

AN ECONOMETRIC MODEL FOR THE
STATE OF OKLAHOMA

By

TIMOTHY CHARLES IRELAND

Bachelor of Science
Phillips University
Enid, Oklahoma
1974

Master of Science
Oklahoma State University
Stillwater, Oklahoma
1976

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
December, 1978

Thesis
1978D
F65e
Cop. 2



AN ECONOMETRIC MODEL FOR THE
STATE OF OKLAHOMA

Thesis Approved:

Gerald M. Lage

Thesis Adviser

John D. Rea

Ronald L. Moomaw

Dayll E. Ray

Norman N. Deukhan

Dean of the Graduate College

1032376

ACKNOWLEDGMENTS

I wish to extend my appreciation to Dr. Gerald M. Lage, Chairman of my graduate committee and Head of the Department of Economics, for his assistance in completing this study and his guidance in my graduate program. His willingness to sacrifice his own time in order to help in the solving of various problems is appreciated by all graduate students as well as myself. A special debt of gratitude is also owed to my other committee members, Dr. John D. Rea, Dr. Ronald L. Moomaw, and Dr. Daryll E. Ray. Their time and valuable suggestions were certainly appreciated.

I am very grateful to the Kerr Foundation, Inc. and the Office of Business and Economic Research at Oklahoma State University for providing the funding necessary to undertake this study. In addition, the Department of Economics is owed a debt of gratitude for providing financial support during my graduate program.

I wish to thank my parents, Dr. and Mrs. John B. Ireland, for their understanding and support of my educational endeavors through these many years of schooling. Their encouragement has always been deeply appreciated.

Finally, the greatest debt of gratitude is owed to the author's wife, Sandi Ireland. Her sacrifice, patience, and understanding along with her cheerful contribution as a typist have enabled this undertaking to become a reality.

TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION	1
Introduction.	1
Objectives.	6
Organization.	7
II. TRENDS IN REGIONAL MODELING.	9
Introduction.	9
Linking Regional and National Models.	11
Alternative Theoretical Specifications.	13
Naive Form	15
National Modeling Form	16
Recent Modeling Form	19
Quarterly vs. Annual Data Use	22
Review of Selected Models	25
Massachusetts.	26
Ohio	27
Northeast Corridor	28
Georgia.	30
Mississippi.	33
Tennessee.	35
Philadelphia SMSA (Philadelphia IV).	37
Basic Characteristics	41
III. EQUATION SPECIFICATIONS.	43
Introduction.	43
Region Definition	43
Model Decisions	45
Suggested Model Structure	46
Model Block Theory.	47
Output	48
Employment	50
Wage and Salary.	54
Personal Income.	57
Tax Revenue.	58
Miscellaneous.	60
Proposed Oklahoma Model	60
IV. MODEL ESTIMATION AND TESTING	63
Introduction.	63
OLS Model	67

Chapter	Page
Output	68
Employment	73
Wage and Salary.	81
Personal Income.	89
Tax Revenue.	95
Miscellaneous.	100
Model Properties	102
Ex Post Simulation	104
Ex Post Forecasting.	111
Simultaneous Equation Model	116
Output	121
Employment	123
Wage and Salary.	126
Personal Income.	129
Tax Revenue.	131
Miscellaneous.	133
Model Results.	133
Ex Post Simulation	134
Ex Post Forecasting.	140
Comparisons of OLS and INST Models.	143
Multiplier and Impact Elasticity Analysis	144
V. FORECAST FOR OKLAHOMA.	151
Introduction.	151
Forecast Assumptions.	153
Forecast Results.	156
VI. ALTERNATIVE FORECASTING TECHNIQUE.	164
Introduction.	164
Box-Jenkins Time-Series Technique	166
Moving Average Models.	167
Autoregressive Models.	168
Mixed Autoregressive-Moving Average Models	168
Homogeneous Nonstationary Processes.	168
Model Construction	170
Time-Series Models for Four Variables	172
Gross State Product.	173
Non-Agricultural Employment.	174
Wage and Salary Income	175
Personal Income.	176
Forecasts Using the Time-Series Models.	177
Composite Forecasting	177
Estimation of Composite Models	180
Forecasting with Composite Models.	182
VII. SUMMARY AND CONCLUSIONS.	185
Summary and Conclusions	185
Suggestions for Future Research	190

Chapter	Page
BIBLIOGRAPHY.	194
APPENDIXES.	200
APPENDIX A - DATA SOURCES.	201
APPENDIX B - GOLDBERGER ALTERNATIVE FOR 1961-1975 OLS SIMULATION	205
APPENDIX C - 1976-1981 FORECAST WITH GOLDBERGER ALTERNATIVE.	208

LIST OF TABLES

Table	Page
I. A Statistical Summary of the Oklahoma Model and Other Selected Regional Econometric Models.	103
II. Oklahoma Econometric State Model OLS Simulation Error Statistics, 1961-1975	106
III. OLS MAPE Error Distribution	109
IV. Comparison of MAPE Errors Over Selected Variables and Regional Models	110
V. Percentage Distribution of MAPE Errors from Sample Period Simulation	112
VI. 1976 OLS Simulation Errors.	113
VII. 1976 OLS Simulation Errors: Goldberger Alternative	117
VIII. Oklahoma Econometric State Model INST Simulation Error Statistics, 1961-1975	135
IX. INST MAPE Error Distribution.	139
X. 1976 INST Simulation Errors	141
XI. A Time Dimension of Impact Elasticities for the OLS Model Due to 1% Changes in Six National Variables	146
XII. National Exogenous Variables Assumptions.	154
XIII. State Exogenous Variables Assumptions	155
XIV. Oklahoma Forecasts of Selected Variables Assuming Control Solution Plus Carter Energy Program	157
XV. Forecasted Increases of Selected Variables from 1976-1981	159
XVI. ARIMA Model Forecasts for Four Variables, 1976-1981	178
XVII. Percentage Increases from 1976-1981 of the Four Forecasted Variables from the ARIMA Process	178

Table	Page
XVIII. Composite Estimation Results.	181
XIX. Composite Model Forecasts for Four Variables, 1976-1981 . .	183
XX. Percentage Increases from 1976-1981 of the Four Forecasted Variables from the Composite Process	183
XXI. Comparison of 1976 MAPE Errors Across Forecasting Techniques.	183

FIGURE

Figure	Page
1. Flow Chart of the Oklahoma Economy	61

LIST OF SYMBOLS

ABRI	-	Alcoholic beverage rate index (percent)
ABT	-	Alcoholic beverage taxes (millions)
ABWPI	-	Alcoholic beverage wholesale price index (percent)
CCE	-	Contract construction employment (thousands)
CCE1	-	Lagged contract construction employment (thousands)
CCWS	-	Contract construction average annual wages and salaries (thousands)
CLFA	-	Civilian labor force (thousands)
CLFA1	-	Lagged civilian labor force (thousands)
CRI	-	Cigarette rate index (percent)
CSSY	-	Contributions to social insurance (millions)
D	-	Fiscal year disposable personal income (millions)
DD	-	Demand deposits (millions)
DIRY	-	Dividend, interest, and rent income (millions)
DPYA	-	Disposable personal income (millions)
DPYAD	-	Real disposable personal income (1972 millions)
FEDMW	-	Federal minimum wage
FGCE	-	Federal government civilian employment (thousands)
FGCE1	-	Lagged federal government civilian employment (thousands)
FGCWS	-	Federal government civilian average annual wages and salaries (thousands)
FIREE	-	Finance, insurance, and real estate employment (thousands)
FIREWS	-	Finance, insurance, and real estate average annual wages and salaries (thousands)

FIREWSD - Real finance, insurance, and real estate average annual wages and salaries (1972 thousands)
 FIT - Federal income taxes (millions)
 FMWS - Federal military wages and salaries (Millions)
 FPI - Farm proprietors income (millions)
 FRDR - Federal Reserve discount rate (percent)
 FWS - Farm wages and salaries (millions)
 GPI - Gasoline price index (percent)
 GPT - Gross production taxes (millions)
 GSP - Gross state product (1972 millions)
 HLE - Hired labor expense (millions)
 I - Fiscal year implicit price deflator for consumption
 IPDC - Implicit price deflator for consumption
 MAAA - Moody's Aaa bond rate (percent)
 MFGE - Manufacturing employment (thousands)
 MFGE1 - Lagged manufacturing employment (thousands)
 MFGWS - Manufacturing average annual wages and salaries (thousands)
 MFT - Motor fuels taxes (millions)
 MINE - Mining employment (thousands)
 MINE1 - Lagged mining employment (thousands)
 MINWS - Mining average annual wages and salaries (thousands)
 MVET - Motor vehicle excise taxes (millions)
 MVWPI - Motor vehicle wholesale price index (percent)
 NAE - Non-agricultural employment (thousands)
 NPY - Non-farm proprietors income (millions)
 NPY1 - Lagged non-farm proprietors income (millions)
 OASI - Social security tax rate (percent)
 OIWS - Other industries wages and salaries (millions)

OKAUNR - Oklahoma unemployment rate (percent)
 OKAUNR1 - Lagged Oklahoma unemployment rate (percent)
 OKIUNR - Oklahoma insured workers unemployment rate (percent)
 OKOLWS - Ratio of other labor income to total wage and salary income
 OLY - Other labor income (millions)
 OTHER - Other taxes (millions)
 P - Fiscal year personal income (millions)
 PCNAE - Percentage change in non-agricultural employment (percent)
 PO - Fiscal year state population (thousands)
 POP - State population (thousands)
 PYA - Personal income (millions)
 PYAD - Real personal income (1972 millions)
 QCC - Contract construction output (1972 millions)
 QFARM - Farm output (1972 millions)
 QFG - Federal government output (1972 millions)
 QFG1 - Lagged federal government output (1972 millions)
 QFIRE - Finance, insurance, and real estate output (1972 millions)
 QFIRE1 - Lagged finance, insurance, and real estate output (1972
 millions)
 QMFG - Manufacturing output (1972 millions)
 QMFG1 - Lagged manufacturing output (1972 millions)
 QMIN - Mining output (1972 millions)
 QSER - Service output (1972 millions)
 QSLG - State and local government output (1972 millions)
 QSLG1 - Lagged state and local government output (1972 millions)
 QTCPU - Transportation, communication, and public utilities output
 (1972 millions)

QWRT - Wholesale and retail trade output (1972 millions)
 QWRT1 - Lagged wholesale and retail trade output (1972 millions)
 RAY - Resident adjustment factor (millions)
 RELAUNR - Relative unemployment rate
 RGDPG - Real gross domestic product government (1972 billions)
 RGDPMA - Real gross domestic product manufacturing (1972 billions)
 SE - Service employment (thousands)
 SE1 - Lagged service employment (thousands)
 SERWS - Service average annual wages and salaries (thousands)
 SLGE - State and local government employment (thousands)
 SLGE1 - Lagged state and local government employment (thousands)
 SLGWS - State and local government average annual wages and salaries (thousands)
 SLGWSD - Real state and local government average annual wages and salaries (thousands)
 ST - Total state taxes (millions)
 SUT - Sales and use taxes (millions)
 TCPUE - Transportation, communication, and public utilities employment (thousands)
 TCPUE1 - Lagged transportation, communication, and public utilities employment (thousands)
 TCPUWS - Transportation, communication, and public utilities average annual wages and salaries (thousands)
 TCPUWSD - Real transportation, communication, and public utilities average annual wages and salaries (thousands)
 TD - Time deposits (millions)
 TIME - Time in years
 TL - Total loans (millions)
 TPT - Tobacco products taxes (millions)

TRY	-	Transfer payments (millions)
TSIT	-	Total state income taxes (millions)
UDIRY	-	U. S. dividend, interest, and rent income (billions)
USMFGW	-	U. S. manufacturing average annual wages and salaries (thousands)
USTRY	-	U. S. transfer payments (billions)
USUR	-	U. S. unemployment rate (percent)
WRTE	-	Wholesale and retail trade employment (thousands)
WRTE1	-	Lagged wholesale and retail trade employment (thousands)
WRTWS	-	Wholesale and retail trade average annual wages and salaries (thousands)
WRTWSD	-	Real wholesale and retail trade average annual wages and salaries (thousands)
WSY	-	Total wage and salary income (millions)
Z	-	Relative labor costs

CHAPTER I

INTRODUCTION

Introduction

Forecasting economic activity on the regional level has become a significant facet of regional economic research and study. Regional policymakers, both on the state and local level, need accurate forecasts of many important economic variables in order to derive proper plans or courses of action in the future. These key variables usually involve output figures, employment conditions, income levels, and tax revenue possibilities.

Two major types of models have been developed and used in the past to forecast some of these important regional variables. These well-known techniques are economic base and input-output models. Both techniques suffer from certain deficiencies which make their use a difficult matter in regional analysis.

Economic base theory is developed around the notion that a region's activity level is determined primarily by its level of exports. The regional economy is divided into two primary sectors, the basic and nonbasic. The basic or export sector is the motivator which allows exchange to take place with other regions and ultimately lead to growth and development in the region. The nonbasic or service sector is seen as portraying a supportive role to the basic sector and can grow itself only in response to expansion in the basic sector. Economic base models

are very simple to construct and very inexpensive to undertake but they are beset with many conceptual and technical problems: the use of improper units of measurements, imprecise identification of sectors, weak assumptions concerning the stability of the basic/service ratio, and the problem of lags (24, p. 20).

A second type of model, the input-output variety, is much more complex in form and structure than the naive economic base model. This theoretical model states that each producing sector is dependent upon the activity of every other producing sector in the economy and it uses fixed technical relationships in its construction. An output of one sector may be used as an input for another sector and vice-versa. This model allows one to more closely follow the ripples that are caused by an exogenous shock to the economic system. In contrast, the economic base model can judge the impact of a shock to an economic system but not the manner in which effects are passed on from one industry to another.

Input-output models are also plagued by certain inherent problems. The use of constant production coefficients or fixed technical relationships between sectors effectively rules out any possibility of economies of scale. Since localization and urbanization economies should be accounted for in regional analysis, this assumption presents quite a perplexing problem to regional analysts (24, p. 34). In the short-run, the assumption of these fixed relationships is not so serious but with any desire of long-term forecasting in mind, the seriousness increases. The amount and type of data needed for construction of an input-output model also poses potential problems. Data on origination and destination of sales and purchases for each firm in a region is a prerequisite for this type of research. Since data of this type is not collected for

regions it must be collected on a primary basis. This can be done only at a great cost. Substituting national coefficients into a model, instead, enters a certain amount of bias into a study and may reduce any results to utter nonsense.

To counteract the deficiencies suffered by the previously mentioned models, new types of models and modeling have been established in the regional area. The most notable of these varieties is the econometric model. An econometric model is a structure (sometimes simultaneous in nature) which estimates macroeconomic relationships from historical or time series data by the use of regression techniques. The validity of the derived relationships are then judged by the use of special statistical tests.

Econometric models are composed of a series of equations which attempt to quantify cause-effect relationships among economic variables. A typical equation can be visualized in the following manner:

$$Y_{it} = f(Y_{jt}, X_{kt}, \epsilon_t) \quad (1.1)$$

where Y_{it} = the i^{th} endogenous variable in time period t ,
 Y_{jt} = the j^{th} endogenous variable in time period t ,
 X_{kt} = the k^{th} exogenous variable in time period t , and
 ϵ_t = the random error term in time period t .¹

An entire model or system of m equations could be denoted in matrix notation as:

$$Y\beta + X\gamma = E \quad (1.2)$$

¹Endogenous variables are ones whose values are determined within the model. Exogenous variables are pre-determined or determined outside of the model.

where β = a $m \times m$ nonsingular matrix of coefficients for the endogenous variables,

Y = a $t \times m$ matrix of endogenous variables for each time period,

γ = a $k \times m$ matrix of coefficients for the exogenous variables,

X = a $t \times k$ matrix of exogenous variables for each time period, and

E = a $t \times m$ matrix of random error terms for each time period
(assumed to have the normal regression properties).

By finding the reduced form of the model when it is linear in nature or through the use of iterative techniques when the model form is non-linear, the model, once estimated, can be used to simulate and forecast into the future.² Mandatory to this accomplishment is the availability of forecasts of future values of the exogenous variables in order to drive the model. Various assumptions and possibilities can be entered into an analysis of the future through these exogenous variables. Parameter changes in the model can also be an important component of study in this light.

Econometric models appear to be a good compromise between economic base models and input-output models (24, p. 38). Economic base models are relatively cheap to construct but they relate very small amounts of information to the analyst. Input-output models, on the other hand, are very complex and relate much greater amounts of information but only at a much greater cost. Econometric models fit between the two in both the areas of information generated and costs to undertake.

Econometric models were initially developed primarily for national economies. The Klein-Goldberger Model of the United States (37) was the

²The reduced form of a model involves manipulating the equations so that each endogenous variable can be written as a function of all the exogenous variables.

first of its kind for this nation. It was a broad model of the economy developed from data over the 1929-1952 period except for the war years. It has been followed by many other models of the United States' economy (17) (22) (23) (50) (57) (58). National econometric models also exist for many of the developed countries and are now being extended and constructed for many developing countries.

Many of the national models of the United States have been shown to have very good forecasting credentials (2) (15) (21) (47). This fact has added prestige to the econometric model in the search for a forecasting tool, especially on the regional level, that overcomes the deficiencies of other forecasting techniques (like input-output and economic base models). Norman J. Glickman (25, p. 1) has accented the belief that econometric models can provide an excellent tool for regional research by saying: "In part, due to the reasonably good performance of their national counterparts, regional econometric models have been seen as a means of fulfilling these needs."

The development and use of forecasts from regional or local econometric models has not seen as much usage considering the extensive possibilities. The principal reason for the slower development of the regional econometric model has been due to the lack of large amounts of appropriate data. Consistent time series for regional variables can usually be collected back for only 20 to 25 years. This is in contrast to the data accessibilities for national variables which date back 50 to 60 years. Another problem in the regional model area concerns the fact that some key variables in an economic analysis are not collected at all on the subnational level. Data on regional imports, exports, and non-manufacturing investment are virtually nonexistent.

Even with these obstacles to regional econometric model building, some states have now developed econometric models which seem to be quite good (1) (6) (11) (12) (31) (39) (41) (43). Modeling on the subnational level has also been extended to the multi-state as well as the sub-state level (16) (24). The forecasting abilities for many of these sub-national models are unknown since many of these studies have failed to report this area of investigation.

Oklahoma has had one previous attempt at modeling the State economy (42). This model was extremely aggregative in nature and did not provide any private industrial breakdown in its structure. It consisted of 25 behavioral equations and eight identities. Only one exogenous variable was present in the model and it involved state population. This model involved no interaction with the national economy. This latter fact removes the possibility of investigating the effects on the state economy from a change in the national policy of the country. Also, the extreme aggregative structure of the model prevents the determination of the effects on each major industry from a state or national policy change.

