMPTOTIC STANDARD ERRORS FOR THE STRUCTURE
SPECIFIED WITH ZERO-DEGREE LAG POLYNOMIALS
AND SIXTY BACK YEARS OF TECHNOLOGY

Identification ^a	Estimated Coefficients and Standard Errors of the Relationships ^b			
	Labor Demand	Labor Supply	Product Demand	Product Supply
<u>Constant Term</u>	-6.564 (1.192)	5.185 (.07966)	5.308 (1.008)	0.2662 (.7477)
<u>Regional 0-1 Dummies</u>				
North	0.3067 (.04304)	-0.1937 (.04105)	-0.2450 (.03724)	0.09530 (.02172)
Atlantic				
East North Central	0.07275 (.02934)	0.1522 (.03554)	0.1362 (.03224)	0.02929 (.01478)
West North Central	0.05250 (.03011)	-0.05007 (.03740)	.4674 (.02723)	.07300 (.01656)
South Atlantic	0.03199 (.02028)	-0.2143 (.02636)	-0.1482 (.02276)	-0.01670 (.01063)
Western	0.09005 (.03488)	-0.3128 (.04208)	0.2462 (.03239)	0.1250 (.01846)
<u>Continuous Variables</u>				
Farm Wage	-2.231 (0.1399)	0.3836 (.2034)		-1.106 (.1037)
Ag. Product Price	3.025 (0.5085)		-2.557 (.4957)	-0.9044 (.3266)
Land Stock	0.6256 (0.04587)			0.3033 (.02670)
Other Expenditures	0.3082 (0.05524)			0.9873 (.02888)
Alternative Wage		-1.654 (.1814)		

TABLE 1 (continued)

Variable Identification ^a	Estimated Coefficients and Standard Errors of the Relationships ^b			
	Labor Demand	Labor Supply	Product Demand	Product Supply
Unemployment Rate		-0.02852 (.04693)		
Population			1.298 (.1891)	
Disposable Income			-0.5627 (.1507)	
Home Consumption			0.05835 (.04366)	-0.1678 (.02117)
Government Payments				-0.1042 (.01402)
Technology	0.0002066 (.0008889)			0.004281 (.004535)

^aThe "dependent" or normalized, left-hand endogenous variable for each of the labor market equations was the quantity of farm labor. The "dependent" variable for each of the product market equations was the quantity of farm product.

^bAsymptotic standard errors appear in parentheses directly below the coefficients.

labor was 0.012603. The inference was (if one ignored the statistical nonsignificance) that a 10 percent increase in technology leads to a 0.126 percent increase in the demand for labor. The positive effect is distributed through 61 years of time, each with a lag weight of .016393.³ As stated before, all estimated lag distributions were evaluated as being unreasonable, but this example was chosen because it at least had all positive lag weights which was a necessary requirement to qualify as a reasonable distribution. However, this estimated distribution was rejected ultimately because the effect of technology likely is not distributed evenly over time, i.e., the true individual annual lag weights for all lagged years of technology in all probability are not equal.

Also, the sign of the overall effect of technology on the demand for labor disagreed with prior expectations. The positive sign on the coefficient for technology probably was reasonable given the estimated own-price elasticity of -2.557 in the product demand equation. Bauer (1968, pp. 42-45 and 52-53) and Wallace and Hoover (1966, pp. 330-332) showed that by adjusting their estimated coefficient of technology for the effect of the product market in each of their two-equation systems of the labor market, the adjusted technology coefficient was positive as long as the own-price elasticity of the demand for product was elastic. The coefficient became negative when the own-price elasticity of the demand for product became inelastic. Whether or not the same reasoning applied to the estimated example of this study was another

³ The determination of the values of the overall effect of technology and the lag weights are demonstrated below.

question. The model of this study was specified in part as an attempt to avoid the simplifying assumptions which the other two studies employed. That is, the model of this study did not require that the demand for labor be equal to the value of marginal product. Also a specific and more complete product demand function was specified rather than the general function used in the Bauer and Wallace-Hoover studies.

Finally, the lag parameter estimate was not greater than its standard error. Desired results would have included a significant, negative, and higher (absolute) effect of technology on the demand for labor that was distributed over a rather long period of years with unequal but positive annual lag weights in the lag distribution in conjunction with an own-price elasticity of product demand that was negative and inelastic. Of course, comparison of combinations of signs and magnitudes of actual estimates with the desired theoretical combinations would have depended on the ability of the specified structure to exemplify the true behavioral relationships and the linkages between these relationships in the economic system. Furthermore, questions concerning the dynamic nature of the system; specification, measurement, and aggregation errors; and difficulties in identifying demand and supply curves from economic data in a time series of cross sections prevented the establishment of a set of strongly held prior expectations concerning the expected combinations of relative magnitudes and signs of estimated coefficients on the technology variable in the farm labor demand and the product price variable in the product demand relationship. The same rationale applied to other possible combinations of coefficients within the structure.

Derivation of the lag weights. The annual lag weights or $\bar{w}_1(s)$ for $s = 0 \dots 60$ were computed using the procedure outlined in the previous chapter. The $\hat{w}(s)$, $s = 0 \dots 60$, or estimates of the approximate lag weights, were all positive and were found by utilizing the lag polynomial $\hat{w}_1(s)$ (equation 97). Substitution of $\hat{a}_0$ into (97) for the zero degree case ($\hat{a}_1 = \hat{a}_2 = \dots = \hat{a}_{p-1} = 0$) resulted in equal weights for each of the 61 years of technology with the value of the approximate annual lag weight being .0002066 or $\hat{a}_0$. These weights were only approximate because they did not sum to unity and they included the overall effect (ρ_1) for technology. They were forced to sum to unity by dividing by their sum or 0.012603. The desired estimated weights $\bar{w}_1(s)$, $s = 0 \dots 60$ were $\hat{w}_1(s) \div \sum_{s=0}^{60} \hat{w}(s)$, $s = 0 \dots 60$, or 61 weights each with a magnitude of 0.016393. The sum of the 61 lag weights was unity.

Derivation of the overall effect of technology. The overall effect of technology (ρ_1) on the demand for farm labor was determined by summing the approximate weights $\hat{w}_1(s)$, $s = 0 \dots 60$, i.e., the

$$\sum_{s=0}^{60} \hat{w}_1(s) = 0.012603.$$

The average or mean lag. The average lag in years for the distributed lag effect of technology in the labor demand relationship was determined as explained in Chapter V by utilizing equation (112).

$$(114) \quad \left. \frac{dB(L)}{dL} \right|_{L=1} = \bar{w}_1(1) + 2\bar{w}_1(2) + \dots + 60\bar{w}_1(60)$$

$$\begin{aligned}
&= \sum_{s=1}^{n-1} s \bar{w}_1(s) = \bar{s}_1 \\
&= \sum_{s=1}^{n-1} s(0.016393) \\
&= (1 + 2 + 3 + \dots + 60) (0.016393) \\
&= 30 \text{ years.}
\end{aligned}$$

The Supply of Farm Labor

The estimated coefficients for the farm wage rate and alternative wage rate had signs which agreed with a priori expectations. The sign of the coefficient on the unemployment rate disagreed with the expected direction of the relationship between the rate of unemployment and the quantity of labor supplied. Except for the positive sign on the coefficient for the East North Central Region, the signs of the coefficients for the regional zero-one dummy variables indicating the difference (on a logarithmic scale) in the intercept terms of each included region from the omitted South Central Region were all negative, but greater than -1.0. Consequently, the intercepts for the included regions⁴ after taking antilogs were all at higher levels for labor quantity than the omitted region. All estimated coefficients (absolute values) in the labor supply equation were greater than their respective asymptotic standard errors (absolute values) except the coefficient on the unemployment rate. This estimate was less than its asymptotic standard error.