Objectives

There is a definite need for a good econometric forecasting model in the State of Oklahoma. Good in the sense that it shows capable forecasting credentials, allows for interactions or influences from both the state and national level, and is disaggregative enough to trace out effects on the primary industries. Evidence on the national level and the findings from various state models show that the econometric model can do quite well as a forecasting tool.

Such a model could be used to simulate significant functions of the state's economy and provide a method for more accurately forecasting future economic activity within the state. It could definitely be used for deducing probable impact effects of various governmental policy changes. Such a device should be of great interest to many state agencies which are involved with important decision making. This study will attempt to develop a model which has these characteristics and is able to perform these important functions.

The major objectives of this study are to:

A. Estimate a State econometric model for Oklahoma. This will involve:

1. Using standard economic theory in its construction; and,
2. Performing said estimation with acceptable regression techniques.

B. Test the estimated model to determine its ability to replicate economic activity. This will include:

1. Ex post simulation; and,
2. Ex post forecasting.

C. Use the tested econometric model to produce a future forecast for the State of Oklahoma.

D. Compare the predictive abilities of the econometric model with those of alternative time-series techniques like Box-Jenkins.

Organization

Following this introduction, Chapter II discusses the development and general trend in regional econometric modeling. This discussion contrasts the national modeling approach with what is practical and

attainable on the regional level. The general structure of the Oklahoma model is presented in Chapter III. The theory behind the structure of the blocks in the model is explained at that point. Chapter IV contains the actual estimated model and the testing of its replication abilities. A sample forecast of expected future activity is presented in Chapter V. This forecast derived from national forecasts incorporating a possible energy proposal. Chapter VI presents an alternative time series estimation procedure for a selected number of variables and compares the forecasting abilities of this technique with the econometric model. Chapter VII presents the summary and conclusions suggested by the study.

CHAPTER II

TRENDS IN REGIONAL MODELING

Introduction

The national econometric forecasting models evolved from a structure that closely follows the design of the National Income and Product Accounts. This design is, of course, Keynesian in nature. The point from which the modeling format begins is the basic expenditure identity:

$$Y = C + I + G + X - M \quad (2.1)$$

where $Y = \text{GNP}$,

$C = \text{consumption spending,}$

$I = \text{investment spending,}$

$G = \text{government spending,}$

$X = \text{export activity, and}$

$M = \text{import activity.}$

Some detail is then provided in the explanation of consumption expenditures, investment expenditures, government expenditures, exports, and imports. Equations and section blocks can also be found to explain taxation, production function relationships, wage determination, unemployment conditions, price levels, income variables, monetary and financial sectors, and other relevant factors in the economy. Data for these variables are readily accessible on the national level and can many

times be found on a quarterly basis as well as the usual annual conditions.

While national modeling may seem to be fairly straight forward, regional modeling is not nearly as cut and dried. The availability of data as to the type of variable as well as whether it is recorded on both a quarterly and annual basis introduces the prospect of many alternative forms to the regional modeling effort. Additional possibilities also enter the picture when the appropriate linking procedure between regional and national models is considered. Summed up, regional modeling is much less clear cut, and, therefore, open to various approaches.

The possibilities that need to be examined when considering regional econometric modeling can be partitioned into three areas. These areas need not be mutually exclusive. The first facet of modeling on the regional level that must be scrutinized involves the question of how the regional models should be linked to the national models. It turns out that this is not a very big issue since the origination of modeling on the national level has essentially decided the matter. The second and possibly most prominent issue involves the degree to which economic theory and the national modeling approach are used in the specification of the regional econometric model. Closely aligned with this issue is the third matter or question about the availability and the time framework of the data to be used in the modeling effort.

Each of these issues will be examined as to the possibilities it opens to the regional modeler. After discussion of these questions, some typical regional models that have been constructed will be briefly discussed. A list of general characteristics pertaining to regional econometric models will also be presented.

Linking Regional and National Models

In building and developing regional econometric models and tying them in with national models, the question arises as to whether a "bottom-up" or "top-down" approach should be taken (36, p. 4). Specifically, this involves whether you start your linkages at the national level and work down to the regional step or begin at the regional level and work up to the national totals.

The "top-down" technique involves constructing regional models so as to act like satellites to the already existing national models. Certain variables in the regional model are made dependent on national variables to make this national-regional linkage. National changes then flow down from the top through this linkage to interact with the regional variables. The advantage to this method is that the national models are already in existence and only the satellite regions have to be spun off. The disadvantage to this is that no feedback effects are allowed from the regional model to the national model. For regions that are fairly small, this disadvantage is not of any consequence. However, for much larger areas that do have sustantial feedbacks into the national economy, this "top-down" linkage approach may be quite limiting.

The counterpart technique for linking these models is the "bottom-up" method. This involves developing models for all regions in the nation and then summing their findings into national totals. Lawrence Klein and Norman Glickman (36, p. 5) have noted that, "This approach is clearly more satisfying to the regional researcher, since it enables distance and other spatial variables and relations to enter the model in a meaningful way." Feedback effects and interdependencies

play a very strong role in this linkage approach of the regional to the national level. The primary drawbacks to this construction concern the availability of regional data. Some variables that are reported on the national level are not available on the regional level or are of much poorer quality. This means that by starting with the regional model and building up, you may have to forego certain crucial variables on the national level (because the regional counterpart does not exist) or you may have to use constructed national variables of lower quality than those actually published (due to the regional data series on which you build being less sound).

Of the two methods, the "top-down" approach has been much more widespread in usage and popularity. Again, the major strength of this method is the ability to plug in regional models to already established national models. This is much easier than starting at the bottom and having to construct all regions before one ever has a national framework. Data constraints on the regional level add to the prestige of using the satellite modeling approach on the regional level.

The "top-down" approach to regional modeling has also been espoused by many of the leading authorities on economic modeling. Lawrence Klein (35) in his memorable article on specifying regional models, remarked that:

We have gone far in the building of big systems, at least in relation to present custodial capabilities, and it seems to me that we should try to fill the requests for added detail by building many moderate-size satellite systems instead of trying to put Walras in numbers (p. 105).

So the issue of the appropriate regional-national linkage is seemingly already decided. National models already exist and have at their disposal a wealth of high quality data. The sensible conclusion would

be to create a "top-down" interlinkage where the regional model acts as a satellite to the existing national structure. Certain regional spatial factors may be sacrificed by the usage of this approach but the alternative seems to be even more undesirable. A significant majority of the existing regional models follow this format.

Alternative Theoretical Specifications

The regional modeler, after having determined the appropriate linkage to be made with the national framework, must turn his attentions toward the question of theoretical specification. Since there has already been much work completed on national models, a theoretical specification similar to that of the national models might be considered. If this is not desirable then possible variations concentrating on regional classifications and interests might be examined. Here, as well as of interest in other matters, is the importance of the type and amount of data that is available. Donald Ratajczak (51) has studied this issue and found that:

Econometricians engaged in regional research have continually been confronted with the choice of adapting their models to the available data or constructing data to appropriately specify their models. On the one extreme, econometric techniques have been used to explain the historical performance of available data series. Accounting identities are not used, nor is a consistent economic system specified. At the other extreme is the attempt to create subsystem 'Keynesian' expenditures models even if the expenditures data are unavailable (p. 51).

Although these two extremes certainly present interesting regional modeling specifications, most of the actual models developed on the regional level probably exist between these two points.

No two regions, states, or locales are exactly identical as to the economies they have or to the depth and quality of data they possess

concerning their economy. This notion makes it very difficult to name any existing specification or system as the one most appropriate for a given region's modeling efforts. In fact to compound things, Ratajczak (51, p. 51) notes that; "The tradeoff between data construction and model modification varies with the economic system that is used to describe regional economic activity." In essence, it is no easy matter to specify a regional model without close scrutinization of what data is available, what data can be constructed without incurring large measurement error, and what exactly is desired as output from the model.¹

Even though it seems there would be an infinite number of possibilities for regional modeling specifications (as to data, theory, output, etc.), Ratajczak (51) is able to group all of the previous regional econometric modeling efforts into three divisions:

1. An explanation of prevailing data by whatever variables are available;
2. Strict conformity with specified theory even if substantial data must be constructed; and
3. Some implicit tradeoffs between theory and data so that measurement error and specification error are jointly minimized (p. 56).²

The above categories will be examined individually as to what advantages and disadvantages they may contain. The first division or naive form would appear intuitively to be the weakest of the three and

¹Measurement error is the error that is introduced into an analysis from an equation which involves a variable(s) that has not been correctly measured. This could be due to use of approximating techniques in construction or to mistakes in recording.

²Specification error is the error due to a regression equation misspecifying the true relationships or cause-effect condition among variables. This can be caused by omitting relevant variables or by including irrelevant variables in the equation.

without as much substance as a realistic approach. The second grouping is the one that most closely follows the form used in the national models. It would be expected that the initial attempts at regional modeling would have used this approach. The third technique incorporates within economic theory the fact that certain problems with data may make the national approach unattainable on the regional level. The costs of constructing an expenditure framework may far outweigh the benefits received. This final method represents a more recent approach.

Naive Form

There have been a small number of regional models that have allowed the available data the determination of their structure or form (10) (54). These models seem to be very simple in nature and primarily seem to be ". . . seeking maximization of explained variation for each data series" (51, p. 56). Although some theory is used in estimating each equation, ". . . the model is merely a set of single equation estimations" (51, p. 57). There are usually few or no interdependencies in the models and, therefore, little simultaneity in determination or computation of the models.

In this manner, the regional endogenous variables are often determined solely by national variables without any interaction with other regional variables. A typical equation would appear as in the following form:

$$Y_{it} = f(Z_{kt}, u_t) \quad (2.2)$$

where Y_{it} = the i^{th} endogenous variable in period t ,

Z_{kt} = the k^{th} national exogenous variable in period t , and

u_t = the error in period t (24, p. 39).

Equations can be found explaining output, personal income, employment, retail sales, labor force, taxes, and other economic variables of interest.

This specification form has the advantage or ease of not having to construct much data since only what is published is usually used. Hence, measurement error should be virtually nonexistent. This simplicity in development, though, is likely outweighed by the many problems it also poses. Policy impact analysis is limited in a model like this. Since no consistent interacting economic system is designed in this approach, one cannot examine how a change in policy will be transmitted through various sectors to affect the regional variables. Also, misspecification of some equations may occur from not allowing for the effect of local conditions in the structural determination. Results can become nonsensical. These facts, if your original intent in modeling is to develop a forecasting and policy analysis model, make this approach of little real value.

National Modeling Form

A second modeling form is that of specifying the regional model in a manner very similar to that of the national models. This form seemed to predominate in the early regional modeling attempts and was strongly advocated by Lawrence Klein in a 1969 article (35). Even though fairly strict conformity with specified theory is advocated in this approach, regional models are allowed to incorporate features that are special to

their own region. Klein (35, p. 108) emphasized this by saying, "The typical regional macromodel will be similar to the national macromodel but will have some characteristics of its own." This macromodeling approach for regions was believed desirable because it could easily be linked to the national models.

Klein proposed the satellite modeling approach that was previously mentioned. Regional models designed in this manner could be tied into national models and would use exogenous variables determined by the national structure.

As with national models, Klein adopts an aggregate income-expenditure approach in a national income accounts framework for regional constructs whereby gross regional product (GRP) is the sum of its components: consumption, investment, government, and net exports (24, pp. 56-57).

This can be visualized as:

$$\text{GRP} = \text{C} + \text{I} + \text{G} + (\text{X} - \text{M}) = \text{GRO} = \text{GRI} \quad (2.3)$$

where C = consumption,

I = investment,

G = government,

X = exports,

M = imports,

GRP = gross regional product,

GRO = gross regional output, and

GRI = gross regional income (24, p. 57).

In his suggestion for a regional model, Klein included equations to explain regional consumption, investment, government expenditures, exports, imports, direct state and local taxation, indirect state and local taxation, federal taxation, transfer payments, capital consumption,

disposable income, production, price levels, export prices, wage rates, capital stock, nonwage income, and unemployment. It was stated that the models should be dynamic and that the degree of detail or disaggregation within each category would depend upon the region.

Simultaneous relationships were specified in this approach. Some of the interdependencies involved disposable personal income determining consumption; gross regional product affecting investment, imports, and indirect regional taxation; employment influencing nonwage income, unemployment, state and local direct taxation, federal taxation, and gross regional product; and numerous other economic variables. This specification of these interacting forces gives this model a more viable, functioning economy.

The strongest point to this modeling form, ". . . is the ability to use a well developed behavioral theory in regional analysis" (51, p. 57). Since much theory has been developed on many of the key expenditure components (i.e., consumption, investment, etc.), these beliefs can be directly incorporated and tested on the regional level. Other advantages are the ease with which regional models can be joined to national models as well as the enhanced ability to directly trace national policy changes to the regional level. Ratajczak (51, p. 58) has noted that, ". . . as long as the national changes affect the region more than the region's internal activity, the expenditures approach will improve the degree of anticipating regional changes."

Although this approach seems very good at first sight, it is beset by one very formidable problem. Data for some of the components of the expenditure framework do not exist on the regional level. Consumption expenditures, non-manufacturing investment expenditures, exports, and

imports rarely exist data-wise on this level of activity. Attempts at creating such information can be made but, ". . . the necessary data creation is so substantial and the assumptions necessary to develop such information are so limiting that many of the factors that may cause differences in regional development cannot be analyzed" (51, p. 57). In short, on the regional level, the national expenditure-income approach is very fine in theory but not very practical when considering costs and benefits of implementation. In defense of Klein (35, p. 108), it might be noted that he stated at the outset of his discussion on the appropriate regional model that, ". . . I shall not recognize, at this stage, the very real and substantial data problems involved." Unfortunately, this disclaimer does not help overcome the data problems of this method.

Recent Modeling Form

Since the early 1970's, the growing specification trend in regional econometric modeling has been one of tempering the national approach with certain tradeoffs between theory and data. This activity is primarily due to the data problems that are involved in implementing the national-like approach which Klein advocated. This category could conceivably contain all possible combinations of tradeoffs between theory and data that fall in between the naive and national specification forms. Although this category could contain an infinite number of tradeoffs and, therefore, model forms and essentially be only a catchall, it has been dominated by a central approach. This approach, advocated by Norman Glickman and many others, suggests that the regional accounts should ignore the expenditure aspects of the accounting framework and instead concentrate on those stressing output and income (24, p. 57).

This recent trend recognizes the fact that many of the components of the expenditure identity cannot be found on the regional level. Therefore, the expenditure basis must be omitted from the regional specification. The regional accounts can then be structured around gross regional output and/or gross regional income. Although the national accounts show that gross product, gross income, and gross output should all be equal, this need not be so with the remaining elements on the regional level. Since both parties to a transaction do not have to be located within the boundaries of the state:

$$GRO = \sum_r X_r \neq GRI = \sum_s F_s \quad (2.4)$$

where X_r = output in sector r ,

F_s = factor payment of type s ,

GRI = gross regional income, and

GRO = gross regional output (24, p. 58).

Gross regional output and gross regional income need not be equal.

The gross regional output accounting identity sometimes is the only gross activity level that can be determined for a region. Not only does the expenditure approach become impossible to compute but the inability to obtain corporate profits may make the gross regional income identity useless. Three potential identities for a regional framework quickly reduce to one and data for it are not achieved easily.

Another change to the model specification proposed by Klein (24, p. 58), ". . . concerns relations between the region and the rest of the world." This change, which is in part due to the movement away from the expenditure accounts approach, involves making implicit to the model what Klein specified as explicit. Klein related export

activity in a region to a national activity variable like GNP while imports for the region were linked to local activity (i.e., gross regional output). Since export and import activity is of interest in a region, the elimination of the expenditure accounts would destroy these relationships between the nation and the region. This problem has been overcome in the more recent specification form.

"Instead of specifying such equations, research workers have estimated demand-type equations relating industry output to the relevant output markets" (24, p. 59). Export-oriented industries in the output sector have equations that are related to national activity, just as an export equation would in a national model. In a similar manner, industries that are more local in nature have their outputs tied to local activity variables like personal income or gross regional output. Hence, linkages between the nation and the region that would arise due to exporting activity are found in the output equations of the exporting industries. Import associations are likewise located in the output specifications of residentiary industries. "Thus, the lack of data has resulted in the implicit specification of activity that ought to be explicit" (24, p. 59).

Although techniques may differ in the way the rest of the key variables of a regional model are formulated, the recent approach does simultaneously model personal income activity, employment by sectors, unemployment, taxation, wage and salary determination, financial categories, demographic variables, etc. No significant differences (as in the expenditure accounts controversy) separate the forms to modeling the rest of the economic sectors.

The most recent trend in regional econometric modeling has been one of trying to overcome the data deficiencies that mar the national approach to regional modeling. Expenditure frameworks and thoughts of final demand information have been replaced by structures that can be formed with existing data. Some economic theory must be sacrificed when using this form but that may be better than not being able to construct a good regional forecasting model because of the bad quality data used in its estimation. Little may be lost by centering analyses around employment and output changes rather than expenditure movements. Ratajczak (51, p. 62) has possibly best summed up the intent behind the trend away from the national modeling approach by saying, ". . . while the creation of a theoretical framework without regard to data factors may be useful in stimulating the search for new data sources, it provides little comfort to the regional analyst."

Quarterly vs. Annual Data Use

Another important area of interest when considering the construction of a regional econometric model involves the choice of time reference for the data to be used. Specifically, should annual or quarterly data be used in the estimation? This decision will be affected by the degree of theoretical content desired in the model as well as certain desired statistical and regional features.

Annual data are very abundant on the regional level for most economic variables. Quarterly data, while now available for many economic variables, are not quite as common. The best source of quarterly data for a state or region is provided from statistics on the insured labor force. Wage rates, employment, and unemployment numbers

can be located for this group. These figures may unfortunately omit three other groups: self-employed workers, government employees, and agricultural workers (51, p. 52). Personal income also exists on a quarterly basis and most of the wage and salary sector can be formed from the insured labor force information. Many other significant economic variables can be located or constructed on a quarterly basis with one major exception. Output data are only available on an annual basis. This omission may be very crucial in the choice of data selection for a regional model.

If an explicit output sector is desired in a regional model then only annual data can be used in the modeling. The issuance of value-added data for manufacturing as well as the components to construct gross product originating by industry via the Kendrick-Jaycox technique (34) are only available on an annual basis. An explicit output sector precludes the use of quarterly data. If an output sector is not of upmost importance in the modeling effort then quarterly data may be considered. The inclusion of an output sector does make for a more theoretically complete model.