⁴The intercept for each included region was computed by summing the antilogs of the estimated overall intercept and the respective estimated coefficient on the regional dummy. The overall intercept was taken as the intercept for the omitted South Central Region.

The Demand for Aggregate Farm Product

The magnitude of the intercepts for all the included regions in the estimating equation were greater than the intercept of the omitted South Central Region after the antilogs of coefficients were taken. The signs of the estimated coefficients on the continuous variables, aggregate product price and consumer population, agreed with preconceived expectations in Chapter II, while the signs of the coefficients on disposable income and home consumption did not. The estimates (absolute values) of the coefficients of all the continuous and dummy variables were greater than their asymptotic standard errors. As was expected, population and disposable income were highly correlated, having a simple correlation of 0.9773. Nevertheless, the estimates were significantly different from zero, even with the more restrictive t-test. The income elasticity of aggregate farm product should be, in reality, positive and inelastic rather than negative and inelastic.⁵

The Supply of Aggregate Farm Product

The estimated intercept terms (antilogs of the values in Table 1, page 110, plus the antilog of the overall intercept) for all the regions included in the estimating equation for the supply of farm product were greater than the omitted South Central Region. The signs of the estimated coefficients on the continuous variables--farm wage rate,

⁵ One interesting discovery was that estimation of the same example without including the regional intercept dummies resulted in some quite different results with respect to signs. That is, population had a negative coefficient while income had a positive coefficient.

land stock, other expenditures,⁶ home consumption and technology--were in agreement with prior theoretical expectations. However, the signs of the coefficients on aggregate product price and government payments did not agree with those hypothesized in Chapter II. All absolute values of the coefficient estimates except the technology and the overall intercept coefficients were greater than the absolute value of their respective standard errors. The estimate of the technology and overall intercept coefficients were smaller than their respective asymptotic standard errors.

Again, as in the similar case of the labor demand function above, the fact that the coefficient estimate of the lag parameter ($\hat{b}_0$) was not greater than its asymptotic standard error was ignored for the purpose of demonstrating the magnitudes of the desired annual lag weights $\bar{w}_2(s)$, $s = 0 \dots 60$; overall effect of technology (ρ_2); and the average or mean lag in years.

Derivation of the lag weights. The $\hat{w}_2(s)$, $s = 0 \dots 60$, or approximate weights were calculated by using equation (98) in Chapter V. Substitution of $\hat{b}_0$ into (98) for the zero-degree case ($\hat{b}_1 = \hat{b}_2 = \dots = \hat{b}_{p-1} = 0$) resulted in equal lag weights for each of the 61 years of technology. The magnitude of each of the approximate weights was .004281 or $\hat{b}_0$. Desired lag weights $\bar{w}_2(s)$, $s = 0 \dots 60$, were calculated by dividing each annual weight $\hat{w}_2(s)$, $s = 0 \dots 60$, by the sum of the $\hat{w}_2$'s. The estimates of the desired annual lag weights were each equal to

⁶ Assuming that land stock and other expenditures represent quantities.

0.016393. The sum of the 61 desired weights $\bar{w}_1(s)$, $s = 0 \dots 60$ was unity.

Derivation of the overall effect of technology. The overall effect (ρ_2) of technology on the supply of aggregate product was positive. The sign agreed with theoretical expectations. The magnitude of the overall effect (ρ_2) was 0.261141 and was determined by summing the approximate lag weights $\hat{w}_2(s)$, $s = 0 \dots 60$. This estimate implied that a 10 percent increase in technology would lead to a 2.6 percent increase in the quantity supplied of product with all other things held constant.

Derivation of the average or mean lag. The average lag in years was found by substituting estimated values for the desired weights into equation (113). That is,

$$(115) \quad \sum_{s=1}^{n'-1} s \bar{w}_2(s) = \sum_{s=1}^{60} s(0.016393) = 30 \text{ years}$$

III. SOME FINAL OBSERVATION ON THE RESULTS

One striking general characteristic concerning the parameter estimates of all the estimated alternative structures was the relative stability of the estimates of the structural parameters for variables other than the technology variable. One outstanding example was the own-price elasticity of product demand. The sign remained negative in all estimated structures and it was also quite elastic, usually around -2.0. As mentioned above in the discussion of the demonstration example, a negative, inelastic value was desired. Consequently, another

specification of the general statistical model was tried in an attempt to get a desired inelastic measure of the product price elasticity in combination with a negative effect of technology in the labor demand equation. The price of aggregate product was treated as the variable of normalization (the left-hand endogenous variable) in the product demand equation, while the wage rate became the variable of normalization in the labor supply equation. The altered model was estimated at the same combinations as in the first phase of estimation. The reciprocals of the estimated own-price flexibilities in the demand for aggregate product resulted in derived elasticities that were negative and even more elastic than the initial specifications in the first phase. The parameter estimates of the continuous variables except technology in the various estimated structures using the altered model continued to remain fairly stable as in the specifications of the first phase. The original general statistical model proposed in Chapter V was judged to be superior to the altered model, so further analysis with the altered model was abandoned.

Another characteristic was that the estimated lag parameters for the distributed lag effects of technology on the demand for farm labor and the supply of product in the various estimated structures were significantly greater than their asymptotic standard errors in most cases, even in the presence of multicollinearity induced by the method of reducing the parameter space. The above example structure failed to demonstrate this characteristic, but it was chosen as an example because the lag weights were all positive. Other higher order estimated lag functions having lag parameters significantly

different from zero were eliminated from consideration as an example because of the negative weights. Other zero-degree lag polynomial specifications had lag parameter estimates ($\hat{a}_0$'s and $\hat{b}_0$'s) that were greater than their respective asymptotic standard errors. For example, the lag parameter estimates for the zero-degree lag specifications at $n = n' = 1, 2, \dots, 10, 11, 21, \text{ and } 31$ were all greater than their respective standard errors.

In all of the zero-degree lag specifications a positive relationship was found to exist between the quantity of labor demanded and technology rather than the hypothesized negative relationship. A desired positive relationship was found between the quantity of product supplied and technology.

CHAPTER VII

SUMMARY, CONCLUSIONS, AND SUGGESTIONS FOR FUTURE RESEARCH

This chapter presents a summary of the study, some conclusions regarding the stated objectives, possible causes for unreasonable lag distributions, an assessment of the analytical approach, and a possible redirection. Finally, some suggestions for future research are listed.