An interesting question concerning the output sector arises when considering this quarterly vs. annual data confrontation. Although an output sector may make a model more solid theoretically in its linkages, it does involve certain problems. These problems involve the fact that there is a two to three year lag in the availability of value-added data on output for the manufacturing sector of the economy. This means that models which explicitly incorporate output sectors (of which manufacturing is usually very prominent) are actually making four to five year forecasts in the future when they forecast two to three years from

the present. As with any model, the longer one forecasts into the future, the greater ones chance of error and the compounding of it. The dilemma arises in whether one wants to cut off ones data series two to three years before most of them end against the ability to have actual forecasts of output levels in the regional economy. Although more theoretically satisfying, the inclusion of an output sector may not only preclude the usage of quarterly data but also reduce the length of the existing annual data series and enhance forecasting error.

Besides the output consideration, regional modeling can probably be achieved using quarterly data over annual data with only a few minor drawbacks. There are certain model characteristics involving various strengths and weaknesses of each that may aid in the decision of which data approach is more advantageous. An econometric forecasting model for the State of Delaware has established a list of the costs or drawbacks associated with using each approach.

Major drawbacks associated with annual models which are avoided when quarterly models are constructed include the following: 1) Annual models have limited numbers of degrees of freedom because long time series on regional variables are not available. . . The consequences of a limited number of degrees of freedom are: a) a reduction in the statistical reliability of estimated coefficients and b) a reduction in the number of explanatory variables in any equation . . . 2) Annual models cannot reflect more rapid than annual adjustments to changing economic conditions and thus obscure the sensitivity and speed of adjustment of the regional economy to economic changes . . . 3) Annual models can only adjust for fiscal year phenomena, such as federal, state, and local government budgets with some difficulty (39, pp. 3-4).

The annual model problems are then contrasted to those from quarterly usage.

The consequences of choosing a quarterly specification also include several costs: 1) Quarterly time series data are available for fewer economic magnitudes than are annual data. This is a particularly serious problem when specification of the model proceeds from a set of regional product accounts

which require regional output data that are available only on an annual basis . . . 2) By specifying a quarterly model we are forced to provide a much more detailed description of the behavior of the sectors modeled than an annual model requires . . . 3) The higher variance of quarterly than annual totals for many economic magnitudes means that our standard errors of estimate for some equations will be higher than would be the case in an annual model (39, pp. 4-5).

The decision on the appropriate usage of annual or quarterly data swiftly becomes a crucial issue when considering the strengths and weaknesses of each approach. If an output sector is desired (and there are many good reasons for including one), then the matter is already decided. Annual data has to be used. If an output sector is not desired, then the issue is wide open. Reasons for and against can be found for each technique. The more complete annual models when including the output sector probably produce more accurate forecasts of impacts on the regional economy. In contrast, the quarterly models are potentially more sensitive to short-run cyclical fluctuations and seasonal changes. Turning points can probably be better predicted for the economy using the quarterly model. These issues along with those involving the appropriate linkage to the nation and specification alternatives must be considered together when trying to select the appropriate model form for a region.

Review of Selected Models

In this section, a brief review will be made of some of the existing regional econometric models. The models that were selected to be discussed here are considered to typify the past and current research trends in this area. Models will be discussed according to the framework they have developed, the behavioral equations they estimate, the sample period they use for data, estimation techniques, and any other features

of interest. The regional models to be presented cover areas involving: Massachusetts, Ohio, Northeast Corridor, Georgia, Mississippi, Tennessee, and Philadelphia SMSA.³

Massachusetts

The Massachusetts model was developed by Frederick W. Bell (6). This was one of the earliest regional econometric modeling attempts and it signifies this by its resemblance to the national forecasting models. This model is very limited in its features since it is composed of only 14 equations. Of these equations, only eight are behavioral.

Equations were established to explain export income, local consumption income, manufacturing investment, nonmanufacturing investment, production, expected labor supply, migration and real wage determination. The model involves very little simultaneity in the determination of the endogenous variables and seems to be highly recursive in nature. GNP plays a large role in the construction of these equations. GNP determines export income and also indirectly affects local service income. The investment functions, in turn, are determined by export and service income.

The wage determination equation for the region was handled in an interesting manner by making real wages solely a function of time. A Phillips relation was tested earlier and found to be insignificant. Bell (6, p. 116) stated that:

³ Some of these models are initial attempts in ongoing projects. As this project is the initial modeling attempt for Oklahoma, it would seem appropriate to examine these early formulations.

On the basis of the available information, we shall postulate that real wages will increase secularly on the basis of technological change and capital-labor substitution and exhibit no pronounced reaction to unemployment.

The average growth of real wages was estimated to be 1.7 percent per year for Massachusetts.

The model was estimated from annual data over the 1947-1962 time period. There were 16 observations used for the three estimating techniques that were performed: ordinary least squares (OLS), two-stage least squares (TSLS), and reduced-form least squares (RFLS). No error statistics for tests of the ability to replicate economic activity were listed. Glickman (24, p. 46), in a review over certain regional models, questioned the ability of this model to forecast and summed up this modeling effort by remarking, ". . . it is questionable whether this work represents a market improvement over the base model."

Ohio

The Ohio model, constructed by W. L. L'Esperance, G. Nestel, and D. Fromm (41), recognized the data problems involved with the expenditure and income approach to constructing gross state product (GSP) accounts. Therefore, they resorted to the technique developed by Kendrick and Jaycox (34) to construct GSP by industry. The GSP accounts provided the framework upon which this 27 equation model was built.

The Ohio model is composed of three separate blocks. Blocks I and III are recursive while Block II is simultaneous in nature. Block II contains the gross state product components. GSP is composed of activity in contract construction; finance, insurance, and real estate; manufacturing; services; wholesale and retail trade; federal government; state and local government; and a category for other industries. Equations

describing activity in these sectors are found with the exceptions of the government categories. Also included in the simultaneous block are equations for automobile sales, retail sales other than automobiles, personal income, federal income taxes, and the identities to compute disposable personal income. Included in the recursive blocks are information on motor fuels consumption, automobile registration, retail sales tax, retail sales, investment in plant and machinery for manufacturing, and internally generated funds in manufacturing.

The most conspicuous omissions from this model are equations for employment and wage considerations. Nowhere are such matters introduced into the analysis. Also, not much detail is provided on the state taxation situation.

The model was estimated using OLS and TSLS. The 16 behavioral equations and 11 identities were determined from annual data generally occurring from 1949-1963. Tests were made using a Q^2 statistic to measure the accuracy of fit of the actual and predicted endogenous variables over the sample period (5). It is hard to determine how well the Ohio model replicated activity in a relative sense since few of the other regional models use this Q^2 statistic. In general, the Ohio model seemed to adequately determine the gross state product aggregate variables but some of the other variables did not fare as well.

Northeast Corridor

The model for the Northeast Corridor of the United States is the largest model in terms of area to be discussed in this section.⁴ This

⁴The Northeast Corridor consists of New Hampshire, Massachusetts, Rhode Island, Connecticut, New York, New Jersey, Pennsylvania, Delaware, Maryland, the District of Columbia, and Virginia.

ten state area plus the District of Columbia model was established by R. T. Crow (16) in conjunction with a transportation study sponsored by the U. S. Department of Transportation. This 66 equation model is quite novel in that it attempts a triple entry accounting structure. Gross regional product (GRP) is determined through the gross product originating in each of ten industrial sectors, gross expenditures, and gross income. This last approach is not quite as fully specified as the other two.

The ten industrial sectors determining GRP through the gross product originating approach are agriculture; mining; contract construction; manufacturing; wholesale and retail trade; transportation; communication and public utilities; finance, insurance, and real estate; services and miscellaneous; and government. This ten sector breakdown involves not only output but also equations for employment and annual wage rates. A large degree of simultaneity enters into these sectors (output, employment, and wages) as both the gross product originating and annual wage rates act as the prime determinants of employment by sector while the gross product originating by industry helps determine wage rates by industry.

The expenditure sector of the model involves equations for consumption, non-residential fixed investment, residential construction, state government purchases, local government purchases, federal government defense purchases, and federal government non-defense purchases. These equations along with a residual catchall determine the GRP expenditure identity. National and regional variables both play a large role as determinants in this sector.

Income data comprises other labor income, self-employed income, property income, employee contributions to social insurance, and transfer

payments. These variables help compute a personal income and a disposable personal income variable that interacts many times with other parts of the model. Miscellaneous equations also encompass population, consumer price indexes, and the full employment labor force. No allowance seemed to be made for unemployment conditions.

The annual data used for this model ran from 1949 to 1963. The number of observations was expanded by pooling the 15 observations across a three region delineation. The estimation was done using a TSLS approach where the first stage was computed by reducing the exogenous variables into eight principal components. This procedure greatly saved degrees of freedom. Simulation and forecasting were performed with the model with very reasonable results. Mean absolute percentage errors (MAPE) for most variables were acceptable. Since this model was constructed with policy-oriented applications in mind, a simulation was performed for alternative military spending policies.

Georgia

The satellite model for Georgia was originally created by A. Ray Grimes, Jr. (31). This model is also constructed along a gross state product framework in which gross product originating by industry forms the basic identity. This Georgia effort includes a fairly disaggregated breakdown of the manufacturing industry which, therefore, affects the output, employment, and wages equations of the model. This model is one of the larger ones constructed for a region and it contains 114 equations. Eighty-one of the equations are stochastic while 33 are identities. This modeling effort is part of an ongoing project in Georgia and this presentation was the initial effort in that project.

The Georgia model is composed of eight blocks of equations: output; employment; wages, income and prices; state government; manufacturing investment; demographic; banking; and retail trade (31, p. 5). Ample feedback and interdependence are allowed for between sectors and exogenous policy variables also play a major role in this construction.

The real output sector of this model is composed of 19 equations. The traditional industry breakdown into mining; agriculture; contract construction; wholesale and retail trade; finance, insurance, and real estate; transportation, communication, and public utilities; services; government; and manufacturing is included. Also, manufacturing is decomposed into the following two-digit industries: food; textiles; paper; apparel; chemicals; stone, clay, and glass; fabricated metals; transportation equipment; and other manufacturing. "The output equations are demand equations with demand being represented by both interindustry and final demand components" (31, p. 11). Gross state product is determined by the sum of the output of the basic industries.

Employment is computed for the same industries and manufacturing breakdown as in the output sector. The equations are of the labor demand variety where the major industries are determined by industry output and industry average annual wages. Lack of data for the two-digit manufacturing industries' wages prevent the same specification for the manufacturing subsector. The civilian labor force is also estimated in this block with unemployment and the unemployment rate determined through identities.

The wages, income, and prices block involves equations for each. Average annual wages are computed for the major industries and are related primarily to the activity in manufacturing. Manufacturing, itself,

is determined via national wages in manufacturing and mining. Local conditions are allowed for in the individual equations by use of the unemployment rate. Income is broken down into other labor, proprietors except farm, property, transfer, social insurance, wage and salary, and agricultural. Each equation is related to the appropriate national and local activities. Prices in the Atlanta SMSA are also determined.

Another large sector of the Georgia model is the state government sector. Equations are presented for each of the major revenue services of state government, intergovernmental revenues, miscellaneous charges, along with an equation for general expenditures. Most of the revenue equations are estimated as a function of some local activity variables while a few also include a tax rate variable. The revenue areas involve: general sales, motor fuels, alcoholic beverage, tobacco, insurance and other selective taxes, motor vehicle licenses, other license fees, personal income, corporation income, and other taxes.

Other sectors included in the model such as manufacturing investment, demographic, retail sales, banking, and miscellaneous seem to play a supportive role to the rest of the model. Manufacturing investment is composed of structures and equipment and both are primarily related to manufacturing output. Total population is estimated in the demographic section while migration is derived by identity. Retail sales are estimated for automobile and non-automobile sales and these variables enter back into some of the output formulations. Banking information is provided for demand and time deposits as well as commercial loans and investments.

The model was estimated using annual data from 1951-1968. Ordinary least squares was used as the estimating technique. At the time of this

publication, the model had not been tested as to its forecasting capabilities. This has been done in later efforts and with quite reasonable success (40).

Mississippi

The regional econometric model for Mississippi was formulated by F. Gerard Adams, Carl G. Brooking, and Norman J. Glickman (1). This model is very similar in its basic construction to the Georgia model as well as other contemporary modeling efforts. As the authors noted, and the growing trend in regional modeling indicates:

In place of the final demand identity in the typical Keynesian model, the basic account identity for regional model building is the so-called 'third' or output approach. Gross State Product (GSP) is the sum of gross output by sectors. This significant modification in the model structure reflects lack of certain critical data (p. 286).

The model is centered around this key output block with blocks also existing for employment, wage rates and personal income, and taxes. The Mississippi model consists of 39 equations of which 29 are behavioral relations.

The output section consists of equations for the basic industries of the economy. In addition, the manufacturing sector is divided into the categories of durables and non-durables. The industries are split into export-oriented and internally-oriented and their respective outputs are determined by the demand for goods produced by them. Export industries are related to national variables while local conditions determine the internal industries. Relative unit cost variables also play a significant role.

Labor demand relationships determine the employment block. Industry employment is a function of industry output and real wages. Time

variables were also used to denote technological change. The unemployment rate equation is related to the national unemployment rate.

Wage rates in the Mississippi model are very aggregative and do not involve much detail. Wage rates exist for manufacturing, agriculture, and non-manufacturing non-agriculture. Manufacturing wages are determined by their U. S. counterpart and, in turn, affect the other two categories. Total wages and salaries along with other labor income, property income, proprietors non-farm income, farm income, transfers, and social insurance contributions determine personal income.

The tax sector is determined as a function of a tax rate and tax base. Proxies for the tax base are used since no such precise tax base information exists. Equations for sales tax, retail sales, federal income tax, income tax effective rate, and income tax on a fiscal basis compose this section. Not much detail is allowed for in this section.

The Mississippi model was estimated from annual data over the sample period 1953-1970. The model was log-linearly specified using ordinary least squares (OLS) and iterated instrumental variables (IIV). The authors noted that, "An analysis of the results indicates that there is very little gain, measured by reduced root mean square error, to be found using IIV" (1, p. 291). The more sophisticated simultaneous equation estimation technique did not seem to add anything to the effort.

Simulation error tests were made and the model was found to "... successfully track the growth path of the major economic aggregates in Mississippi over the sample period" (1, p. 291). Multiplier studies were made and a control forecast for the Mississippi economy from 1973-1980 was presented. Rapid expansion was predicted for Mississippi during this time.

Tennessee

The Tennessee model was constructed by Hui S. Chang (11). This log-linear model follows the lines of construction that were used by the Georgia and Mississippi models. GSP is formed by the gross product originating approach for industries. The model is fairly large and it contains 77 equations and 120 variables. The major sectors of the economy are output, employment, wage and nonwage income, tax revenue, and retail sales.

The same basic industry breakdown is involved in the output, employment, and wage and salary equations. In addition, government is divided into a federal component and a state and local component. As in the Mississippi model, manufacturing is decomposed into durable and nondurable manufacturing for these sectors except for wages and salaries. A composite manufacturing wage and salary variable is determined at that point.

Output equations as in the Mississippi model are categorized as to whether they are export-oriented or internally-oriented and then related to the appropriate activity variable. "Such a distinction of course cannot be unequivocal" (11, p. 8). Some sectors are affected by both national and local activities. Employment is basically specified as a function of output and real wages in the industry. The wage sector is centered around activity in manufacturing. Manufacturing wages are determined by U. S. manufacturing wages and then affect all other industries' wages. Local labor conditions are also allowed for. These specifications in output, employment, and wages are all similar to those in some of the previously mentioned models.

The personal income sector of the Tennessee model also follows a common theme established by earlier models. Total wages and salaries are determined by identity and join with nonfarm proprietors' income, farm proprietors' income, property income, other labor income, transfer payments, contributions to social insurance, and a resident adjustment to form total personal income. The income sector also includes a number of other important income concepts that are computed through identities. These concepts include per capita personal income, disposable personal income, real disposable personal income, per capita disposable income, real per capita disposable income, and disposable income in fiscal years.

The equations for the tax sector primarily involve state tax revenues, however, a federal income tax equation is also estimated. Revenues are generally related to tax bases and tax rates. Proxies for these features are sometimes substituted. Tax behavioral equations involve state taxes of the following nature: sales and use, gross receipts, cigarette, gasoline, alcoholic beverages, corporation excise, motor vehicles, and other tax revenue.

The retail sales sector is very small and consists of only automobile and nonautomobile retail sales. Disposable personal income plays a major role in both equations. Sales tax rates and interest rates also determine automobile sales.

The Tennessee model was constructed from annual data over the period 1952-1972. Some of the employment equations involved shorter periods from 1964-1972. Ordinary least squares was used as the regression technique. This model was rigorously tested for its replication ability and it proved highly successful. MAPE errors for the Tennessee model seemed to be slightly lower than those associated with the Mississippi,

Georgia, and Philadelphia IV models. Multiplier simulations were made and, "The results showed that the model correctly reflects the effect of changes in national and state exogenous variables on the Tennessee economy" (11, p. 65). Baseline forecasts as well as those for two alternative possibilities for the 1976-1982 time period were determined and their effects on the Tennessee economy were compared. This model underwent an extensive list of tests and seemed to perform quite well.

Philadelphia SMSA (Philadelphia IV)

The development of the regional econometric model of the Philadelphia SMSA has been an ongoing project of Norman J. Glickman (24) (25) (26) (27).⁵ The initial model contains 26 equations and ". . . was essentially a three-sector model involving 1) manufacturing, 2) wholesale and retail trade and selected services, and 3) all other activity" (24, p. 76). These sectors included equations for output, wages, and employment and along with the model equations for income, labor force, demographics, prices, and government conditions composed the entire model. The model was expanded along industrial lines and was enhanced by the deepening of the breadth of all sectors. The current Philadelphia IV model contains 228 equations of which 123 are identities and 105 are behavioral relationships.

The Philadelphia SMSA model actually involves three models. There is the overall SMSA model along with separate submodels for the City of Philadelphia and the suburbs. This makes this model very spatial in nature. The entire structure is composed of 14 separate blocks or

⁵ The Standard Metropolitan Statistical Area (SMSA) of Philadelphia includes the counties of Bucks, Chester, Delaware, Philadelphia, and Montgomery in the state of Pennsylvania along with the counties of Burlington, Camden, and Gloucester in New Jersey.

divisions. These blocks contain manufacturing output; non-manufacturing output; manufacturing employment; non-manufacturing employment; wages, prices and income; federal and local government; manufacturing investment, demographics; retail sales; banking; Philadelphia City; suburbs; consumption; and quarterly equations.