I. SUMMARY

Changing technological conditions both within and outside the agricultural sector, along with other structural changes, have had and are continuing to have great impact on the labor input in agriculture. These effects do not occur instantaneously, but are spread or distributed over time. The technological horizon in agriculture appears unlimited from the vantage point of the present. Consequently, a better understanding of the direction and magnitude of effects of technology on the farm labor input both with respect to economic and human elements was deemed desirable from two standpoints. These were (1) the possible advancement of economic knowledge and (2) the utilization of results as input in various policy applications. Therefore, the objectives were (1) to estimate the effect of technology on the farm labor market and (2) to estimate the distributed lag function or time path of this effect.

The conceptual model, consisting of four behavioral equations,

the demand and supply of farm labor and the demand and supply of aggregate farm product, and two market equilibrium identities, was specified. Expected signs of structural parameters were hypothesized and discussed. Various available methods for reducing the parameter space of a general distributed lag function so that it can be estimated statistically, were contrasted and compared. The general polynomial method utilizing only lagged exogenous variables was selected, based upon the evaluation of the presently available methods.

The statistical model was derived after performing the data search and evaluating some of the economic and statistical limitations. The simultaneous system was estimated using the two-stage least-squares estimation method. Several alternative specifications were estimated so that the optimum specification of the model could be selected. The selection procedure was based on certain criteria leading to the optimum lag distribution and reasonable estimates of other structural parameters.

No reasonable lag distributions for the effects of technology in the farm labor and product markets were estimated because the estimated lag distribution for each alternative specification of the model failed to satisfy one or more of the selection criteria. The signs of a majority of estimates of other structural parameters were consistent with economic theory, but some were not. Magnitudes of the estimated coefficients for most structural variables other than technology seemed reasonable. Parameter estimates for most structural variables other than technology remained quite stable over all the alternative specifications.

II. CONCLUSIONS

None of the stated objectives of the study were attained. Consequently, no substantive conclusions regarding the distributed lag effects of technology on the farm labor input could be drawn from the results of this study.

Some conclusions were drawn regarding the analytical approach taken. The results indicated that in this specific application of general polynomial, finite distributed lags in a simultaneous equation framework, the use of regression analysis in a structure which did not constrain or restrict the signs and magnitudes of the lag weights permitted too much freedom in the estimated structures by allowing the data to determine the optimum distribution.¹ This conclusion was in opposition to the preconceptions of the researcher prior to estimation. However, empirical realities forced an ex post reassessment of preference for possible redirection from that of an unconstrained finite lag distribution, which lacked desirable properties, to that of a constrained distribution with the possibility of desirable properties.

Another conclusion was that the general polynomial form for reducing the parameter space of a finite distributed lag still was very appealing conceptually in spite of the unreasonable results. The method of reducing the parameter space probably was not at fault, in general,

¹Other unpublished research currently in progress by C. Hancock and L. Bauer in the Department of Agricultural Economics and Rural Sociology at the University of Tennessee, Knoxville, was experiencing similar difficulties. They were using general polynomial, finite lags in a single equation system with similar data. This information was obtained in a personal interview with the researchers.

other than multicollinearity, which was intensified when using the general polynomial reduction method. The primary shortcoming in this application of the reduction technique probably was the failure to restrict the range of magnitudes of the estimates of the lag parameters ($\hat{a}$'s and $\hat{b}$'s) so that the lag weights would be forced to be nonnegative. Also, certain other constraints such as restriction of the own-price elasticity of product demand to the inelastic range may have proven fruitful in obtaining estimates of reasonable lag distributions. As pointed out in Chapter III, the use of such restrictions would have been an application of quadratic programming. To convert to quadratic programming in this study would have required considerable redirection of the analysis. The additional research and computational costs were too high. However, this redirection along with some others listed below as suggestions for future research should prove fruitful if given further investigation in future studies.

Other experimentations with the model short of this redirection to quadratic programming could have been performed to research the possibilities of misspecification, measurement error, aggregation error, and so on. However, the model was deemed to be the most appropriate with respect to such considerations as choice of functional form, variables to include, deflators, and time unit of observation, under the existing limitations. To engage in such experimentations as trying different deflators and other functional forms prior to the complete redirection of the analysis toward the seemingly more important consideration of a priori constraints on the lag weights and some of the structural parameters, such as the price elasticity of

product demand, probably would have contributed little toward reaching the stated objectives of this study. The proper sequence of a proposed future analysis probably would be to first constrain the lag weights to be nonnegative for the alternative lag distributions, along with the negative inelastic restriction on the own-price elasticity of product demand. Following this major redirection, other modifications could be made where necessary.

Finally, one implicit and unstated objective was met, even though the stated objectives were not attained. Economic knowledge was advanced at least to the extent that the outcome of the study demonstrated some of the many pitfalls to research in a relatively new and untouched area--the quantitative analysis of distributed lags in a simultaneous system. Moreover, the researcher, having experienced these pitfalls, hopefully will avoid making the same recognized mistakes (at least in the same form) again in future research endeavors. The reader perhaps will be more fortunate by avoiding the pitfalls altogether as a result of having read this dissertation.

III. SUGGESTIONS FOR FUTURE RESEARCH

A suggestion for further research that probably would be very helpful would be the improvement of certain data such as the quantity of labor and the "quantity" of aggregate farm product. Also, the statistical model obviously could be improved by including additional variables as indicated by the conceptual model of Chapter II.

Another important extension that might help in breaking the empirical deadlock of the present model and also to be more useful in

policy applications would be to separate the United States into regions so that the regional parameters of the effect of technology and its distributed lag function for each of the regions could be estimated. The model of this study accounted only for regional intercept differences. However, a disaggregated regional analysis could provide separate estimates for both intercepts and coefficients in the various regions.

Distributed lag analysis as utilized in this problem setting was temporal in nature, i.e., the effect of technological change on the farm labor market through time was examined. Another possibility would be an investigation of the spatial distribution of the effect of technological change on the farm labor market from one region or state to another region or state.

The previous problem area is related to the more general problem area mentioned earlier concerning the proper interpretation of a time series of cross sections. More econometric knowledge is necessary on the expected properties of the parameters of such a system. Doubts associated with such questions as whether or not too much is being asked of the data need to be resolved among others.

Possible avenues for future empirical research on this particular model that were considered originally prior to specifying the statistical model of this study include:

1. The application of quadratic programming to the model with constraints that force the lag weights to be nonnegative and imposition of restrictions on certain parameters (e.g., the own-price elasticity of product demand restricted to a range between 0 and -1).

2. Use of a lag distribution of a known form such as the declining geometric lag or the inverted "V" (De Leeuw, 1962).²
3. Utilization of a geometric mean of current and lagged technology (Wallace and Hoover, 1966).
4. Respecification of the model as a bloc recursive system with the labor market equations making up one bloc and the product market equations making up the other bloc. Utilization of a bloc recursive system would permit the utilization of the rational polynomial method of reducing the parameter space.³

Future needs of research in general on technology and its effect on the farm labor market should emphasize the effect of all forms of technology, originating from research and extension activities in both the public and private sectors and not just the public sector. A closely related, but more general problem, is the need for more adequate measures of technology. In addition, more empirical research on the importance of the present and future impact of technology in the agricultural sector as it relates to other structural characteristics of the sector and the economy is needed, so that the most pressing agricultural problems can be identified and research priorities established to seek solutions to these problems.

²See Evenson (1967) for an application of the inverted "V" in an agricultural problem setting.

³The reader can recall that Bauer (1968) obtained reasonable results using this method on similar data.

REFERENCES

REFERENCES

- Agricultural Stabilization and Conservation Service. 1949-1964. Agricultural Conservation Program Statistical Summary, annually. U.S. Government Printing Office, Washington, D.C.
- Agricultural Stabilization and Conservation Service. 1966. Feed Grain and Wheat Program, 1965 Statistical Summary. U.S. Government Printing Office, Washington, D.C.
- Agricultural Stabilization and Conservation Service. 1970. Final Report Conservation Reserve Program Summary of Accomplishments 1956-1972. U.S. Government Printing Office, Washington, D.C.
- Almon, S. 1965. "The Distributed Lag Between Capital Appropriations and Expenditures," Econometrica, 22(1):178-196.
- Bauer, L. L. 1968. "The Effect of Technology on the Farm Labor Market," Unpublished Doctoral dissertation, North Carolina State University at Raleigh, Raleigh.
- Biederman, Paul (ed.). 1967. Economic Almanac 1967-1968. The Macmillan Company, New York.
- Bishop, C. E. 1967. "Dimensions of the Farm Labor Problem," pp. 1-17. In C. E. Bishop (ed.), Farm Labor in the United States. Columbia University Press, New York.
- Breimyer, Harold F. 1962. "The Three Economies of Agriculture," Journal of Farm Economics, 44(3):679-699.
- Bretzfelder, R. B., Q. F. Dallavalle, and D. A. Hirschberg. 1969. "Personal Income, 1968, and Disposable Income, 1929-68, by States and Regions," Survey of Current Business, U.S. Department of Commerce, Office of Business Economics, 49(4):16-32. U.S. Government Printing Office, Washington, D.C.
- Cagan, P. 1956. "The Monetary Dynamics of Hyper-inflation," pp. 25-117. In Milton Freidman (ed.), Studies in the Quantity Theory of Money. University of Chicago Press, Chicago.
- Chetty, V. K. 1971. "Estimation of Solow's Distributed Lag Models," Econometrica, 39(1):99-117.
- Cochrane, Willard. 1958. Farm Prices: Myth and Reality. University of Minnesota Press, Minneapolis.
- Council of Economic Advisors. 1971. Economic Report of the President. U.S. Government Printing Office, Washington, D.C.

- De Leeuw, F. 1962. "The Demand for Capital Goods by Manufacturers: A Study of Quarterly Time Series," Econometrica, 30(3):407-423.
- Edwards, C. 1962. "Non-Linear Programming and Non-Linear Regression Procedures," Journal of Farm Economics, 44(1):100-114.
- Evenson, R. 1967. "The Contribution of Agricultural Research to Production," Journal of Farm Economics, 49:1415-1425.
- Foote, R. J. 1958. Analytical Tools for Studying Demand and Price Structures. U.S. Department of Agriculture Handbook No. 146. U.S. Government Printing Office, Washington, D.C.
- Gisser, M. 1965. "Schooling and the Farm Problem," Econometrica, 33(3):582-592.
- Goldberger, A. S. 1964. Econometric Theory. John Wiley and Sons, Inc., New York.
- Griliches, F. 1967. "Distributed Lags: A Survey," Econometrica, 35(1):16-49.
- Hoover, D. M. 1961. "Land Prices in United States Agriculture, 1910-1950." Unpublished Doctoral dissertation, University of Chicago. University Microfilms, Ann Arbor, Michigan.
- Johnson, T. 1969. "A Note on Distributed Lags in Polynomial Form." Unpublished paper presented to an econometric workgroup at North Carolina State University in August, 1969.
- Jorgenson, W. 1966. "Rational Distributed Lag Functions," Econometrica, 34(1):135-149.
- Klein, L. R. 1953. A Textbook of Econometrics. Row, Peterson and Company, Evanston, Illinois.
- Klein, L. R. 1958. "The Estimation of Distributed Lags," Econometrica, 26(4):553-565.
- Koyck, L. M. 1954. Distributed Lags and Investment Analysis. North-Holland Publishing Co., Amsterdam.
- Learn, E. W. and W. W. Cochrane. 1961. "Regression Analysis of Supply Functions Undergoing Structural Change," pp. 63-72. In E. O. Heady, C. B. Baker, H. G. Diesslin, E. Kehrberg, and S. Staniforth (eds.), Agricultural Supply Functions--Estimating Techniques and Interpretations. Iowa State University Press, Ames.
- Myers, L. H. 1968. "Economic Analysis of the Monthly Structure of the Hog-Pork Sector of the U.S." Unpublished Doctoral dissertation, Purdue University, Lafayette.

- Nerlove, M. 1956. "Estimates of the Elasticities of Supply of Selected Agricultural Commodities," Journal of Farm Economics, 38(2):496-509.
- Nerlove, M. 1958. Distributed Lags and Demand Analysis. U.S. Department of Agriculture Handbook 141. U.S. Government Printing Office, Washington, D.C.
- Schuh, G. E. 1962. "An Econometric Investigation of the Market for Hired Labor in Agriculture," Journal of Farm Economics, 44(2): 307-321.
- Schuh, G. E. 1966. "The Influence of Technological Advance Upon Farm Factor Demands, Supplies, and Income Shares," pp. 65-89. In W. B. Back, D. M. Hoover, and J. A. Martin (eds.), Income Distribution Analysis. Agricultural Policy Institute Series 23, North Carolina State University, Raleigh.
- Schultz, T. W. 1967. "National Employment, Skills, and Earnings of Farm Labor," pp. 53-69. In C. E. Bishop (ed.), Farm Labor in the United States. Columbia University Press, New York.
- Solow, R. M. 1960. "On a Family of Lag Distributions," Econometrica, 28(2):399-406.
- Stout, T. T. and V. W. Ruttan. 1958. "Regional Patterns of Technological Change in American Agriculture," Journal of Farm Economics, 40(2):196-207.
- Stroud, A., A. Zellner, and L. C. Chau. 1963. Revised by M. Thornber and A. Zellner. 1965. "Program for Computing Two- and Three-Stage Least Squares Estimates and Associated Statistics," Social Systems Research Institute, University of Wisconsin, River Falls.
- Tweeten, L. G. 1969. "Theories Explaining the Persistence of Low Resource Returns in a Growing Farm Economy," American Journal of Agricultural Economics, 51(4):798-817.
- Tyrchniewicz, E. W. and G. E. Schuh. 1969. "Econometric Analysis of the Agricultural Labor Market," American Journal of Agricultural Economics, 51(4):770-787.
- U.S. Bureau of the Census. 1942. Census of Agriculture: 1940. Volume 1, parts 1-6. U.S. Government Printing Office, Washington, D.C.
- U.S. Bureau of the Census. 1949-1970. Statistical Abstract of the United States, annual issues. U.S. Government Printing Office, Washington, D.C.