The manufacturing output and employment blocks are composed of the 12 major manufacturing industries in the SMSA. They include food and kindred products; textile mill products; apparel and related products; printing and publishing; chemicals and allied products; petroleum and coal products; primary metal industries; fabricated metal products; machinery, except electrical; electrical machinery; transportation equipment; and other manufacturing. As in many other models, the export-oriented industries have their output related to national activity variables while the locally-oriented industries are related to local activities. In some cases both are involved. The manufacturing employment block related industry employment to industry output and lagged employment.

The non-manufacturing output and employment blocks contain the other major industries of the economy. For the most part, the output variables of this category are related to local variables and the employment variables involve the same inverse production function specification used in the manufacturing employment equations.

The wages, prices, and income block involves a slightly different approach than that used by some of the other models. Instead of average annual wages being estimated for each industry and then summed into a total wage and salary figure, average money earnings are computed for only the breakdown of manufacturing and non-manufacturing. The

other typical components of personal income are present in this model. Equations exist for other labor income, proprietors' income, property income, transfer payments, and contributions to social insurance. These equations involve the usual specifications. An equation for the consumer price index in the SMSA is estimated in the model and it is related primarily to national consumer price movements and unit labor costs.

The federal and local government block is one of the most expansive sectors in this regional model. The large number of equations is partially due to the fact that most of the variables of this block are converted by identity from nominal to real terms. This section does present some relationships which are not often estimated in a regional model. Along with the appearance of federal income taxes there are behavioral equations for municipal government revenues, school district expenditures, school district enrollment, and the market value of property. The tax revenues for the municipalities and the school districts depend upon the market value of property while the expenditure variables of these items are constrained by their revenues collected. Many of the variables of this sector interact in a manner similar to those above.

The blocks for manufacturing investment, demographics, retail sales, and banking all tend to be very small in size. The manufacturing investment equation is determined by gross regional output, manufacturing output, and capital stock considerations. Demographics include behavioral equations for population and the unemployment rate. The labor force and the number unemployed are determined by identities. These variables depend upon local conditions as well as the national

unemployment situation for the unemployment rate equation. Retail sales are divided into automobiles, food, drugs, gasoline, general merchandise, and other sales. These sales equations are all related to disposable personal income. The banking block deals with time deposits and demand deposits. These deposits are determined by local activity and interest rate considerations. Total loans and investments are, in turn, driven by the amount of total deposits.

The submodel for Philadelphia City is constructed in a fashion similar to the overall regional model. Equations exist for personal income, employment, investment, output, municipal government, and school district effects. A large government sector is produced for the city. The suburban submodel has its variables determined as the residual from the overall region and the City of Philadelphia.

Cross-sectional information from 1960 and 1961 provide the basis for consumption expenditures equations. The consumption equations are all a fixed percentage of personal income.

The last block of the model involves some quarterly equations for some of the significant variables of the region. Most of the relationships are in a percentage change format. These equations cover specifications for employment, unemployment, and consumer prices. Most of the functional forms are similar to the annual specifications.

The model was estimated basically from data over the period 1947-1971. The Philadelphia IV version was specified using OLS while many of the earlier efforts also incorporated simultaneous equation techniques. The model was thoroughly tested and found to replicate economic activity very successfully. MAPE errors were found to be lower than some of those in other published models. Multiplier tests were made and several

policy impact simulation forecasts were made. The Philadelphia SMSA model appeared to be a very valuable forecasting tool.

Basic Characteristics

It has been noted previously that there are many avenues of approach when constructing a regional econometric model. Different possibilities were discussed as to linkages with national models, specification formats, and the time reference of the data to be used. Although there are these potential ways model constructions can vary, many of the previously estimated regional models do possess some of the same basic features or characteristics. Norman J. Glickman (24) has derived the following list of 11 major characteristics of regional econometric models:

- a) Many of the important problems in the development of regional econometric models have revolved around the availability of data. One of the data constraints has been the lack of data on a basis more frequent than annual. As a result, most models are estimated on annual data . . .
- b) Because of the use of annual data, there are relatively few observations . . . most have approximately 15-17 observations.
- c) The unfortunate fact that there are very few variables for which there are lengthy time series constitutes another main data constraint . . . the models have been relatively small. Although, the number of equations in these models ranges from 14 to 228, most are 35 or fewer.
- d) The combination of annual data and few variables with long time series has not only produced small models but ones which are relatively simple--often consisting of sets of bivariate relationships . . . This results in part from the fact that there are relatively few statistical degrees of freedom; thus explanatory variables that ought to be included in equations must be omitted. When this is so, such equations are subject to errors in specification.

- e) The models are relatively static. With so few observations, there is little room for accurately specifying the lag relationships that may be relevant, many of which hold for periods of less than one year.
- f) In addition, the models are highly recursive.
- g) Consistent with Klein's suggestion, the models are heavily linked with the national economy . . . The presence of large numbers of national exogenous variables and the highly recursive nature of many regional models means that they are structurally highly dependent on national movements, and they do not constitute, to any considerable degree, internally generated systems.
- h) As in other kinds of empirical research, the availability of data often influences the direction of research . . . there are a large number of variables of great interest to regional analysts and public policy-makers that are missing, including exports, imports, migration, and various land use variables.
- i) . . . there is relatively little spatial disaggregation and, thus, little analysis of intraregional phenomena.
- j) Again due to data inavailability very few models have been estimated for small areas. Most have been constructed for states or even larger areas.
- k) There are also significant problems relating to the use of constructed data: most researchers use the Kendrick-Jaycox method which tends to mask differences in regional production functions as well as wage versus non-wage industrial income. Thus interesting differences in employment productivity and wages are lost in this data construction process (pp. 61-64).

With these important differences and similarities of regional econometric models noted from this chapter, the task of the next chapter will be to specify a framework to use in constructing a model for Oklahoma. In deciding upon an appropriate specification for Oklahoma, realistic theory and data accessibilities will have to be considered.

CHAPTER III

EQUATION SPECIFICATIONS

Introduction

In the last chapter, various alternatives to the construction of a regional econometric model were discussed. Issues over the appropriate equation specifications, data usages, and national-regional linkages were presented with the intent of using this information to derive a framework for constructing a regional econometric model for Oklahoma. Before proceeding to this point, it might be worthwhile to first consider whether Oklahoma is a valid economic region and, therefore, of significance in modeling. This introduces the more general question of what exactly constitutes a region.

Region Definition

A noted regional economist, Harry W. Richardson (52) has commented on the importance of the question of what determines a true region.

Consideration of what constitutes a region and of how the national economy may be sub-divided into a system of regions would appear to be an essential prerequisite for the analysis of regional economic phenomena (p. 223).

Richardson points out that even though there are a number of different approaches to the task of defining regions, a few basic methods are most prominent. "Virtually all these fall within three main categories: uniform or homogeneous regions; nodal regions; and programming or planning regions" (p. 224).

The three approaches to defining regions do not involve concepts which are mutually exclusive. Some congruencies and similarities may be found among the different approaches.

The homogenous or uniform approach to defining regions involves forming regions from units which have similar characteristics. Production, consumption, labor, or attitudinal similarities may constitute the construction of a region. The nodal approach to deriving regions emphasizes the importance of the interplay of different units. The functional interdependence of components as observed by flows of activity is the distinguishing feature. The third major category for definition of a region, the planning or programming technique, views decision-making as the important determinant. Richardson (52, p. 229) has remarked that in this category, ". . . regions are defined in terms of the coherence and unity of economic decision-making." Political jurisdictions become crucial in this partitioning. It is possible that these three different approaches may design regional boundaries that are very similar in nature.

The above delineation may be used to attempt to answer the question of whether the State of Oklahoma constitutes a true economic region. In light of either the homogenous or nodal approaches to defining regions, much debate could arise when this question is considered. Interesting discourse could occur on whether production or consumption patterns are similar in the metropolitan areas of the state as compared to the Panhandle portion of the state and whether Amarillo, Wichita, or Dallas-Fort Worth have any influence over certain parts of the state. Although it might appear that Oklahoma could never be considered a valid region when using these approaches, the matter can be seemingly resolved by

considering the planning or programming approach to demarcating regions. Oklahoma is a well-defined political jurisdiction and, therefore, shows a certain amount of solidarity in its economic decisions. By relying upon this latter consideration, Oklahoma can be described (at least in a planning sense) as a valid economic region. The development of a regional econometric model for Oklahoma will then proceed with the knowledge that Oklahoma is a true region by at least one definition.

Model Decisions

In designing a structure for a regional econometric model, certain decisions must be made about the alternative issues that were discussed in the second chapter. Key to these decisions may be the broader question of what is exactly wanted from the model. The desired output or achievable ends from constructing the model may play a large role in determining what the model is actually composed of.

The intent behind the construction of the Oklahoma model is to develop a structure which shows the interaction of the key economic sectors in the state. The key variables which are of importance to this plan involve the levels of output, personal income, wages and salaries, governmental tax revenue, and employment occurring within the state. It is desired to develop a model which will show the interdependence of these sectors and, therefore, be a tool for predicting changes in the sectors when some sort of shock is introduced into the economic system. Forecasting expected future activity within the state is also highly desirable.

The wish to include an output sector in the Oklahoma model and have estimates of output levels available for analysis removes some of

the choices pertaining to the formation of the model. As mentioned before, output data only appears on an annual basis. Quarterly output data do not exist. If an output sector is to be specified in a model, the entire model will probably have to be constructed from annual data. Hence, the Oklahoma model by the inclusion of an output sector will be estimated solely from annual data.

There is no real decision to be made concerning the issue of the appropriate national-regional linkage for the Oklahoma model. The "top-down" approach is the only sensible method since the construction of national models predates that of any regional model development. Usage of the "bottom-up" method would require the construction of regional models for all regions before national totals could be computed and appropriately linked to the regions. This approach would involve much greater costs to employ. Therefore, the Oklahoma model will be linked to one of the already existing national models.

The specification form of the model will follow the recent trend away from the national modeling approach and be based around the gross regional product output identity. The triple accounting identity for expenditures, income, and output that was advocated by Klein will not be used. This choice reflects the data constraints that are placed on most regional modelers. Many of the expenditure components as well as corporate profit measures for Oklahoma are simply not available. The Oklahoma model will ignore expenditure activity and instead concentrate on output, employment, and income conditions.

Suggested Model Structure

Klein and Glickman (36) in a recent article have described what

they consider to be a typical satellite regional model. The main core of their proposed model contains equations for production by sector, employment by sector, local wage rates, personal income, public receipts, and expenditures. It is noted that the expenditure accounts cannot be fully developed but, ". . . should be filled out to the extent possible" (36, p. 8).

With one exception, the structure that was proposed by Klein and Glickman will be utilized in the Oklahoma model formulation. The Oklahoma model will be centered around six blocks of equations. These blocks will be for output, employment, personal income, wage and salary, tax revenue, and miscellaneous activity. The suggested expenditure block of Klein and Glickman will be omitted due to the inability to obtain consistent data series for any of the major components. The available retail sales data for Oklahoma is very inadequate but it is superior to that for the other private expenditures. Hence, the Oklahoma model will be restricted to a form that is very similar to that employed by some of the previously mentioned regional models (1) (11) (31).

Model Block Theory

In this section, a detailed discussion will be presented on the development of each of the six equation blocks which makeup the Oklahoma econometric model. These six blocks as mentioned earlier are for output, employment, wage and salary, personal income, tax revenue, and miscellaneous economic activity. The discussion for each block will center on the equations used to explain activity, the hypothesized specification form of those equations, and the types and sources of data used for the endogenous variables of the block. As expected, the specification of

each block will call upon the use of economic theory whenever possible.

After the development of the blocks, a flow chart will be developed showing the flows of activity hypothesized for the economy. The specification form will be tempered in the final product by the determined statistical significance of the hypothesized relationships.

Output

The real output sector of the econometric model will feature equations for the following industries: manufacturing; mining; contract construction; wholesale and retail trade; finance, insurance, and real estate; transportation, communication, and public utilities; services; state and local government; and federal government. The variables in this sector will be in the form of gross product originating by industry in 1972 dollars. The sum of gross product originating by industry for the above industries as well as an exogenous agriculture sector will determine the principal identity of the model--gross state product (GSP).

Gross product originating for each industry except manufacturing will be constructed using the Kendrick-Jaycox method of output determination (34). This technique involves applying national ratios to state income-received data. Manufacturing is handled in a slightly different manner since value-added data is already published for this industry. Data components for constructing these output measures can be found in the Survey of Current Business (63) and the Annual Survey of Manufacturers (62).¹ Gross product originating for each industry

¹A complete listing of data sources can be found in Appendix A.

is initially computed in nominal terms and then deflated using the appropriate industry deflator. Real GSP is then found by summing the individual industry real output components.

For the output sector, demand-type relationships between the industry and its relevant output market will be estimated. Some similarities to economic base theory arise in this sector as some decision has to be made concerning which industries are export-oriented and which are domestically-oriented. An arbitrary choice has to be made in classifying some industries as to whether they service markets outside the state or within the state. This is because some industries contain components of both and the industry breakdown provided within the model is fairly aggregative. An industry has to be classified here generally as either all export or all local in its production. A more industrially disaggregative model would have more flexibility in this area.

It is believed generally that manufacturing and mining are export-oriented industries. To some degree, federal government output for the state may also be classified in this division. Hence, these industries will be modeled with this assumption in mind while all other industries in this sector will be considered to be internally determined.

The manufacturing, mining, and federal government industries will be determined primarily by their U. S. counterparts. Along with this national influence, these industries will be modeled to also depend upon some measure of local activity. The manufacturing sector, in particular, will also be affected by the competitiveness of production in the state. A prominent cost differential affecting competition is the labor cost. Hence, a relative labor cost variable will be used in the manufacturing equation estimation. The mining industry output will

also be associated with local mineral production or oil well completions. A variable of this nature will be included in the mining equation. For these two industries as well as all of the other industries in the block, lagged output variables will be included to allow for adjustment towards desired output. This variable reflects the Koyck lag effect.

The remaining industries in the real output sector--contract construction; wholesale and retail trade; finance, insurance, and real estate; transportation, communication, and public utilities; services; and state and local government will be modeled to respond primarily to internal demands. The principal measures of local demand are gross state product, personal income, and disposable personal income. In addition to these major activity variables, other influences such as population in the finance, insurance and real estate equation and Moody's Aaa rate in the construction industry will be used where applicable. These variables represent conditions which are key to specific industries' activities.

Employment

The employment block of the model is actually composed of two areas of related activity. First of all, there are equations specifying employment activity for eight of the nine industrial sectors of the economy. These sectors are manufacturing; mining; contract construction; wholesale and retail trade; transportation, communication, and public utilities; finance, insurance, and real estate; services; and government. Government is also broken down into a federal and state and local component. The second part of the employment block concerns the civilian labor force and unemployment conditions.

The employment equations for the eight industries are of the labor demand variety and are determined from a CES production function assuming profit maximization. The profit function for a producer is determined by the difference between his total revenue and his total costs.² Equation 3.1 demonstrates such an equation:

$$\pi = P \cdot Q - w \cdot L - r \cdot K \quad (3.1)$$

where π = profits,

P = product price,

Q = output,

w = wage rate,

L = labor,

r = capital rental rate, and

K = capital stock.

A CES production function takes the following form:

$$Q = \delta [\alpha K^{-\rho} + (1 - \alpha) L^{-\rho}]^{-\beta/\rho} \quad (3.2)$$

where δ , α , ρ , and β are all parameters.

The first order condition for profit maximization involves taking the partial derivative of output with respect to labor for equation (3.1). This nets the following equation:

$$\frac{\partial \pi}{\partial L} = P \cdot \frac{\partial Q}{\partial L} - w = 0 \quad (3.3)$$

which can be transformed into the following:

²The components for the following theoretical derivation can be found in many microeconomics textbooks such as that of James M. Henderson and Richard E. Quandt (32).

$$\frac{\partial Q}{\partial L} = \frac{w}{P} \quad (3.4)$$

For profit maximization, the marginal product of labor $\left(\frac{\partial Q}{\partial L}\right)$ should be equated to the real wage rate (w/P). The marginal product of labor can be computed from the production function in equation (3.2). This derivation occurs in the following fashion:

$$\begin{aligned} \frac{\partial Q}{\partial L} &= \delta \left(-\frac{\beta}{\rho}\right) [\alpha K^{-\rho} + (1 - \alpha)L^{-\rho}]^{-\frac{\beta}{\rho} - 1} [(1 - \alpha)(-\rho)L^{-\rho - 1}] \\ &= \delta \beta (1 - \alpha) L^{-\rho - 1} [\alpha K^{-\rho} + (1 - \alpha)L^{-\rho}]^{-\frac{\beta}{\rho} - 1} \\ &= \beta (1 - \alpha) L^{-(1 + \rho)} Q [\alpha K^{-\rho} + (1 - \alpha)L^{-\rho}]^{-1} \\ &= \delta^{-\frac{\rho}{\beta}} \beta (1 - \alpha) L^{-(1 + \rho)} Q \left(1 + \frac{\rho}{\beta}\right) \\ \frac{\partial Q}{\partial L} &= \gamma L^{-(1 + \rho)} Q \left(1 + \frac{\rho}{\beta}\right) \end{aligned} \quad (3.5)$$

where $\gamma = \delta^{-\rho/\beta} \beta (1 - \alpha)$.

By substituting equation (3.5) into equation (3.4), one finds:

$$\gamma L^{-(1 + \rho)} Q^{1 + \rho/\beta} = w/P \quad (3.6)$$

By solving for L and taking the logarithmic form, the basic employment specification form can be established with sector employment determined as a function of sector output and the sector's real wage rate. In functional notation this can be written as:

$$E = f(Q, w/P) \quad (3.7)$$

where E represents employment. As can be shown from the derivation, one would expect to deduce a positive output effect and an inverse real wage relationship with employment.

In addition to the derived specification of above, the employment equations will also be estimated by allowing for a time variable and lagged employment. The time variable is used to determine whether significant capital-labor substitutions have occurred over time. The lagged employment variable is established as a Koyck lag where, ". . . a Koyck lag is used to reflect the adjustment process of actual employment toward optimal employment, given the level of output" (11, p. 30).

The employment data to be used in this sector is taken from the Oklahoma Employment Statistics (48). The sum of the employment figures for the eight industries will be called non-agricultural employment and this variable will interact in other sectors of the model to explain activity.

The second half of the employment block will feature the estimation of equations for the civilian labor force and the unemployment rate. With these two areas estimated, the actual number of unemployed could be computed by multiplying the unemployment rate times the civilian labor force. The level of unemployment was chosen to be computed in this residual fashion because of the problems in acquiring a consistent data series for this variable.

The estimation of the unemployment rate for Oklahoma will involve a form slightly different than that employed in other regional modeling attempts. Instead of a single unemployment rate for the state, two rates will be estimated. The regular state unemployment rate will be modeled along with an insured workers' unemployment rate. This form is being employed because of the fact that some sectors of the state economy respond more strongly to changes in the insured unemployment rate. In addition to their nominal forms, the unemployment rates will

be converted by identity to a relative configuration. This form will be relative to the U. S. unemployment rate.