- U.S. Bureau of the Census. 1961. Census of Agriculture: 1959.
Volume 1, parts 1-48. U.S. Government Printing Office, Washington,
D.C.
- U.S. Bureau of the Census. 1966. Long Term Economic Growth 1860-
1965. U.S. Government Printing Office, Washington, D.C.
- U.S. Bureau of the Census. 1967. Census of Agriculture: 1964.
Volume 1, parts 1-48. U.S. Government Printing Office, Washington,
D.C.
- U.S. Department of Agriculture. 1949-1965. Farm Labor, various issues.
U.S. Government Printing Office, Washington, D.C.
- U.S. Department of Agriculture. 1949-1970. Agricultural Statistics,
annual issues. U.S. Government Printing Office, Washington, D.C.
- U.S. Department of Agriculture. 1959. Farm Income Situation.
Economic Research Service, FIS 175. U.S. Government Printing
Office, Washington, D.C.
- U.S. Department of Agriculture. 1962. Number of Farms 1910-1959,
Land in Farms 1910-1959. Statistical Reporting Service, Statistical
Bulletin 316. U.S. Government Printing Office, Washington, D.C.
- U.S. Department of Agriculture. 1965. Agricultural Prices, 1964
Annual Summary. Statistical Reporting Service, Pr. 1-3(65).
U.S. Government Printing Office, Washington, D.C.
- U.S. Department of Agriculture. 1966. Farm Income Situation.
Economic Research Service, FIS 203 Supplement. U.S. Government
Printing Office, Washington, D.C.
- U.S. Department of Agriculture. 1968. Number of Farms and Land
in Farms. Statistical Reporting Service, Sp Sy 3(1-68). U.S.
Government Printing Office, Washington, D.C.
- U.S. Department of Agriculture. 1970a. Farm Income State Estimates
1949-1969. Economic Research Service, FIS 216 Supplement.
U.S. Government Printing Office, Washington, D.C.
- U.S. Department of Agriculture. 1970b. 1970 Handbook of Agricultural
Charts. U.S. Department of Agriculture Handbook No. 397. U.S.
Government Printing Office, Washington, D.C.
- U.S. Department of Labor. 1964. Manpower Report of the President.
U.S. Government Printing Office, Washington, D.C.
- U.S. Department of Labor. 1965. Employment and Earnings Statistics for
the States and Areas, 1939-1964. Bureau of Labor Statistics Bulletin
1370-72. U.S. Government Printing Office, Washington, D.C.

U.S. Department of Labor. 1970. Manpower Report of the President.
U.S. Government Printing Office, Washington, D.C.

Wallace, T. D. and D. M. Hoover. 1966. "Income Effects of Innovation:
The Case of Labor in Agriculture," Journal of Farm Economics,
48(2):325-336.

APPENDICES

APPENDIX A

SOURCES OF DATA AND CONSTRUCTION OF VARIABLES

A large quantity of data was used; consequently, the data are not listed here. A listing of the data may be obtained from the author. A verbal explanation of the construction and description of each of the variables of the model follows, along with some specific considerations on some of the variables that were not discussed in the text.

Data for 39 states or groups of states over the years 1949-1964 were calculated to measure the variables of the model except technology. State data for the technology variable were over the years 1889-1964. The groups of states were:

1. Maine, Vermont, New Hampshire, Massachusetts, Connecticut, Rhode Island
2. Delaware, Maryland
3. Arizona, New Mexico
4. Nevada, Utah, Wyoming

These states were grouped because of incomplete data for one or more of the variables, i.e., the data were either missing for a state in the group for a few of the years, or they were originally collected and recorded at the group level instead of the state level. The remaining 35 of the contiguous 48 states were considered individually.

I. LABOR¹

Total farm labor was measured in thousands of man-days by state

¹The labor variable, along with farm wage, technology, land

by year. Total farm labor was computed by summing derived estimates of the operator, family, and hired labor components of the farm labor force.

Operator Labor

Data for this component did not exist, so an estimate of this component was derived using the following variables on which data were available--the number of farm operators in each of the Census of Agriculture classifications of off-farm work for each census year (1950, 1954, 1959, and 1964) and the total number of operators (assumed to be equal to the number of farms) for each state for each census and noncensus year. Specifically, the number of farm operators in each census class of off-farm work for each census year (1950, 1954, 1959, and 1964) (U.S. Bureau of the Census, 1961, State Table 4 and 1967, State Table 18) was divided by the number of operators reporting (assumed to equal the number of farms reporting) for each state for each respective census year (U.S. Bureau of the Census, 1961, State Table 4 and 1967, State Table 18), i.e.,

$$\frac{\text{operators in each census class of off-farm work}}{\text{total number of operators}} = \text{proportion of operators in each census class}$$

The proportion of operators in each census class then was multiplied by the corresponding number of days the operator worked on a farm. This number of days of on-farm work was assumed to be 300 days less the mid-point of days of off-farm work in each census classification. Summing the products of these two variables for each state for each

stock, other expenditures, alternative wages, and unemployment, were the same as in Bauer (1968).

census year resulted in the average number of days of on-farm work per operator in each state for each census year. This product was multiplied by the number of operators (or the number of farms) (U.S. Department of Agriculture, 1962, 1968) for the respective census years and for the two years on either side of the census year. This multiplication resulted in an estimate of the total annual number of days of on-farm work performed by operators.

Family Labor

Estimates by state by year for this component were derived by multiplying the average number of unpaid workers, obtained by subtracting the number of farm operators (or the number of farms) (U.S. Department of Agriculture, 1962, 1968) from the annual average number of family workers (U.S. Department of Agriculture, 1949-1965), times one-half the average number of days worked by farm operators obtained above.

Hired Farm Labor

Total hired farm labor by state by year was determined by dividing the total expenditure for hired farm labor (U.S. Department of Agriculture, 1966, Table 10) by the composite farm wage rate (U.S. Department of Agriculture, 1949-1965) times 10 to get an average daily wage rate.²

II. FARM WAGE

The farm wage rate was taken as the average composite wage rate

² A 10-hour workday was assumed.

per hour (U.S. Department of Agriculture, 1949-1965) deflated by the implicit price deflator for total gross national product (Council of Economic Advisors, 1971, Table C-3). A weighted average of the farm wage rate was obtained where individual states were summed. The weights were expenditures for hired farm labor (U.S. Department of Agriculture, 1966, Table 10).

III. TECHNOLOGY

Technology was represented by the proxy variable, public dollar expenditures for research and extension. Data on both research and extension expenditures were obtained from many federal and state sources too numerous to list individually here. A list of sources is available from the author.

Annual research expenditures from 1889 through 1964 were computed as the grand total of federal and nonfederal sources of experiment station income less the carry-over balance of federal and nonfederal sources to the following year (total unexpended income). Extension expenditures included the expenditures of funds from all federal sources and from sources within the state such as state and college, county, farm organizations, etc., from 1915 through 1964. Expenditures on Farmers' Institutes (an early form of extension activity that was in force in many states prior to the passage of the Smith-Lever Act of 1914, which established the Cooperative Extension Service) were included for the years 1899 through 1930. No extension expenditures were included directly for 1889 through 1898. Extension expenditures in these years were assumed to be included in research funds expended

by experiment stations. Many activities which presently are considered to be extension activities were conducted by the state experiment stations prior to 1915. Consequently, research expenditures included extension expenditures in these earlier years.

The data from 1929 through 1964 were deflated by the total gross national product implicit deflator (Council of Economic Advisors, 1971, Table C-3) changed to a 1957-59 base. The total gross national product implicit deflator was not calculated by the Office of Business Economics, Department of Commerce, for the years prior to 1929. The author of this study calculated annual values for the implicit deflator for the years prior to 1929.

Annual values of the deflator for 1909 through 1928 were obtained by dividing gross national product expressed in current prices by the gross national product expressed in constant (1958) prices (Biederman, ed., 1967, p. 130). The index numbers then were converted to a 1957-59 base.

Annual values of the implicit price deflator for the years 1889 through 1908 were derived by multiplying the annual value of the implicit price deflator for gross national product (1929 = 100) originally calculated by Kendrick of the National Bureau of Economic Research (Bureau of the Census, 1966, p. 200) by .5064, the value of the implicit price deflator for gross national product in 1929 (1958 = 100), calculated by the Office of Business Economics (Bureau of the Census, 1966, p. 200). The annual values were then changed to a 1957-59 base.

The method used to derive the annual values of the deflator for 1889 through 1908 had some obvious shortcomings. One was that gross

national product was measured differently between the two series. Hence, each series measured the price movements of a different definition of gross output. Another problem was that the ratio between the two series probably did not remain constant at .5064 (the ratio for 1929) for all years, 1889-1908. A comparison of the annual ratios between the two series over the years, where they overlapped (using the researcher's calculated values for 1909-1929 and the Kendrick values for the same years), indicated that annual ratios ranged from .4849 to .5393. However, the mean of the annual ratios for the two series for 1909-1929 was .5067 which was very close to .5064, the assumed annual ratio for 1889-1908.

Nevertheless, the decision was made to employ the implicit price deflator as determined above, even with its shortcomings, rather than employing some other deflator such as the wholesale price index, which was available over the necessary years. The implicit price deflator was used because it was judged to be a better representation of movements in the general price level than any other available, but more narrowly defined deflator.

IV. LAND STOCK

Land stock was measured as a value aggregate in thousands of dollars of value. Total farmland acreage in pasture equivalents was obtained by converting the acreages of each of three classifications of land, i.e., cropland, pastureland and range, and woodland and other (Agricultural Stabilization and Conservation Service, 1949-1964) to pasture equivalents by multiplying each acreage classification

by their respective value from Hoover's (1961, Table 10) dissertation. Woodland and other was obtained by subtracting cropland and pastureland and range from total farmland. Furthermore, as Bauer (1968, p. 68) explained:

The acreage of cropland diverted out of production under the feed grain (Agricultural Stabilization and Conservation Service, 1966, Tables 63 and 64) and Soil Bank . . . [(Agricultural Stabilization and Conservation Service, 1970, Table 6)] programs were converted to pasture equivalents by multiplying the number of acres diverted times the cropland coefficient less the pastureland coefficient from Hoover's (1961, Table 10) thesis. This figure was then subtracted from total pasture equivalents.

This adjusted pasture equivalent figure was then multiplied by the 1940 value of one pasture equivalent expressed in 1957-59 dollars. This value was found by taking the total value of land in 1940 (U.S. Bureau of the Census, 1942, State Table 2), dividing it by the number of pasture equivalents in 1940, and multiplying by the wholesale price index for all commodities . . . [(U.S. Bureau of the Census, 1949-1970, specifically 1969, Table 499)] which was 233 for 1957-59 if the 1940 value was 100.

Wallace and Hoover (1966, pp. 334-335) explained:

This procedure allows the estimated stock of land to be free of any influence on values which may have resulted from the present supply control program. Only harvested cropland was weighted by the cropland pasture equivalent value in estimating the stock each year. Thus cropland which was diverted to fallow or diverted under a variety of voluntary and mandatory programs entered the computation of stock as pastureland to reflect its less intense use.

V. OTHER EXPENDITURES

The other expenditures variable was measured as a value aggregate in millions of dollars (U.S. Department of Agriculture, 1966, Table 10). Other expenditures were assumed to be the sum of expenditures on feed, livestock, seed, fertilizer and lime, repairs and operation of capital items, miscellaneous, and depreciation and other consumption of farm capital. The measure was converted to 1957-59 dollars by using the

national "Items Used for Production" price index (U.S. Department of Agriculture, 1965, p. 10).

VI. ALTERNATIVE WAGES

Alternative nonfarm wages were measured in dollars per hour and were taken from "Gross Average Hourly Earnings of Production Workers on Manufacturing Payrolls" (U.S. Department of Labor, 1965, Table 16). A weighted average wage rate was computed where states were grouped using "Employer or Manufacturing Payrolls" as weights. The measure was converted to 1957-59 dollars by using the implicit price deflator for total gross national product (Council of Economic Advisors, 1971, Table C-3) after changing the deflator to a 1957-59 base.

VII. UNEMPLOYMENT

Insured unemployment as a percent of average covered employment (U.S. Department of Labor, 1964, Table D-5 and U.S. Department of Labor, 1970, Table D-6) was used as a measure of unemployment. A weighted average of unemployment was computed where states were combined by using as weights the numbers of insured unemployed found in the same data sources.

This series does not equal actual unemployment. However, as Bauer pointed out:

. . . it is very highly correlated with the available actual unemployment rates, and was used because the actual unemployment data available did not cover the necessary time span.

VIII. FARM PRODUCT PRICES

Farm product prices were measured as state indices of prices received by farmers of the respective states. Data in both published and unpublished forms were obtained through personal correspondence by the author with Crop Reporting Boards and/or university agricultural economics departments for most of the states. The list of sources on product prices was not included in this study because of the large number used. A list of the sources is available from the author.

Some problems arose concerning the comparability of the price indices of the various states. As would be expected, a variety of methods was used by researchers of the various states to measure the indices of aggregate or all-commodities farm prices. This variety occurred partially because the construction of the index required decisions on the following among others:

1. The commodities to include in the index based on some criterion of importance with respect to the contributions of the commodities to the farm cash receipts of the state.

2. The total number of commodities to include based on some criterion of percentage of total cash receipts represented by the commodities included in the index.

3. The relevant price of each commodity to use in the index, i.e., questions on simple versus weighted averages for commodity grades, time of measurement, etc.