The unemployment rate equations will be related to national levels of unemployment along with local activity variables such as GSP, personal income, or wages and salaries. This specification will allow for the effect of national and local conditions. A lagged unemployment rate variable will also be included in order to test the "discouraged worker hypothesis" (65). This hypothesis suggests that high unemployment rates in the past may cause some workers to leave the labor market. The civilian labor force will be specified in a simple form also.

Wage and Salary

The wage and salary block for the Oklahoma econometric model will contain behavioral equations for the following nine industries: manufacturing; mining; contract construction; wholesale and retail trade; finance, insurance and real estate; transportation, communication, and public utilities; services; government--state and local as well as federal civilian; and farm. The variables used in this category are average annual wages and salaries and they are formed initially by dividing a sectoral wage bill by the sector's total employment. The data concerning wages comes from that published in the Survey of Current Business (63).

The specification of the wage and salary block centers around activity in the manufacturing industry. As is assumed in some of the other regional models (1) (11) and suggested by Wilbur R. Thompson (61), regional wages tend to move in line with national wage movements via the export sector. The manufacturing sector, as was discussed in the

output block, is believed to be the primary export sector of the state economy. Hence, the manufacturing sector is viewed as the key center of activity for determining wage and salary levels for all industries.

The manufacturing industry is initially modeled and then all other industries are estimated based upon the manufacturing sector's wages and salaries. This specification assumes at the outset that the:

Manufacturing industry is part of the national labor market, and consequently its wage rate is determined in relation to national manufacturing wages. Local labor market conditions are represented through the state unemployment rate (1, p. 290).

Hence, the manufacturing industry is formulated to be dependent upon the national situation as represented by the U. S. manufacturing wage rate and the local atmosphere through the local unemployment rate. The filtering of the national conditions down to the regional wage determination through the export sector incorporates what Thompson described as "intra-area wage rollout" (61, p. 71).

The concept of "intra-area wage rollout" concerns the fact that the manufacturing sector must compete with the other industrial sectors for the local labor supply. And in like fashion, the other industries of the economy must compete with manufacturing for its' employees. As Thompson (61) stated:

If the local export industries, those selling outside the local market, pay high wages, one would expect the contagion of a high wage rate to run throughout the whole labor market, if we assume some significant amount of labor substitution between industries and occupations and some significant resistance to migration (p. 71).

The assumptions of labor substitution between industries and occupations as well as some resistance to migration are not so restrictive as to invalidate this formulation.

Therefore, all of the other industries of the state economy except farming are estimated primarily as a function of the manufacturing sector's wages and salaries. Local employment conditions are entered into the formulation by use of unemployment rates. The usual inverse relationship between unemployment rates and wage rates would be expected. This would reflect the assumption that tight labor markets (low unemployment rates) correspond to rising wage rates and just the opposite for loose labor markets (high unemployment rates). Other labor activity variables may be included where their use seems appropriate. The farm sector is modeled in a slightly different manner. It is, instead, related to the hired labor expense in farming for the state. This is thought to be a more relevant formulation for this industry.

Another feature of the wage and salary block is the fact that its inclusion along with the employment block establishes a complete labor market. Labor demand conditions were established in the employment block as ". . . a derived demand obtained by setting marginal revenue product equal to the wage" (19, p. 266). Its counterpart, the labor supply, ". . . is incorporated through adjustments in the wage rate which responds to unemployment" (19, p. 266). The development of the wage and salary block with its dependence upon the unemployment situation allows the introduction of labor supply aspects into the model. This aspect is also important because the labor supply must ". . . be estimated to derive the wages used in the labor demand equation and to generate the wage bill for the region" (51, p. 60). Hence, an implicit formulation helps to establish a complete labor market.

Personal Income

The personal income block will be composed of eight basic components: total wages and salaries; non-farm proprietors income; other labor income; contributions to social insurance; dividend, rent, and interest income; transfer payments; farm proprietors income; and a resident adjustment. Behavioral equations will be determined for all of them except total wages and salaries and farm proprietors income. Total wages and salaries will be determined by identity while farm proprietors income is exogenous to this model. The sum of these eight components will determine total personal income which interlinks into the model at other places as an explanatory variable. The data for these dependent variables are taken from the Survey of Current Business (63).

The total wages and salaries component of personal income will draw upon information from the wage and salary block and the employment block for its construction. It will be computed by summing for all industries the product of their average annual wages and their employment. Along with these products, exogenous variables like federal military wages and salaries, farm wages and salaries, and other industries wages and salaries will be added in to arrive at the final identity. Total wage and salary income also interacts with some of the other components of personal income in their behavioral determinations.

For the most part, the other components of personal income are determined by a local income variable or by a corresponding U. S. variable. Other labor income, which ". . . consists of supplementary types of labor income paid out or accruing in the current period" (56, p. 61), is linked to total wage and salary income along with the U. S. ratio of other labor income to wage and salary income. The total wage and salary

variable would seem to play a major role since contributions to health plans, insurance plans, and compensation for injuries (other labor income components) are highly dependent upon the amount of wages earned. The ratio variable would attempt to measure the national influence in this area.

Transfer payments and dividend, rent, and interest income are specified to be primarily dependent upon their national counterparts. In addition, local conditions such as employment or population may enter into their behavioral relations. Nonfarm proprietors income is hypothesized to be dependent upon some total income or employment variable portraying the local activity conditions of the time. In a similar fashion, contributions to social insurance are linked to the OASI tax rate and a local activity variable such as wage and salary income. Resident adjustment is simply associated with its past behavior.

In addition to the computation of total personal income, disposable personal income is also included in the model. This variable is derived through an identity with federal income taxes and state income taxes being the subtractions from personal income. Disposable personal income plays an important role in determining other equations of the model.

Tax Revenue

Certain data limitations will restrict the tax revenue block to a small number of equations. Behavioral equations will be estimated for sales and use taxes, tobacco products taxes, alcoholic beverage taxes, motor fuels taxes, motor vehicle excise taxes, total state income taxes (corporation and personal), and federal income taxes. The above taxes except for federal income taxes will be summed along with two exogenous

tax measures to arrive at a total state tax measure. The two exogenous taxes in this block are gross production taxes and all other taxes.

Data for these variables comes from a special Oklahoma tax report (55).

"Theoretically, tax collections for each category can be explained as the product of the appropriate rate and base" (1, p. 290). If possible, each behavioral equation in this block will be estimated as a function of the appropriate tax rate and tax base. As easy as this may seem, there are problems with this specification. Precise tax base data do not exist for the tax components and proxies must be used in their place. Also, at times multiple tax rates may exist for any category (1, p. 290). This makes the selection of any one rate as the appropriate tax rate somewhat difficult. At other times, no tax rate change has occurred over the sample period and, therefore, cannot be used in a statistical analysis.

Tax bases will be proxied by three local activity variables. These variables are personal income, disposable personal income, and total population. The rate variables, where they can appear, will be represented by rate indexes which are characteristic to each category.

Effects of price movements also have a great influence upon tax collections. Therefore, price indices for appropriate tax categories as well as the implicit price deflator for consumption will be used to measure the price and inflationary influences, respectively. Substitution effects from relative price movements will be denoted by the price indices for particular tax categories and inflationary movements will be tracked through usage of the implicit price deflator for consumption. Negative substitution effects and positive inflationary movements would be expected.

Miscellaneous

A final block of equations to be included in the model concerns financial activity within the state. The equations of this block will essentially stand alone and will not interact with any other blocks of the model. Relationships will be determined for demand deposits, time deposits, and total loan activity within the state. Data for this sector are supplied by a publication of the U. S. Federal Deposit Insurance Corporation (64).

The purpose of this block is to mainly provide forecasts in this financial area. Hence the equations will be specified with forecasting in mind and not the impact of specific policies. The general functional form of these equations will be to relate the financial variables to a general income variable and an interest rate. The deposit variables will be attached to total personal income which proxies general activity within the state. In addition, they will be related to some sort of competing interest rate such as Moody's Aaa bond rate. The loan equation will also be formulated as dependent upon general activity and some sort of interest rate. Personal income will again proxy general conditions while the prime commercial paper rate and the Federal Reserve discount rate may provide adequate interest rates.

Proposed Oklahoma Model

The Oklahoma econometric model as proposed in the previous discussion will be constructed to contain 65 equations. Of these proposed equations, 47 of them will establish behavioral relationships while 18 will contain identities. Figure 1 presents the actual model of the

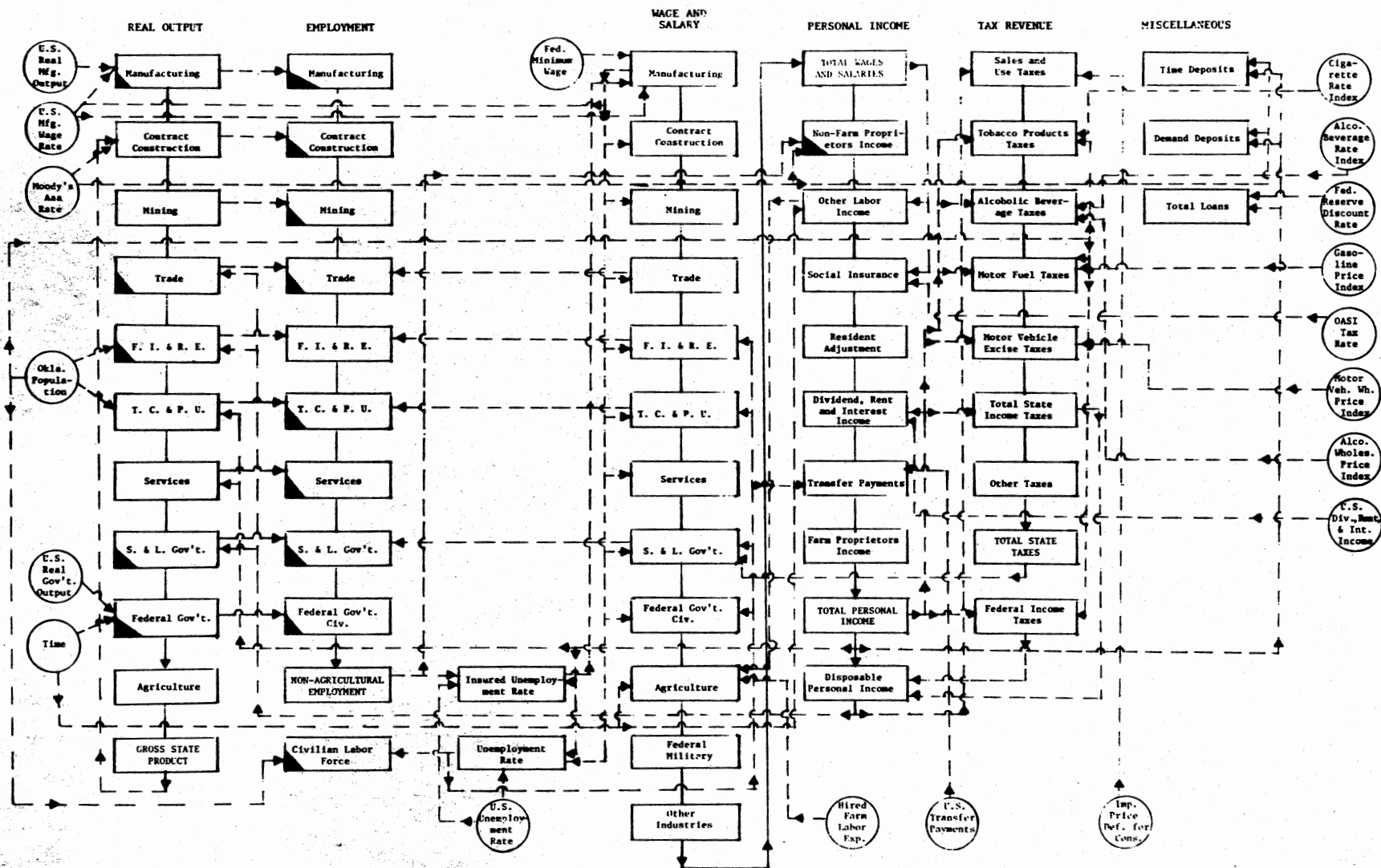


Figure 1. Flow Chart of the Oklahoma Economy.

economy determined in the next chapter. The proposed model and the actual model deviate slightly due to the results of the statistical testing performed in the following chapter.

Columns are established in the flow chart to represent the basic equation blocks in the model. The rectangular boxes in each column signify (with a few exceptions) the presence of an endogenous variable.³ Circular drawings represent exogenous variables. Identities are denoted by solid lines in the chart and behavioral relationships are signified by dashed lines. In addition, the inclusion of a lagged endogenous variable is explained by the shading of the lower left-hand side of a rectangular box.

From Figure 1, it can be seen that the output sector will contain eight behavioral equations. Employment will entail 12 behavioral equations. The wage and salary block will be composed of 10 behavioral equations. Five behavioral equations make up the personal income block. The tax revenue block will contain seven behavioral equations while the miscellaneous block will be composed of three behavioral equations. In addition, numerous identities serve to solidify the model. Appropriate economic flows can be observed for each block of equations by use of the explanations issued above.

The proposed Oklahoma model as outlined in this chapter will be empirically tested in the next chapter. Statistical techniques will be used to derive the final Oklahoma econometric model. The model will be tested in great detail as to its forecasting capabilities and these results will also be noted in the fourth chapter.

³Rectangular boxes in which no flow arrows are shown entering into the equation denote an exogenous variable.

CHAPTER IV

MODEL ESTIMATION AND TESTING

Introduction

In the third chapter, a theoretical specification for the econometric model of Oklahoma was presented. It is the objective of this chapter to empirically estimate and test the hypothesized specification using acceptable regression techniques. After the model has been initially estimated, it will need to be tested as to its ability to replicate economic activity during the sample period as well as out of the sample period. A further test of the model will be made by observing the response of the system of equations to a multiplier analysis. The results of this manner of model estimation and model testing will be presented in this chapter.

The Oklahoma econometric model will be estimated from annual data over the 1958-1975 time period. This time span is dictated by the availability of data built upon a common benchmark and the lags present in reporting historical values of certain variables.

A question arises, at the outset, as to what the appropriate regression technique should be for estimating a simultaneous equation model. The large-sample properties of simultaneous equation estimators such as two-stage least squares (TSLS), full-information maximum likelihood (FIML), and limited-information maximum likelihood (LIML) methods are superior to those of the single equation estimators such as ordinary

least squares (OLS). OLS yields inconsistent estimates in simultaneous equation situations (33). However, small sample properties are of greater interest to this model construction effort since only 18 observations are available for estimation. Several studies have been made into small sample properties of these estimators and Monte Carlo results "... indicates that OLS, while often more biased than the other procedures, exhibits the property of minimum Mean Squared Error. Thus, this method should not be dismissed for small sample models . . ." (26, p. 24).¹ Since the one major intent of forecasting with an econometric model is to minimize error, the OLS estimation procedure should definitely not be ruled out. For these reasons the Oklahoma econometric model will be estimated using a single equation estimator (OLS) and a simultaneous equation estimator.² Other regional modeling efforts have employed the usage of both a single equation and a simultaneous equation estimator in their constructions (1) (12) (26) (41). Some of these studies have found no real gain as measured by reduced prediction error from using the simultaneous equation estimator over the single equation estimator (1) (12). The Oklahoma model will be estimated using both techniques and the two estimators will be judged as to their ability to replicate economic activity.

Once estimation has been completed, the testing of the model's replication capabilities is made by simulating economic activity over

¹For a list of Monte Carlo studies and small sample property investigations, see reference (26).

²The simultaneous equation estimator will involve an instrumental regression process (INST) and will be discussed later.

the sample period.³ As mentioned earlier, this simulation is attainable by finding the reduced form of the model when it is linear or by using the Gauss-Seidel iterative technique if the model is nonlinear. This can be visualized for the linear case by recalling that the entire model can be denoted in matrix notation via equation (1.2) as:

$$Y\beta + X\gamma = E$$

where all matrices are as defined before. If the matrix of coefficients for the endogenous variables (β) is nonsingular, then there exists a matrix β^{-1} such that:

$$Y\beta^{-1} + X\gamma\beta^{-1} = E\beta^{-1}.$$

This equation can be transformed into the following reduced form notation:

$$Y = -X\gamma\beta^{-1} + E\beta^{-1} \quad (4.1)$$

$$Y = X\pi + U$$

where $\pi = -\gamma\beta^{-1}$ and $U = E\beta^{-1}$.

This reduced form states that all of the endogenous variables can be written as a function of all of the exogenous variables. If the model to be estimated is nonlinear, then the reduced form cannot be achieved. It can be approximated by an iterative technique such as the Gauss-Seidel method.

³Simulation can be described as the mathematical solution of a simultaneous set of difference equations (49).

The actual simulation process is achieved by initializing starting values for the endogenous variables and providing a time series for the exogenous variables in the model. A solution for the first year is determined via the reduced form or by iterative processes provided that the model is stable and prespecified convergence criterion are met. Estimated endogenous variables are used to update lagged endogenous variables if they exist and a new set of exogenous variables are called upon to generate a solution for the next period. This process can be repeated for as many periods as there are exogenous variables provided to the simulation program.⁴

By usage of this simulation process, a set of predicted endogenous variables can be generated and compared with actual historical values so as to determine the ability of the model to replicate activity. A series of simulation error statistics can be computed to test the model. Since no standardized tests exist to evaluate these error statistics, comparison with other models is the only way to determine the worth of one's estimates.

In summary, the testing of the entire model will be achieved by 1) calculating the reduced form of the model if linear or approximating it by the Gauss-Seidel iterative technique if nonlinear, 2) using actual values of the exogenous values and generated lagged endogenous variables during the sample period to calculate a set of predicted endogenous variables, and 3) to compare the predicted with the actual endogenous variables during and after the sample period. These predicted and

⁴A discussion of numerical solution methods and computer algorithms can be found in a work edited by T. H. Naylor (45).

actual values will be contrasted by computing the following error statistics: mean absolute percentage error (MAPE), mean absolute error (MAE), root mean square error (RMSE), and Theil's "U" coefficient.

In addition, the model will be tested through a multiplier and impact elasticity analysis by shocking the entire system of equations with a change in one or more of the exogenous variables. This perturbed solution can be compared with a baseline solution to determine dynamic multipliers or impact elasticities. The results can be used to see if the model performs as expected with regard to certain exogenous changes.

This chapter will proceed, first, with the presentation of the OLS model and its testing. Secondly, this will be followed by the simultaneous equation model and its tests of replication abilities. Finally, a multiplier analysis will be presented.