4. The base period.

5. The weight base period(s).

6. The basic index formula to use.

7. Method of weighting in aggregation from commodity to commodity groups to all commodities.

8. Method of adjusting for short-season crops.

Decisions on these questions by the price analysts of the different states resulted in a wide variety across states of commodities to include base periods, weight base periods, etc. Consequently, the author of this study was faced with deciding whether or not the data were sufficiently comparable in their present forms, or if not, could they be made comparable by recomputation. The latter possibility was ruled out because of the magnitude of the computations that would be required. So with the lack of no other feasible alternatives except eliminating the variable from the model, the decision was made to use the indices as originally calculated after converting the data of all states (where necessary) to the 1957-59 base. This was done simply by dividing each index number in the series for a state by the average value of that series in 1957-59. A weighted average of price indices was used where states were grouped with annual cash receipts from farm marketing by state (U.S. Department of Agriculture, 1970a, Table 6) serving as weights.

Another problem was that price indices data were not complete for all states. Three states had one annual observation missing and one state lacked two annual observations over the 16-year period. Those missing observations were either at the beginning or at the end of the period. A relatively crude method of estimating these missing observations was employed. The method involved extrapolation using predicted values estimated by a simple least-squares regression of price on time.

A more serious problem was that all-commodity price indices did not exist for the states of California, Texas, Florida, Kentucky, Delaware, New Hampshire, Rhode Island, Connecticut, and Maine. The last five states were grouped with other states (as pointed out above) which had index data on prices received by farmers at the all-commodity level of aggregation. Therefore, the existing price index data for a state or states in a group served as a proxy measure of the price index for the group, e.g., data were available on the index of prices received by Maryland farmers, so this index represented the price index of the Maryland-Delaware group. A weighted average of the index data for Massachusetts and Vermont served as a measure of the index of prices received for New England. The weights used were cash receipts of the two states.

The author of this study computed indices of prices received by farmers for California, Texas, Florida, and Kentucky. A modified Laspreyes formula³ with a 1957-59 base and weight base period was used. The calculated indices were somewhat limited because only the nine or ten most important commodities (where importance was based on relative value of cash receipts to the state) were included in the construction of the index. The commodity lists for California and Texas each included ten commodities, while the lists for Florida and Kentucky each contained nine commodities. Admittedly, the indices, if more commodities would have been included, probably would have been

³The formula was modified only to the extent that quantity of the numerator, the current aggregate, and the quantity and price of the denominator, the base aggregate, were three-year averages (1957-59) instead of single-period values.

more adequate as measures of the prices received by farmers in the various states. However, given the computational task, some reasonable limits on the number of commodities to include had to be imposed. The included commodities accounted for about 60 percent of the 1958 total farm cash receipts for California, 86 percent for Texas, 71 percent for Florida, and 91 percent for Kentucky (U.S. Department of Agriculture, 1959, Table 10). Prices and quantities for the commodities included in the indices were obtained from several issues of Agricultural Statistics (U.S. Department of Agriculture, 1949-1970, specifically 1949-1966).

IX. QUANTITY OF AGGREGATE FARM PRODUCT

No state data on the index of quantity of aggregate farm product were available, so value aggregate data were used as a measure of quantity. State cash receipts from farm marketings (U.S. Department of Agriculture, 1970a, Table 6) in millions of dollars were utilized as a measure of product quantity. This value aggregate was deflated by the national index of prices received by farmers for all farm products (U.S. Department of Agriculture, 1949-1970, specifically 1966, Table 686) after changing the index to a 1957-59 base.

Use of a value aggregate raised a particular question, namely the reliability of the estimate of aggregate quantity after accounting or attempting to adjust for changing price by deflating with a price index. As Foote pointed out (1958, p. 36),

The reliability of the estimate of changes in quantities obtained by the foregoing procedure depends primarily on the composition and internal stability of the value series in relation to the particular

price index employed. Changes in value may reflex shifts between the various products included in the value aggregate and not a change in total quantity or price. If such shifts are important, serious errors may arise, from using a price index incorporating weights that remain fixed over relatively long periods of time. In addition, the weights used in the price index may not necessarily be related to the relative importance of the products in the value aggregate even in the base period. To the extent that they differ, the estimate of quantity differs from the actual quantity. However, if the relative importance of the products in the value aggregate remains relatively stable over time, changes in the derived quantities should approximate changes in actual quantities. A somewhat similar consideration applies if the prices in the two series differ.

The problem of questionable reliability was intensified further by using a national index to deflate state data. Perhaps a more appropriate deflator (in general) would have been the state index of prices received by farmers. However, the state index of prices received by farmers was not used because it was specified as a jointly dependent variable in the system along with cash receipts, labor quantity, and the wage rate. Deflation of the cash receipts of a state (the normalized, left-hand endogenous variable) by use of the respective state price index (a right-hand endogenous variable) might have induced negative correlation between the quantity variable and the price variable.⁴ The national index was correlated with the price indices of the various states, so its use as a deflator also possibly could have induced correlation between the quantity variable and farm price in the equations. However,

⁴The problem would originate because of regressing a variable on itself, *i.e.*, the reciprocal of price times undeflated cash receipts being regressed on price. The extent of statistical difficulty probably would depend on how well cash receipts divided by the price index represented the true quantity variable. If the deflated cash receipts are an adequate measure of quantity rather than total revenue (price times quantity), the problem probably would not be too serious.

the extent of this "induced" correlation probably was not nearly as great because the national index was only correlated with, but not equal to, the state indices. The national deflator also was used in an attempt to remain consistent, because all other data on other variables were deflated by national deflators.

To utilize deflated cash receipts as a measure of quantity required some assumptions about the natures of the value and price series. The assumption was made that the composition and internal stability of the cash receipts for each state have remained relatively stable for the most part at least for the years included in this study. Furthermore, the relative importance of the products in the value aggregate as compared with the actual or "true" quantity series was assumed to have remained fairly stable over the period for most of the states. Consequently, the use of deflated cash receipts as a proxy for actual quantity was deemed reasonable, especially since no other measurement of aggregate quantity by state was available. Cash receipts were summed where states were grouped and then deflated.

X. GOVERNMENT PAYMENTS

Annual direct government payments to the farmers of each state (U.S. Department of Agriculture, 1970a, Table 6) were measured in millions of dollars. Government payments by state were summed where states were grouped. The series was deflated by the total gross national product implicit price deflator (Council of Economic Advisors, 1971, Table C-3) after changing the index to a 1957-59 base.

XI. HOME CONSUMPTION

Home consumption was measured as the value of home consumption in millions of dollars (U.S. Department of Agriculture, 1970a, Table 6). Home consumption by state was summed where states were grouped. The data were deflated by the national index of prices received by farmers (U.S. Department of Agriculture, 1949-1970, specifically 1966, Table 686) after changing the index to a 1957-59 base.

XII. INCOME

Disposable income by state (Bretzfelder, Dallavalle, and Hirschberg, 1969, pp. 24-25) less total net farm income⁵ (U.S. Department of Agriculture, 1970a, Table 6) was utilized as a measure of income in millions of dollars. The "adjusted" or nonfarm disposable income was summed where states were grouped. The data were deflated by the total gross national product implicit deflator (Council of Economic Advisors, 1971, Table C-3).