OLS Model

The OLS model contains 63 equations. Of this number, 45 of the equations are behavioral relations while 18 of the equations are identities. Involved within the model are 63 endogenous variables and 45 predetermined variables. The 45 predetermined variables can be further partitioned into 29 regular exogenous variables and 16 lagged endogenous variables. The model is nonlinear.

The equations of this section were estimated by use of ordinary least squares (OLS). Where serial correlation posed a problem, the equation(s) were corrected by use of the Cochrane-Orcutt technique (13).

Serial correlation was a problem for 16 of the equations.⁵

The equations of this model will be presented in a block by block fashion. A brief discussion of the estimated equation will be presented along with information on other variables that were tried but found unacceptable. Statistical information involving the following will be listed: t statistics for each variable (under each coefficient), coefficients of determination (R^2), adjusted coefficients of determination (\bar{R}^2), the standard error of the model ($S_{y.x}$), Durbin-Watson statistics (D.W.), first-order autocorrelation values where a serial correlation correction has been made (ρ), and a statistic suggested to detect serial correlation when lagged endogenous variables are present in an equation (h).⁶

Output

The variables of this block are expressed in 1972 dollar figures. The real output sector is composed of nine equations. Eight of the equations are behavioral relations and one equation is an identity. This identity is the gross state product (GSP) computation which summarizes output activity in the economy. The GSP identity is composed of activity in 10 industries. Eight of the industries have their output

⁵Serial correlation is normally tested by use of the Durbin-Watson statistic. This statistic is biased, though, when lagged endogenous variables are present in an equation. For equations with lagged endogenous variables, a substitute h statistic suggested by Durbin (18) was used.

⁶Equations containing lagged endogenous variables which are found to involve serial correlation will be corrected in a similar fashion to that of the other equations. All corrected equations will be presented with the Durbin-Watson statistic. Equations with lagged dependent variables that are corrected will not include an h statistic.

endogenously determined within the model while two industries, farming and mining, are exogenous. The mining output was made exogenous because of the inability to successfully model this activity. The equations of this sector now follow.

Manufacturing Output (QMFG).

$$\text{QMFG} = 2851.79 + 2.563307 \text{ RGDPMA} - 3425.314 \text{ Z} + .800466 \text{ QMFG1} \quad (0.1)$$

(2.207) (3.422) (-2.455) (11.272)

$$R^2 = .9852 \quad \bar{R}^2 = .9818 \quad S_{y.x} = 82.1893 \quad \text{D.W.} = 2.6117 \quad h = -1.3997$$

The manufacturing output equation is related positively to real manufacturing output activity within the U. S. (RGDPMA). This reflects the assumption that manufacturing output is nationally oriented. Output activity is also shown to be negatively related to a relative labor cost variable (Z). This shows that as manufacturing wages in Oklahoma fall relative to the U. S. wages for manufacturing, manufacturing output increases. The positive relationship with lagged manufacturing output (QMFG1) denotes a movement towards desired output. Another variable that was tried but excluded was the lagged effect of the labor cost variable.

Contract Construction Output (QCC).

$$\text{QCC} = 246.7236 + .0365817 \text{ GSP} - 16.21966 \text{ MAAA} \quad (0.2)$$

(6.218) (4.010) (-1.512)

$$R^2 = .7872 \quad \bar{R}^2 = .7588 \quad S_{y.x} = 26.5793 \quad \text{D.W.} = 1.9203$$

The contract construction output equation is determined primarily by activity within the state as proxied by gross state product (GSP). A cost variable or credit indicator as represented by Moody's Aaa bond

rate (MAAA) denotes the negative effect this condition plays. As MAAA increases and credit gets tighter, construction output is curtailed. Other variables that were tried for this equation were lagged construction output, real contract construction awards, and real personal income.

Wholesale and Retail Trade Output (QWRT).

$$\text{QWRT} = 22.02929 + .1379727 \text{ DPYAD} + .393101 \text{ QWRT1} \quad (0.3)$$

(.296) (2.969) (1.655)

$$R^2 = .9769 \quad \bar{R}^2 = .9736 \quad S_{y.x} = 58.7666 \quad \text{D.W.} = 1.7949 \quad h = -1.436$$

The wholesale and retail trade output equation is dependent upon local activity as represented by the real disposable income variable (DPYAD). This positive influence demonstrates that as DPYAD increases so does QWRT. The positive sign on the lagged dependent trade variable (QWRT1) shows the movement towards desired output. Other variables that were tried included gross state product, real personal income, and state population.

Finance, Insurance, and Real Estate Output (QFIRE).

$$\text{QFIRE} = -785.3264 + .087029 \text{ DPYAD} + .464736 \text{ POP} + .238244 \text{ QFIRE1}$$

(-1.298) (3.898) (1.482) (1.289) (0.4)

$$R^2 = .9898 \quad \bar{R}^2 = .9875 \quad S_{y.x} = 29.5981 \quad \text{D.W.} = 2.0633 \quad h = -.523$$

κ

The finance, insurance, and real estate output equation is also locally oriented. The positive influence of real disposable income (DPYAD) and state population (POP) represent local activities which affect the output of this industry. Also, the lagged dependent variable

(QFIRE1) demonstrates the move towards desired output. Gross state product and real personal income are variables that were tried but not included in the final estimation equation.

Transportation, Communication, and Public Utilities Output (QTCPU).

$$QTCPU = - 418.2575 + .126065 PYAD + .132032 POP \quad (0.5)$$

$$(-1.698) \quad (14.475) \quad (1.047)$$

$$R^2 = .9957 \quad \bar{R}^2 = .9951 \quad S_{y.x} = 17.9777 \quad D.W. = 2.2483$$

The transportation, communication, and public utilities output equation is determined by two local activity variables. Real personal income (PYAD) has a strong, positive influence upon the output of this industry. In addition, population (POP) has a weaker, positive impact upon QTCPU. Other variables that were tried include real disposable income, gross state product, and a lagged endogenous variable.

Service Output (QSER).

$$QSER = 126.5232 + .110563 PYAD \quad (0.6)$$

$$(4.779) \quad (34.774)$$

$$R^2 = .9853 \quad \bar{R}^2 = .9844 \quad S_{y.x} = 26.5084 \quad D.W. = 1.9499$$

The service output equation is determined solely by real personal income (PYAD). This local activity variable has a strong, positive influence on the service sector. Variables that were tried but not acceptable were real disposable income, gross state product, population, and a lagged endogenous variable.

State and Local Government Output (QSLG).

$$\text{QSLG} = 43.18252 + .014564 \text{ DPYAD} + .837139 \text{ QSLG1} \quad (0.7)$$

$$(1.833) \quad (1.792) \quad (8.104)$$

$$R^2 = .9931 \quad \bar{R}^2 = .9921 \quad S_{y.x} = 11.9025 \quad \text{D.W.} = 1.3513 \quad h = 1.223$$

The state and local government output equation is found to be dependent upon the positive influence of local activity and the lagged dependent variable (QSLG1). Real disposable income (DPYAD) denotes the activity variable in this case. Population, state taxes, and real personal income were also tried but were found to provide not as satisfactory a fit for the equation.

Federal Government Output (QFG).

$$\text{QFG} = -1225.893 + 21.32422 \text{ RGDPG} + .101787 \text{ QFG1} - 74.95338 \text{ TIME} \quad (0.8)$$

$$(-9.021) \quad (10.169) \quad (1.040) \quad (-10.439)$$

$$R^2 = .9733 \quad \bar{R}^2 = .9671 \quad S_{y.x} = 19.9748 \quad \text{D.W.} = 1.9071 \quad h = -.018$$

The federal government output sector is determined primarily by U. S. real government output (RGDPG). National conditions have a strong, positive influence on this sector. The lagged endogenous variable (QFG1) also indicates the movement towards optimal output. The strong negative effect of the time variable (TIME) demonstrates the fact that in Oklahoma the federal government has been playing a diminishing role in state activity. Other variables that were tried but not included for various reasons were real personal income and real disposable income.

Gross State Product (GSP).

$$\begin{aligned} \text{GSP} = & \text{QMFG} + \text{QCC} + \text{QMIN} + \text{QWRT} + \text{QFIRE} + \text{QTCPU} + \text{QSER} + \\ & \text{QSLG} + \text{QFG} + \text{QFARM} \end{aligned} \quad (0.9)$$

The key aggregate of the output sector is the gross state product (GSP) identity which is found by summing output activity from the above estimated industries along with the two exogenous sectors; mining (QMIN) and farming (QFARM). This aggregate summarizes production activity within the state and is presented in real terms to differentiate price or inflation effects from real activity.

Employment

The employment sector of the model is made up of 15 equations. Behavioral relations involve 12 equations while the other three equations are composed of identities. The same basic industrial breakdown is provided for in this block of equations as was used in the real output block with the exception that farm employment is not included. The summation of employment over the industries of this block is denoted as non-agricultural employment (NAE). One other minor difference from the output block is that federal employment is composed solely of civilian labor and does not include the military sector. A total civilian labor force equation along with an unemployment rate and an insured unemployment rate equation are presented in this block. The equations now follow.

Manufacturing Employment (MFGE).

$$\text{MFGE} = 32.035269 + .0208308 \text{ QMFG} + .447105 \text{ MFGE1} \quad (0.10)$$

$$(2.959) \quad (2.712) \quad (2.316)$$

$$R^2 = .9405 \quad \bar{R}^2 = .9314 \quad S_{y.x} = 4.2017 \quad \text{D.W.} = 1.1828 \quad \rho = .3524$$

The manufacturing employment sector is found to be positively related to manufacturing output (QMFG) within the state. In addition, the positive effect on the lagged dependent variable (MFGE1) indicates the movement towards desired employment. This lagged variable derives from the Koyck lag effect. The equation was found to be suffering from serial correlation in its original form and was corrected using the Cochrane-Orcutt technique (13). Other variables that were tried but not included were real manufacturing wages and a time variable.

Contract Construction Employment (CCE).

$$\text{CCE} = .516177 + .0455712 \text{ QCC} + .347654 \text{ CCE1} \quad (0.11)$$

$$(.095) \quad (5.270) \quad (2.680)$$

$$R^2 = .7707 \quad \bar{R}^2 = .7354 \quad S_{y.x} = 1.3106 \quad \text{D.W.} = 1.2915 \quad \rho = .5917$$

The contract construction employment equation is modeled as dependent upon contract construction output (QCC) in a positive fashion. Increases in construction output bring about increases in construction employment. The positive effect of the lagged dependent (CCE1) again reflects the movement towards desired employment. This equation was corrected using the Cochrane-Orcutt technique (13) for serial correlation. Real contract construction wages and a time variable were tried but they were found to be unacceptable.

Mining Employment (MINE).

$$MINE = 4.746045 + .00174419 QMIN + .841191 MINE1 \quad (0.12)$$

$$(.916) \quad (.419) \quad (11.162)$$

$$R^2 = .8881 \quad \bar{R}^2 = .8721 \quad S_{y.x} = 1.1644 \quad D.W. = 1.3751 \quad h = 1.083$$

The mining employment equation is determined by output in the mining sector (QMIN) as well as the lagged endogenous variable effect (MINE1). Again, positive relationships exist between the output and employment variables in this sector. Other variables that were tried but not included were real mining wages and a time variable.

Wholesale and Retail Trade Employment (WRTE).

$$WRTE = 24.6237 + .0254352 QWRT - 4.20548 WRTWSD + .741497 WRTE1 \quad (0.13)$$

$$(2.135) \quad (4.128) \quad (-1.543) \quad (8.331)$$

$$R^2 = .9968 \quad \bar{R}^2 = .9961 \quad S_{y.x} = 1.5185 \quad D.W. = 1.8711 \quad h = -.0598$$

The wholesale and retail trade employment sector is determined by not only the sector's output (QWRT) and the lagged endogenous variable (WRTE1) but also the sector's real wages (WRTWSD). The trade output has a positive influence upon employment while the sector's real wages have a negative impact upon employment. If real wages rise in the trade sector, then it becomes more costly to operate and employment in the trade sector is reduced. The lagged endogenous variable indicates the movement towards desired employment. A time variable was also tried in the estimation process but it was found to be unsatisfactory.

Finance, Insurance, and Real Estate Employment (FIREE).

$$\text{FIREE} = 7.401434 + .0264784 \text{ QFIRE} - 1.248905 \text{ FIREWSD} \quad (0.14)$$

$$(2.105) \quad (18.370) \quad (-1.554)$$

$$R^2 = .9893 \quad \bar{R}^2 = .9879 \quad S_{y.x} = .7657 \quad \text{D.W.} = 2.0183$$

The finance, insurance, and real estate employment equation is found to be determined by the sector's output (QFIRE) and its real wages (FIREWSD). Output has a strong, positive influence on employment while real wages have a negative or inverse relationship with employment. Other variables that were tried were a lagged endogenous variable and a time variable.

Transportation, Communication, and Public Utilities Employment (TCPUE).

$$\text{TCPUE} = 36.123834 + .0184767 \text{ QTCPU} - 2.078265 \text{ TCPUWSD} + .285549 \text{ TCPUE1} \quad (0.15)$$

$$(1.593)$$

$$R^2 = .9197 \quad \bar{R}^2 = .8997 \quad S_{y.x} = .7877 \quad \text{D.W.} = 1.1330 \quad \rho = .3523$$

The transportation, communication, and public utilities employment equation is determined by the sector's output (QTCPU), the sector's real wages (TCPUWSD), and a lagged endogenous variable (TCPUE1). Output has a positive effect upon employment while real wages have a negative effect. The lagged endogenous variable indicates the movement towards desired employment. A serial correlation problem was corrected by the Cochrane-Orcutt technique (13). A time variable was also tried but it was excluded from the final equation.

Service Employment (SE).

$$SE = - 1.995903 + .0195087 QSER + .856232 SE1 \quad (0.16)$$

$$(-1.240) \quad (2.604) \quad (12.366)$$

$$R^2 = .9980 \quad \bar{R}^2 = .9977 \quad S_{y.x} = 1.1066 \quad D.W. = 1.8367 \quad h = .344$$

The service employment sector is determined by service output (QSER) and a lagged endogenous variable (SE1). The positive influence of the output of this sector demonstrates that as output increases so will employment in the service sector. The lagged variable shows the movement towards desired employment. Other variables that were tried include a time variable and real service wages.

State and Local Government Employment (SLGE).

$$SLGE = - 3.246237 + .090344 QSLG - 2.583695 SLGWSL + .538705 SLGE1 \quad (0.17)$$

$$(-.531) \quad (5.020) \quad (-1.422) \quad (4.504)$$

$$R^2 = .9972 \quad \bar{R}^2 = .9966 \quad S_{y.x} = 1.3213 \quad D.W. = 1.8099 \quad h = .4218$$

The state and local government employment sector is dependent upon the sector's output (QSLG), the sector's real wages (SLGWSL), and a lagged dependent variable (SLGE1). Output in this government sector has a positive impact upon employment while real wages have a negative effect. Again, the lagged endogenous variable denotes the movement towards desired employment. A time variable was also tried but it was found to be statistically insignificant.

Federal Government Civilian Employment (FGCE).

$$FGCE = 2.075992 + .0316244 QFG + .327548 FGCE1 \quad (0.18)$$

$$(.820) \quad (7.077) \quad (3.731)$$

$$R^2 = .9589 \quad \bar{R}^2 = .9530 \quad S_{y.x} = 1.1370 \quad D.W. = 1.5404 \quad h = .958$$

The federal government civilian employment equation is found to be dependent upon federal government output (QFG) and a lagged endogenous variable (FGCE1). Output portrays a strong, positive influence upon employment while the lagged variable indicates the movement towards desired employment. Real federal government civilian wages and a time variable were also tried but they were found to be unsatisfactory.

Non-Agricultural Employment (NAE).

$$NAE = MFGE + CCE + MINE + WRTE + FIREE + TCPUE + SE + SLGE + FGCE \quad (0.19)$$

The aggregate employment variable found in this block is non-agricultural employment (NAE). This variable is formed by summing the above estimated employment equations. It is composed of nine industries' employment figures.

Percentage Change in Non-Agricultural Employment (PCNAE).

$$PCNAE = \frac{NAE - NAE1}{NAE1} \quad (0.20)$$

This percentage change in non-agricultural employment identity (PCNAE) is found by subtracting from the current non-agricultural employment

(NAE) total the total from last period (NAE1) and by dividing by the last period (NAE1) total. This variable is later used as an explanatory variable.

Civilian Labor Force (CLFA).

$$\begin{aligned} \text{CLFA} = & -157.6885 - 4.83948 \text{ OKAUNR1} + .160336 \text{ POP} + .787984 \text{ CLFA1} \\ & (-1.224) \quad (-.802) \quad (1.397) \quad (4.281) \quad (0.21) \\ R^2 = & .9645 \quad \bar{R}^2 = .9563 \quad S_{y.x} = 17.0405 \quad \text{D.W.} = 2.1624 \quad h = -.8005 \end{aligned}$$

The civilian labor force equation was found to be dependent upon the lagged unemployment rate (OKAUNR1), state population (POP), and a lagged endogenous variable (CLFA1). The negative impact of the lagged unemployment rate demonstrates the "discouraged worker" hypothesis mentioned in the previous chapter. If the unemployment rate was high in the previous period, then some people tend to become discouraged about work and drop out of the labor force. The observed significance level of this variable does indicate, though, that this effect is not very strong. The positive impact of state population indicates that as population rises, the civilian labor force will also increase. The lagged endogenous variable indicates an adjustment process. Other variables that were tried include a lagged insured workers unemployment rate as well as changes in the unemployment rates.

Unemployment Rate (OKAUNR).

$$\begin{aligned} \text{OKAUNR} = & -4.450471 + .846005 \text{ USUR} - .000639532 \text{ PYA} + \\ & (-3.099) \quad (12.944) \quad (-2.828) \\ & 1.312868 \text{ MFGWS} \\ & (2.980) \quad (0.22) \\ R^2 = & .9194 \quad \bar{R}^2 = .9022 \quad S_{y.x} = .3228 \quad \text{D.W.} = 1.5149 \end{aligned}$$

The unemployment rate equation for the state indicates that this measure is dependent upon national activity as well as local conditions. National effects seem to have a strong, positive influence as is indicated by the positive sign on the U. S. unemployment rate (USUR). As the U. S. rate rises, Oklahoma unemployment also increases. Local conditions are reflected by the state personal income variable (PYA) and the manufacturing sector's nominal wages (MEGWS). The negative effect of personal income seems to indicate that as income rises, the unemployment rate falls. This could be due to the notion that some members of the family drop out of the labor market (actively seeking a job) when other members are earning more income. The positive sign on manufacturing's wages seems to point out that as wages are pushed up in this key industry, employment is cut back. Some people are then pushed out of a job and the unemployment rate rises. Other variables that were tried include a lagged endogenous variable and a percentage change in non-agricultural employment variable.

Relative Unemployment Rate (RELAUNR).

$$\text{RELAUNR} = \frac{\text{OKAUNR}}{\text{USUR}} \quad (0.23)$$

The relative unemployment rate relates the ratio of the Oklahoma unemployment rate to that of the U. S. This variable is used as an explanatory variable in other equations.

Insured Unemployment Rate (OKIUNR).

$$\text{OKIUNR} = 2.323406 + .647716 \text{ USUR} - .000245771 \text{ PYA} - 18.77973 \text{ PCNAE} \quad (0.24)$$

$$(6.042) \quad (10.645) \quad (-13.827) \quad (-3.523)$$

$$R^2 = .9703 \quad \bar{R}^2 = .9634 \quad S_{y.x} = .2132 \quad \text{D.W.} = 1.7969$$

The insured worker's unemployment rate is also controlled by national as well as local conditions. The U. S. unemployment rate (USUR) again has a positive impact upon this unemployment rate measure while personal income (PYA) and the percentage change in non-agricultural employment (PCNAE) have a negative effect. The personal income impact is negative because of the notion that people who are possibly seeking a job, but don't have one, leave the labor market when others in their family provide greater earnings. The negative sign on the percentage change in non-agricultural employment variable indicates that as PCNAE rises (which implies employment is expanding), the unemployment rate decreases. As employment is expanded, more of the unemployed workers are able to find jobs. Other variables that were tried include a lagged endogenous variable, the lagged U. S. unemployment rate, and manufacturing wages.

Wage and Salary

The wage and salary block of the Oklahoma econometric model is made up of 15 equations. Ten of the equations are behavioral while five are identities. The same basic industrial breakdown as before is provided for in this block. All of the industries except farming will be presented in the form of average annual wages. The farm sector will be estimated in terms of total disbursements for wages and salaries. The equations are now presented.

Manufacturing Wages and Salaries (MFGWS).

$$\begin{aligned} \text{MFGWS} = & .2664407 + .710968 \text{ USMFGW} - .0180541 \text{ OKIUNR} + \\ & (2.597) \quad (31.466) \quad (-1.335) \\ & .777179 \text{ FEDMW} \\ & (6.239) \end{aligned} \quad (0.25)$$

$$R^2 = .9974 \quad \bar{R}^2 = .9968 \quad S_{y.x} = .0485 \quad \text{D.W.} = 1.3617 \quad \rho = .5316$$

The manufacturing wages and salaries equation is modeled to allow primarily for the effects of national manufacturing but also the effects of other national and local conditions. Oklahoma's wages and salaries for manufacturing are determined for the most part by U. S. manufacturing's wages and salaries (USMFGW). The positive influence of this national market indicates that as national wages increase so will state manufacturing wages. This arises from the basic assumption of this block that manufacturing is a nationally-oriented industry. The federal minimum wage (FEDMW) also indicates a positive impact upon wages and salaries for this industry. Local conditions are also allowed for as is demonstrated by the negative effect the insured worker's unemployment rate (OKIUNR) has on wages and salaries. As the unemployment rate rises, upward movements of manufacturing wages are curtailed due to the slackness of the labor market. Other variables that were tried but excluded involved other forms of the unemployment rate variable. A serial correlation problem was corrected using the Cochrane-Orcutt technique (13).

Contract Construction Wages and Salaries (CCWS).

$$\begin{aligned} \text{CCWS} = & - 2.0272912 + 1.317006 \text{ MFGWS} \\ & (-4.868) \quad (22.468) \end{aligned} \quad (0.26)$$

$$R^2 = .9674 \quad \bar{R}^2 = .9652 \quad S_{y.x} = .2396 \quad \text{D.W.} = 1.6008 \quad \rho = .4529$$

The contract construction wages and salaries are directly related to the movements of the manufacturing sector's wages and salaries (MFGWS). The positive coefficient which is greater than one in magnitude indicates that a one unit positive change in manufacturing's wages will cause a greater than one unit change in construction's wages and salaries. Other variables that were tried include the two unemployment rates, in both nominal and relative forms, as well as real contract construction awards. A serial correlation problem was corrected using the Cochrane-Orcutt procedure (13).

Mining Wages and Salaries (MINWS).

$$\text{MINWS} = - 3.4897035 + 1.769487 \text{ MFGWS} \quad (0.27)$$

$$(-6.198) \quad (24.142)$$

$$R^2 = .9717 \quad \bar{R}^2 = .9698 \quad S_{y.x} = .1899 \quad \text{D.W.} = 1.6261 \quad \rho = .7081$$

Mining wages and salaries are related to activity from the manufacturing sector's wages and salaries (MFGWS). The positive coefficient on this variable demonstrates that a greater than unity increase in mining wages will result from a unit increase in manufacturing's wages and salaries. Other variables that were tried include the unemployment rates in their nominal and relative forms. The Cochrane-Orcutt technique (13) was used to correct a serial correlation problem.

Wholesale and Retail Trade Wages and Salaries (WRTWS).

$$\text{WRTWS} = - .4582698 + .769352 \text{ MFGWS} \quad (0.28)$$

$$(-3.098) \quad (37.246)$$

$$R^2 = .9879 \quad \bar{R}^2 = .9871 \quad S_{y.x} = .0792 \quad \text{D.W.} = 1.4310 \quad \rho = .4958$$

The trade sector's wages and salaries are positively related to the manufacturing wages and salaries (MFGWS). The coefficient on this variable indicates that a unit change in manufacturing wages will bring about a less than unit change in the trade sector's wages. The federal minimum wage along with the various forms of the unemployment rates were also tried in the estimation procedure. A serial correlation problem was corrected using the Cochrane-Orcutt technique (13).

Finance, Insurance, and Real Estate Wages and Salaries (FIREWS).

$$\text{FIREWS} = 1.175736 + .909334 \text{ MFGWS} - 2.05131 \text{ RELAUNR} \quad (0.29)$$

$$(3.676) \quad (53.799) \quad (-4.719)$$

$$R^2 = .9939 \quad R^2 = .9931 \quad S_{y.x} = .1208 \quad D.W. = 1.4860$$

The finance, insurance, and real estate wages and salaries are determined by the manufacturing sector's wages (MFGWS) as well as the relative unemployment rate (RELAUNR). The positive but less than unity coefficient on the manufacturing wages variable demonstrates that this sector's wages will increase by less than a unit when manufacturing is increased by a one unit change. The negative unemployment rate variable says that as the unemployment rate rises in Oklahoma as compared to the U. S. rate, the finance, insurance, and real estate wages will decline. Again, slack labor market conditions are responsible. Some of the other unemployment rate variables were also tried in the estimation process.

Transportation, Communication, and Public Utilities Wages and Salaries (TCPUWS).

$$\text{TCPUWS} = - 1.020875 + 1.511178 \text{ MFGWS} - 1.808021 \text{ RELAUNR} \quad (0.30)$$

$$(-2.290) \quad (37.828) \quad (-3.383)$$

$$R^2 = .9884 \quad \bar{R}^2 = .9868 \quad S_{y.x} = .1416 \quad \text{D.W.} = 1.2921 \quad \rho = .5536$$

The transportation, communication, and public utilities wages and salaries are also dependent upon the manufacturing sector's wages and salaries (MFGWS) as well as the relative unemployment rate (RELAUNR). The manufacturing wage coefficient shows that a one unit increase in manufacturing wages will cause a greater than one unit increase in TCPUWS. An increase in the relative unemployment rate will cause upward wage movements in this sector to be curtailed. Other forms of the unemployment rate were also tried for this equation. The Cochrane-Orcutt procedure (13) was used to correct for serial correlation.

Service Wages and Salaries (SERWS).

$$\text{SERWS} = - .2187438 + .738589 \text{ MFGWS} \quad (0.31)$$

$$(-.728) \quad (19.152)$$

$$R^2 = .9557 \quad \bar{R}^2 = .9528 \quad S_{y.x} = .0945 \quad \text{D.W.} = 1.7956 \quad \rho = .7323$$

The service wages and salaries are related positively to wage movements in manufacturing (MFGWS). A one unit increase in manufacturing wages will induce a less than one unit increase in service wages. Other variables that were tried include various unemployment rate forms and the federal minimum wage. Serial correlation was corrected using the Cochrane-Orcutt technique (13).

State and Local Government Wages and Salaries (SLGWS).

$$\begin{aligned} \text{SLGWS} = & 1.244348 + .603212 \text{ MFGWS} + .00176515 \text{ ST} - \\ & (2.755) \quad (4.685) \quad (1.466) \\ & 1.728066 \text{ RELAUNR} \\ & (-4.552) \end{aligned} \quad (0.32)$$

$$R^2 = .9906 \quad \bar{R}^2 = .9884 \quad S_{y.x} = .1020 \quad \text{D.W.} = 1.8047 \quad \rho = .2742$$

The state and local government wages and salaries are determined by manufacturing wages (MFGWS), total state taxes (ST), and the relative unemployment rate (RELAUNR). The manufacturing wage influence indicates that a one unit increase in manufacturing wages will have a less than one unit positive impact on government wages. State taxes tend to have a positive influence upon wages and salaries. Wages tend to rise as total state tax collections increase. The relative unemployment rate causes a curtailment of wage increases when this relative unemployment rate rises in Oklahoma. A federal minimum wage variable was also tried for this equation. The Cochrane-Orcutt method (13) was used to correct for serial correlation.

Federal Government Civilian Wages and Salaries (FGCWS).

$$\begin{aligned} \text{FGCWS} = & -1.3912329 + 1.699109 \text{ MFGWS} - 2.596229 \text{ RELAUNR} \\ & (-1.544) \quad (27.588) \quad (-2.236) \end{aligned} \quad (0.33)$$

$$R^2 = .9784 \quad \bar{R}^2 = .9753 \quad S_{y.x} = .3008 \quad \text{D.W.} = 1.3554 \quad \rho = .3329$$

The federal government civilian wages and salaries equation was found to be dependent upon the manufacturing sector's wages and salaries (MFGWS) and the relative unemployment rate (RELAUNR). The manufacturing

coefficient indicates that a one unit increase in manufacturing wages will lead to a greater than one unit change in federal government wages. The relative unemployment rate variable points out the negative impact rising unemployment rates have on wages. Other variables that were tried in the estimation process involve various unemployment rate forms and a variable measuring the civil service (grade seven) wage level. A serial correlation problem was corrected using the Cochrane-Orcutt procedure (13).

Farm Wages and Salaries (FWS).

$$\begin{aligned} \text{FWS} = & - 1.225348 + 1.001086 \text{ HLE} + 28.21433 \text{ OKOLWS} - \\ & (-.714) \quad (21.386) \quad (.354) \\ & .442545 \text{ TIME} \quad (0.34) \\ & (-3.549) \end{aligned}$$

$$R^2 = .9968 \quad \bar{R}^2 = .9961 \quad S_{y.x} = .7582 \quad \text{D.W.} = 2.9031$$

The total farm wage and salary disbursements are found to be dependent upon Oklahoma agricultural hired labor expense (HLE), the ratio of other labor income to total wages and salaries (OKOLWS), and a time variable (TIME). Farm wages are primarily dependent in a positive fashion upon the hired labor expense.⁷ A unit increase in the hired labor expense provides a minutely higher increase in wages and salaries. The ratio of other labor income to total wages and salaries has a very small effect upon farm wages. Its weak, positive sign indicates that increases in the ratio may induce rises in wages and salaries. This

⁷Hired labor expense is a variable that is forecasted in an agricultural submodel for Oklahoma.

could be due to some notion of compensation for what is being granted in terms of extra benefits in other industries. The time variable indicates the declining nature of the emphasis in the farm sector.

Relative Labor Cost (Z).

$$Z = \frac{MFGWS}{USMFGW} \quad (0.35)$$

This relative labor cost identity is composed of the ratio of manufacturing wages in Oklahoma to that of the U. S.

Real Wholesale and Retail Trade Wages and Salaries (WRTWSD).

$$WRTWSD = \frac{WRTWS}{IPDC} \quad (0.36)$$

Real trade wages are formed by deflating nominal wages by the implicit price deflator for consumption (IPDC).

Real Finance, Insurance, and Real Estate Wages and Salaries (FIREWSD).

$$FIREWSD = \frac{FIREWS}{IPDC} \quad (0.37)$$

Real wages for this sector are formed by deflating nominal wages by the implicit price deflator for consumption (IPDC).

Real Transportation, Communication, and Public Utilities Wages and Salaries (TCPUWSD).

$$TCPUWSD = \frac{TCPUWS}{IPDC} \quad (0.38)$$

Real wages for this sector are formed by deflating nominal wages of the sector by the implicit price deflator for consumption (IPDC).

Real State and Local Government Wages and Salaries (SLGWSD).

$$SLGWSD = \frac{SLGWS}{IPDC} \quad (0.39)$$

State and local government real wages are formed by deflating nominal wages by the implicit price deflator for consumption (IPDC).

Personal Income

The personal income block of the model is composed of 13 equations. Identities contribute eight of the equations while behavioral relations are estimated for the other five equations. The major aggregate of this block is the total personal income identity. In addition, a disposable personal income variable is also computed. These two aggregates are converted into a fiscal year basis in order to be used in the tax revenue block later. The resident adjustment equation which was to be a behavioral relation was made exogenous due to the inability to simulate satisfactorily with the equation. The equations are now presented.

Other Labor Income (OLY).

$$OLY = - 189.06 + .137982 WSY - 15.99744 TIME \quad (0.40)$$

$$(-18.864) \quad (19.677) \quad (-5.670)$$

$$R^2 = .9941 \quad \bar{R}^2 = .9932 \quad S_{y.x} = 9.0831 \quad D.W. = 1.5228 \quad \rho = .4590$$

The other labor income equation is determined by total wage and salary disbursements (WSY) and a time variable (TIME). The positive

influence of the wage and salary income variable indicates that as disbursements rise so will payments for extra benefits. The negative impact of the time variable indicates that by itself time would have a curtailing effect upon other labor income. Other variables that were tried include a lagged endogenous variable and the U. S. ratio of other labor income to total wages and salaries. The Cochrane-Orcutt technique (13) was used to correct for serial correlation.

Non-Farm Proprietors Income (NPY).

$$\begin{aligned} \text{NPY} = & - 512.4071 + 1.492254 \text{ NAE} + .678863 \text{ NPY1} - \\ & (-4.695) \quad (4.820) \quad (4.950) \\ & 54.87685 \text{ MAAA} \quad (0.41) \\ & (-3.016) \\ R^2 = .9689 \quad \bar{R}^2 = .9617 \quad S_{y.x} = 37.4358 \quad \text{D.W.} = 1.3170 \quad h = 1.2276 \end{aligned}$$

The non-farm proprietors income equation is determined by local activity conditions as well as national financial indicators. Local activity is represented by non-agricultural employment (NAE) which has a strong, positive effect upon this income source. National conditions are represented by Moody's Aaa bond rate (MAAA). This variable has a negative impact upon non-farm proprietors income and indicates that when this rate rises, income of this type will fall. This could be due to a general tightening effect in the economy or to losses of sources of revenue for expansion or firm creation. A lagged endogenous variable (NPY1) also discloses a positive impact. Other variables that were tried include U. S. business and professional income, U. S. non-farm proprietors income, and a summation of trade and service employment.

Dividend, Rent, and Interest Income (DIRY).

$$\text{DIRY} = - 85.206669 + 11.00309 \text{ UDIRY} + .0236705 \text{ PYA} \quad (0.42)$$

$$(-3.926) \quad (5.722) \quad (1.056)$$

$$R^2 = .9961 \quad \bar{R}^2 = .9955 \quad S_{y.x} = 22.0916 \quad \text{D.W.} = 1.5625 \quad \rho = .3643$$

Dividend, rent, and interest income is dependent upon both national and local activities. The primary determinant is the positive influence of U. S. dividend, rent, and interest income (UDIRY). As the U. S. variable rises, the state total also increases. Also, local conditions are allowed for in that personal income (PYA) has a positive effect upon dividend, rent, and interest income. Upward movements in total personal income cause increases in this income source. Greater income usually leads to greater investments. Other variables that were tried but excluded from the final equation include a lagged endogenous variable, oil well completions within the state, and the separate effects of U. S. dividends and U. S. rent and interest income. Serial correlation was corrected for by using the Cochrane-Orcutt method (13).

Transfer Payments (TRY).

$$\text{TRY} = - 61.43823 + 11.94558 \text{ USTRY} + 171.5061 \text{ RELAUNR} \quad (0.43)$$

$$(-1.454) \quad (145.457) \quad (2.985)$$

$$R^2 = .9992 \quad \bar{R}^2 = .9991 \quad S_{y.x} = 16.1219 \quad \text{D.W.} = 1.3734$$

The transfer payments equation also reflects the effects of national and local conditions. The equation is primarily determined by U. S. transfer payments (USTRY). An increase in the U. S. figure will bring about a large increase in state payments. Local circumstances are

represented by the relative unemployment rate (OKAUNR). The positive sign on this variable indicates that as unemployment rises in Oklahoma as compared to the nation, local transfer payments are increased. State population as well as various forms of the unemployment rate were also tried in the estimation process.

Contributions to Social Insurance (CSSY).

$$\begin{aligned} \text{CSSY} = & -132.6792 + .0771405 \text{ WSY} + 6.076921 \text{ OASI} & (0.44) \\ & (-9.924) \quad (17.010) \quad (.834) \\ R^2 = & .9943 \quad \bar{R}^2 = .9935 \quad S_{y.x} = 12.0888 \quad \text{D.W.} = 1.9345 \end{aligned}$$

Contributions to social insurance are dependent upon local wages and salaries (WSY) as well as the national tax rate for this item (OASI). Total wages and salaries have a strong, positive effect upon these contributions. As wages rise, social insurance collections also increase. The national tax rate indicates that contributions will increase as that portion of wages that is taxable is also increased. Other variables that were tried include non-agricultural employment and the U. S. ratio of social insurance contributions to total wages and salaries.

Wage and Salary Income (WSY).

$$\begin{aligned} \text{WSY} = & (\text{MFGWS} \times \text{MFGE}) + (\text{CCWS} \times \text{CCE}) + (\text{MINWS} \times \text{MINE}) + \\ & (\text{WRTWS} \times \text{WRTE}) + (\text{TCPUWS} \times \text{TCPUE}) + (\text{FIREWS} \times \text{FIRE}) + \\ & (\text{SERWS} \times \text{SE}) + (\text{SLGWS} \times \text{SLGE}) + (\text{FGCWS} \times \text{FGCE}) + \\ & \text{FMWS} + \text{FWS} + \text{OIWS} & (0.45) \end{aligned}$$

Total wage and salary income for Oklahoma is computed by an identity. Average annual wages for the main industrial sectors are multiplied by

their respective employment totals to arrive at a total disbursements figure for each sector. These figures are summed along with farm wages and salaries (FWS), federal military wages and salaries (FMWS), and other industries' wages and salaries (OIWS) to derive the total wage and salary computation. The latter two wage components are exogenous to the model.