State disposable income is equal to state personal income less personal taxes and nontax payments (Bretzfelder, Dallavalle, and Hirschberg, 1969, p. 18). Personal income includes, among other things, the net income of farmers. Admittedly, farmers are also consumers of farm products, both through the marketplace and in kind through home

⁵ Total net farm income and not net disposable farm income was subtracted from state disposable income. Consequently, a measurement error equal to personal taxes and nontax payments paid by farmers are present in the measure of nonfarm disposable income. Nonfarm disposable income was understated by the amount of personal taxes and nontax payments paid by farmers.

consumption of farm products. Therefore, disposable farm income should be included conceptually either in disposable income or as a separate causal variable in the demand for aggregate farm product. But this inclusion of net farm income presented some possibilities of statistical problems which could have had detrimental effects on the parameter estimates obtained.

The problems were (1) regressing a part of a variable on itself and (2) possible additional collinearity between disposable income and other expenditures, government payments, or value of home consumption. The first problem could have occurred because the variable, undeflated cash receipts from farm marketings, was used, after deflating, by the national index of prices received by farmers, as a measure of quantity of aggregate farm product, an endogenous variable in the system. But cash receipts, along with government payments, value of home consumption, and gross rental value of farm dwellings, are components of realized gross farm income (U.S. Department of Agriculture, 1970a, Table 6). Realized gross income minus farm production expenses equals realized net farm income. Realized net farm income minus net change in farm inventories equals total net farm income. Consequently, undeflated cash receipts is a component of total net farm income. To regress deflated cash receipts on deflated state disposable income, which includes net farm income (less personal tax and nontax payments) could have biased upward the absolute magnitude of the coefficient estimate of the effect of disposable income on the endogenous "quantity" variable in the product demand equation.⁶ The second possible problem

⁶Again, this problem probably would not be a serious one if the variable, deflated cash receipts, adequately represents quantity of farm product.

of collinearity between state disposable income (which contains disposable farm income) and other expenditures,⁷ government payments, or value of home consumption could have occurred because the mentioned variables are all used in determining total net farm income by the U.S. Department of Agriculture.

Resolution of the above statistical problems led to the deletion of total net farm income as a component of state disposable income. The possibility of the problems arose because of having to use deflated value aggregate data as measures of actual quantities and prices. Some compromise was necessary, but conceptually the model probably was weakened as a result. Conceptually, the model, as specified, hypothesized that disposable farm income has no effect on the quantity demanded of aggregate farm product, which is an unrealistic hypothesis. However, one should note that the value of home consumption of farm products, one component of net farm income, was included as an exogenous variable in the product demand equation. Also, the model indirectly (and perhaps incorrectly in a conceptual sense) partially accounted for some of the other components of net farm income since these components, cash receipts, government payments, and other expenditures, were in the model. However, they were not treated as consumer income measures in the model, so no foreseeable interpretation of their effects relative to the effect of income on product demand was recognized.

⁷ Other expenditures is a component of total farm production expense (U.S. Department of Agriculture, 1970a, Table 6).

XIII. POPULATION

Estimated population in thousands of persons by state by year was obtained from various issues of Statistical Abstract of the United States (U.S. Bureau of the Census, 1949-1970, specifically 1958-1969). Estimates for a particular state for a particular year sometimes differed from one issue to the other, so the most current estimate was utilized. This was not a true measure of the actual population of consumers of farm products of a state because states engage in both intrastate and interstate commerce with respect to farm products. Data on the consuming population was unavailable, so state population was utilized instead.

APPENDIX B

AN ADDITIONAL REASON FOR NOT SPECIFYING TECHNOLOGY AS A JOINTLY DEPENDENT VARIABLE IN THE STATISTICAL MODEL

Chapter III showed that in equation (46)

$$V_{t0} = X_t + X_{t-1} + X_{t-2} + \dots + X_{t-(n-1)}$$

$$V_{t1} = X_{t-1} + 2X_{t-2} + 3X_{t-3} + \dots + (n-1) X_{t-(n-1)}$$

$$V_{t2} = X_{t-1} + 4X_{t-2} + 9X_{t-3} + \dots + (n-1)^2 X_{t-(n-1)}$$

.

.

.

$$V_{t,p-1} = X_{t-1} + 2^{(p-1)}X_{t-2} + 3^{(p-1)}X_{t-3} + \dots + (n-1)^{(p-1)}X_{t-(n-1)}$$

With X designated as expenditure on research and extension and considered as a jointly dependent variable in the system, V_{t0} contains both current and lagged values of expenditures, when current expenditures are jointly determined in the system. Conceptually, the lagged values of X could be considered as predetermined, but X_t the current value is jointly determined in the system because technology is assumed to be endogenous. To calculate V_{t0} to use in estimating the lag parameter a_0 , would require that the jointly determined variable X_t be summed with the predetermined lagged values of X in the estimating equation. Difficulty in arriving at an estimate of X_t (or $\hat{X}_t^*$) would result because the X 's are in the estimating equations as V 's which are sums of multiples of X 's. The X 's have lost their individual identity because of their transformation to V 's. Consequently, no feasible means of arriving at $\hat{X}_t^*$ in the first stage of two-stage least squares was recognized.

APPENDIX C

REGIONS AND STATES OR GROUPS OF STATES WITHIN EACH REGION

1. North Atlantic

New England States
New York
New Jersey
Pennsylvania

2. East North Central

Ohio
Indiana
Illinois
Michigan
Wisconsin

3. West North Central

Minnesota
Iowa
Missouri
North Dakota
South Dakota
Nebraska
Kansas

4. South Atlantic

Delaware-Maryland
Virginia
West Virginia
North Carolina
South Carolina
Georgia
Florida

5. Western

Montana
Idaho
Nevada-Utah-Wyoming
Colorado
Arizona-New Mexico
Washington
Oregon
California

6. South Central

Kentucky
Tennessee
Alabama
Mississippi
Arkansas
Louisiana
Oklahoma
Texas

VITA

Steven Darrell Mundy was born in Hickory, North Carolina, on October 4, 1942. He is the oldest son of Mr. and Mrs. Wade Mundy. He attended Oxford Elementary School, Oxford, North Carolina. He graduated from Bunker Hill High School at Claremont, North Carolina, in 1961. The following September he entered North Carolina State University at Raleigh, North Carolina. He received a B.S. degree in agricultural engineering and soils technology in 1965 and a M.S. degree in agricultural economics in 1970.

He was married to the former Sandra Wess Kennedy of Hillsborough, North Carolina, on August 7, 1966.

He taught vocational agriculture in 1967-1968 at Western High School, Elon College, North Carolina.

In September, 1968, he accepted a National Defense Education Act Title IV fellowship in the Department of Agricultural Economics and Rural Sociology at the University of Tennessee at Knoxville. He completed the Doctor of Philosophy degree in June, 1972, with a major in agricultural economics and minors in economic theory and econometrics.

The author is a member of FarmHouse, Gamma Sigma Delta, Alpha Zeta, The American Agricultural Economics Association, and The Southern Agricultural Economics Association.

He began employment as an instructor in the Department of Agricultural Economics and Rural Sociology at the University of Tennessee in March, 1971. After graduation he will be employed as an assistant professor in the same department.