Personal Income (PYA).

$$PYA = WSY + OLY + NPY + DIRY + TRY + RAY + FPI - CSSY \quad (0.46)$$

The major aggregate of this block is the personal income variable. It is arrived at by summing the above mentioned income sources along with two exogenous income sources. These exogenous variables are the resident adjustment factor (RAY) and farm proprietors income (FPI). In addition, contributions to social insurance is subtracted out of this summation process.

Real Personal Income (PYAD).

$$PYAD = \frac{PYA}{IPDC} \quad (0.47)$$

Real personal income is formed by deflating nominal personal income by the implicit price deflator for consumption (IPDC).

Fiscal Year Personal Income (P).

$$P = \frac{PYA + PYA1}{2} \quad (0.48)$$

Fiscal year personal income is formed by averaging income over the present and preceding years. This variable is needed for the tax revenue block. This procedure puts the variables of the tax block on a more consistent framework.

Disposable Personal Income (DPYA).

$$DPYA = PYA - FIT - TSIT \quad (0.49)$$

Disposable personal income is found by subtracting from personal income (PYA) the two sources of income taxes, federal (FIT) and state (TSIT). This variable provides information on that portion of total income which can actually be disposed of as the household sees fit.

Real Disposable Personal Income (DPYAD).

$$DPYAD = \frac{DPYA}{IPDC} \quad (0.50)$$

Real disposable personal income is arrived at by deflating nominal disposable personal income by the implicit price deflator for consumption (IPDC).

Fiscal Year Disposable Personal Income (D).

$$D = \frac{DPYA + DPYA1}{2} \quad (0.51)$$

This fiscal year variable is formed by averaging income over the present and preceding years.

Ratio of Other Labor Income to Total Wage and Salary Income
(OKOLWS).

$$OKOLWS = \frac{OLY}{WSY} \quad (0.52)$$

This variable is made up of the ratio of other labor income in Oklahoma to total wage and salary income in Oklahoma.

Tax Revenue

The tax revenue block of the model is partitioned into eight equations. Seven of the equations are behavioral relations while one equation is an identity. The tax block is primarily composed of state tax variables but there is one behavioral equation pertaining to federal income tax collections. The key aggregate for this section is the total state tax collections. This variable is composed of six types of taxes that are individually estimated along with two exogenous tax measures. These exogenous variables are gross production taxes (GPT) and all other taxes (OTHER). The endogenous variables of this block are all in a fiscal year framework except for federal collections. The equations are now presented.

Sales and Use Taxes (SUT).

$$SUT = - 50.31779 + .00517122 D + 121.9344 I \quad (0.53)$$

$$(-2.513) \quad (2.180) \quad (2.958)$$

$$R^2 = .9968 \quad \bar{R}^2 = .9963 \quad S_{y.x} = 1.9647 \quad D.W. = 1.4914$$

Sales and use tax collections are dependent upon disposable personal income in the fiscal year (D) and the fiscal year implicit price deflator

for consumption (I). The positive sign on the disposable income variable indicates that as households have increased amounts of income to dispose of as they desire, sales and use taxes will increase. The positive sign on the price deflator denotes the fact that inflationary increases will also raise sales and use tax collections. This latter occurrence is easily seen in light of the fixed sales tax rate Oklahoma has observed during the modeled period. Other variables that were attempted in the estimation include personal income, population, non-agricultural employment, and gross state product.

Total State Income Taxes (TSIT).

$$TSIT = - 57.145307 + .0183311 P \quad (0.54)$$

$$(-5.658) \quad (15.552)$$

$$R^2 = .9380 \quad \bar{R}^2 = .9335 \quad S_{y.x} = 8.4184 \quad D.W. = 1.5929 \quad \rho = .4786$$

The total state income tax variable is determined by fiscal year personal income (P). The positive sign on this variable indicates that as personal income rises, income tax collections will also increase. Several other variables were also tried in the estimation of this equation. Population, the implicit price deflator for consumption, and even quadratic forms of personal income were tried in the attempt to gain a better predicting equation. In addition, attempts were made to control for tax law changes by usage of dummy variables. Slope and intercept dummies were used to allow for the effects changes in tax law can have upon the determinants of total state income taxes. These attempts all proved unsatisfactory. A serial correlation problem was corrected using the Cochrane-Orcutt technique (13).

Tobacco Products Taxes (TPT).

$$\begin{aligned} \text{TPT} = & - 61.7441 + .0950102 \text{ CRI} + .000925815 \text{ P} + \\ & (-1.764) \quad (5.195) \quad (1.777) \\ & .0265406 \text{ PO} \\ & (1.646) \end{aligned} \quad (0.55)$$

$$R^2 = .9751 \quad \bar{R}^2 = .9693 \quad S_{y.x} = 2.0957 \quad \text{D.W.} = 2.3499$$

The tobacco products tax collections are determined by a tax rate (CRI), a tax base (P), and fiscal year state population (PO). The tax rate variable is the cigarette rate index. The positive relationship denoted by this variable indicates that as the rate is increased, tax collections will rise. Fiscal year personal income serves as the tax base. It also portrays a positive association with tobacco products taxes. In a similar light, tax collections are found to rise when state population increases. Disposable personal income, non-agricultural employment, gross state product, the cigarette wholesale price index, and the implicit price deflator for consumption were also tried in the estimation process.

Alcoholic Beverage Taxes (ABT).

$$\begin{aligned} \text{ABT} = & 10.17494 - .266408 \text{ ABWPI} + .181028 \text{ ABRI} + .00219897 \text{ P} \quad (0.56) \\ & (2.403) \quad (-5.198) \quad (10.250) \quad (10.958) \end{aligned}$$

$$R^2 = .9911 \quad \bar{R}^2 = .9887 \quad S_{y.x} = .7239 \quad \text{D.W.} = 1.5298$$

Alcoholic beverage taxes are dependent upon a tax rate (ABRI), a tax base (P), and a price index characteristic to this item (ABWPI). The alcoholic beverage rate index serves as the appropriate tax rate

for this item. This variable displays a positive relationship with the total tax collections. Fiscal year personal income serves as the tax base and it also shows a positive association with alcoholic beverage taxes. The price index variable is the alcoholic beverage wholesale price index. The negative sign on this variable indicates that as prices rise for this item, tax collections will decrease. Other variables that were tried include disposable personal income, population, non-agricultural employment, gross state product, and the implicit price deflator for consumption. This category of tax is estimated over the period 1961-1975. Alcoholic beverage taxes did not exist on the same basis before 1961.

Motor Fuels Taxes (MFT).

$$\begin{aligned} \text{MFT} = & -12.57694 + .00750074 P - .227696 \text{ GPI} + .0260409 \text{ PO} & (0.57) \\ & (-.383) & (8.869) & (-6.332) & (1.808) \\ R^2 = & .9928 & \bar{R}^2 = .9911 & S_{y.x} = 1.8228 & \text{D.W.} = 1.7493 \end{aligned}$$

Motor fuels taxes are determined by fiscal year personal income (P), the gasoline price index (GPI), and fiscal year population (PO). The fiscal year variables, personal income and population, both have a positive impact upon motor fuels taxes. Motor fuels taxes increase when either of these variables increases. The gasoline price index has a negative effect upon these tax collections. As prices of gasoline move upward, motor fuels tax collections tend to decrease. Other variables attempted in the estimation process were disposable personal income, non-agricultural employment, gross state product, and the implicit price deflator for consumption.

Motor Vehicle Excise Taxes (MVET).

$$\text{MVET} = 23.63869 - .316962 \text{ MVWPI} + .0034712 \text{ P} \quad (0.58)$$

$$(5.341) \quad (-5.535) \quad (13.424)$$

$$R^2 = .9795 \quad \bar{R}^2 = .9766 \quad S_{y.x} = .9801 \quad \text{D.W.} = 1.8374$$

Motor vehicle excise taxes are estimated as a function of the motor vehicle wholesale price index (MVWPI) and fiscal year personal income (P). The negative sign on the price index shows that as motor vehicles increase in price, this individual effect tends to decrease motor vehicle excise tax collections. A substitution effect can be expected to be taking place. Fiscal year personal income has a strong, positive influence on motor vehicle excise tax collections. With greater income, more purchases and, therefore, higher tax collections would be expected. Disposable personal income, gross state product, population, the implicit price deflator for consumption, the automobile wholesale price index, and the automobile license rate index are other variables that were tried in the estimation process.

Total State Taxes (ST).

$$\text{ST} = \text{SUT} + \text{TPT} + \text{ABT} + \text{MFT} + \text{MVET} + \text{TSIT} + \text{OTHER} + \text{GPT} \quad (0.59)$$

Total state taxes are derived by summing the above estimated tax collections along with other taxes (OTHER) and gross production taxes (GPT).

Federal Income Taxes (FIT).

$$\text{FIT} = -1168.1042 + .0632152 \text{ PYA} + .537503 \text{ POP} \quad (0.60)$$

$$(-1.021) \quad (3.698) \quad (1.067)$$

$$R^2 = .9349 \quad \bar{R}^2 = .9256 \quad S_{y.x} = 38.0461 \quad \text{D.W.} = 1.5079 \quad \rho = .5454$$

Federal income tax collections are determined by personal income (PYA) and state population (POP). Personal income has a strong, positive effect upon federal income tax collections while population has a positive but weaker effect. The implicit price deflator for consumption was also tried in the estimation process. The Cochrane-Orcutt technique (13) was used to correct for serial correlation.

Miscellaneous

The miscellaneous sector includes three financial equations. These equations measure demand deposits, time deposits, and total loans. These equations do not interact with other blocks of the model. This block is, in effect, separate from the other parts of the model. The equations now follow.

Demand Deposits (DD).

$$\text{DD} = 1295.474 + .258195 \text{ PYA} - 55.09383 \text{ MAAA} \quad (0.61)$$

$$(19.440) \quad (18.597) \quad (-2.126)$$

$$R^2 = .9932 \quad \bar{R}^2 = .9923 \quad S_{y.x} = 64.8027 \quad \text{D.W.} = 1.7191$$

Demand deposits are determined by personal income (PYA) and Moody's Aaa bond rate (MAAA). Personal income has a large, positive impact upon demand deposits. Demand deposits increase when personal income

risers. Moody's bond rate poses as an alternative financial interest rate. It represents rates on alternative financial instruments. The negative coefficient for this variable implies that as interest rates for alternative instruments rise, households cut back on their demand deposits.

Time Deposits (TD).

$$TD = -.1338.408 + .600542 \text{ PYA} - 160.0072 \text{ MAAA} \quad (0.62)$$

$$(-13.461) \quad (28.990) \quad (-4.139)$$

$$R^2 = .9970 \quad \bar{R}^2 = .9966 \quad S_{y.x} = 96.6897 \quad D.W. = 2.3707$$

Time deposits are modeled in an identical manner to that of demand deposits. Personal income (PYA) and Moody's Aaa bond rate (MAAA) serve as the determinants of time deposits. Personal income shows a strong, positive influence while Moody's bond rate has a negative impact upon time deposits.

Total Loans (TL).

$$TL = -802.23321 + .808308 \text{ PYA} - 132.3451 \text{ FRDR} \quad (0.63)$$

$$(-5.942) \quad (37.660) \quad (-2.765)$$

$$R^2 = .9957 \quad \bar{R}^2 = .9951 \quad S_{y.x} = 139.5148 \quad D.W. = 1.4075 \quad \rho = .2032$$

Total loans are dependent upon personal income (PYA) and the Federal Reserve discount rate (FRDR). Personal income has a strong, positive effect upon total loans implying that increases in personal income will cause increases in total loans. The Federal Reserve discount rate has a negative association with total loans. As the discount rate is raised to invoke a tighter monetary policy, total loans are decreased due to

the increased cost of borrowing money from the Federal Reserve System. Other variables included in the estimation process were the prime commercial paper rate and a variable measuring the difference in the Federal Reserve discount rate and the prime commercial paper rate.

Model Properties

The basic properties of the OLS Oklahoma econometric model can be compared in general with those of some of the other existing econometric models at this time. Table I provides a summary of the basic properties of some selected regional econometric models along with those of the Oklahoma model.

The Oklahoma model was estimated from 18 annual observations (row 1) and displays 63 equations (row 2). The number of observations compares quite closely to that number used for many of the other studies. The range is from eight to 25. The total equation figure is just less than that of the Northeast Corridor model and greater than all others except the Philadelphia models. Equations range from 14 to 228. For the size of this model, the number of bivariate specifications (row 4) is quite low (6) and the number of stochastic equations with lags (row 5) is fairly high (15). Bivariate specifications number from six to 40 while lagged equations measure from zero to 37. In a proportionate sense, the number of stochastic equations (row 3) for the Oklahoma model (45) is quite similar to those of the other models. Stochastic equations range from eight to 105. The number of actual exogenous variables (row 6) used in the model (29) is relatively large compared to the other models. Exogenous variables number from one to 49. The Oklahoma model has 91% of its stochastic equations showing a \bar{R}^2 value

TABLE I
A STATISTICAL SUMMARY OF THE OKLAHOMA MODEL AND
OTHER SELECTED REGIONAL ECONOMETRIC MODELS

	Massa- chusetts	North- east Corridor	Phila- delphia I	Phila- delphia III	Phila- delphia IV	Puerto Rico	Buffalo	Southern Cali- fornia	Los Angeles	Missis- sippi	Oklahoma ^a	Oklahoma ^b
1. Number of Observations	16	15	17	24	25	17	8-17	9	12	18	18	18
2. Number of Equations	14	66	26	178	228	35	35	18	29	50	33	63
3. Number of Stochastic Equations	8	60	17	91	105	23	23	13	19	40	25	45
4. Number of Bivariate Specifications	6	40	11	15	19	6	15	10	12	6	8	6
5. Number of Stochastic Equations with Lags	2	10	1	29	37	5	4	1	0	28	5	15
6. Number of Exogenous Variables	6	20	5	22	49	22	14	3	4	56	1	29
7. Number of Equations with $\bar{R}^2 \geq 0.90$	5	44	9	53	69	21	14	9	16	38	21	41
8. Number of Equations with $\bar{R}^2 < 0.70$	1	8	1	2	0	1	1	2	0	0	1	0
9. ^c Estimation Techniques	OLS RFLS	OLS TSLS-PC	OLS LISE TSLS	OLS TSLS-PC	OLS	OLS	OLS	OLS	OLS	OLS IIV	OLS	OLS INST-PC

Source: (24, p. 62).

^aThis Oklahoma model was built by Chong K. Liew and Dae K. Kahng in 1971 (42).

^bThis is the estimated model of Oklahoma which is constructed in this manuscript. This will be the model referred to when the Oklahoma model is mentioned.

^cOLS represents ordinary least squares; TSLS stands for two-stage least squares; TSLS-PC denotes TSLS used with principal components in the first stage; RFLS indicates reduced form least squares; LISE represents limited information single-equation; IIV stands for iterated instrumental variables; and INST-PC denotes instrumental variables used with principal components in the first stage.

greater than or equal to .90. In actual numbers, 41 of the 45 equations have this property (row 7). The range for this category is from 53% to 91%. None of the equations for the Oklahoma model have \bar{R}^2 values less than .70 (row 8). The range, for this characteristic of the models is from zero to 15%. Row 9 shows the estimation techniques used in these selected models.

Ex Post Simulation

After completion of the initial estimation of the OLS model, it is necessary to first test the model as to its ability to replicate economic activity during the 1961-1975 sample period.⁸ This testing procedure will be accomplished by simulating the estimated model over this time period and by comparing the generated endogenous variables with the values that actually occurred. The actual and generated endogenous variables are compared using various simulation error statistics.

The error statistics that are used in this section to determine the accuracy of prediction of the OLS model are the mean absolute error (MAE), the mean absolute percentage error (MAPE), the root mean square error (RMSE), and Theil's "U" coefficient. These concepts may be visualized in the following manner:

$$MAE_i = \sum_{j=1}^N \frac{|y_{ij}^p - y_{ij}^a|}{N} \quad (4.2)$$

$$MAPE_i = \sum_{j=1}^N \frac{|y_{ij}^p - y_{ij}^a|}{N} \cdot \frac{1}{y_{ij}^a} \cdot 100 \quad (4.3)$$

⁸The sample period 1961-1975 is used for simulation because prior values of alcoholic beverage taxes are not consistent.

$$RMSE_i = \sqrt{\frac{\sum_{j=1}^N (y_{ij}^p - y_{ij}^a)^2}{N}} \quad (4.4)$$

$$U_i = \sqrt{\frac{\sum_{j=1}^N (\Delta y_{ij}^p - \Delta y_{ij}^a)^2}{\sum_{j=1}^N (\Delta y_{ij}^a)^2}} \quad (4.5)$$

where $i = 1, \dots, M$; $j = 1, \dots, N$; M = number of endogenous variables; $y_{ij}^p = i^{th}$ predicted endogenous variable; $y_{ij}^a = i^{th}$ actual endogenous variable; and N = number of observations (24, p. 68). Since no statistical tests exist in which to determine the significance of the error statistics mentioned above, it is usually necessary to compare one's error statistics with those of other models to determine their worth. The MAPE statistical error is the one most commonly used to pursue this goal. Since some models are estimated in different units, its percentage error format makes it highly desirable for comparison purposes.

The simulation error statistics for the period 1961-1975 are presented in Table II. Values covering all 63 of the equations are included for all four of the error statistics. The output block shows a range of MAPE errors from 1.26% for federal government output to 4.63% for manufacturing output. The employment block lists MAPE errors from .78% for service employment to 5.43% for contract construction employment. The civilian labor force shows a MAPE error of 1.35% while the unemployment rates show errors of between 5% and 7%. The wage and salary block shows excellent simulation results with MAPE errors of between .82% for manufacturing and 3.90% for mining. Income measures MAPE errors of 1.86% for dividends, interest, and rents as a low while non-farm

TABLE II
OKLAHOMA ECONOMETRIC STATE MODEL OLS SIMULATION
ERROR STATISTICS, 1961-1975

Equation	MAE	MAPE	RMSE	Theil's U
QMFG	76.17	4.63%	96.84	.02634
QCC	20.64	3.77%	26.77	.02482
QWRT	56.95	3.27%	70.82	.01989
QFIRE	24.59	1.78%	31.72	.01132
QTCPU	13.07	1.27%	17.09	.00835
QSER	17.67	1.79%	23.46	.01074
QSLG	12.65	1.52%	16.51	.00997
QFG	13.68	1.26%	17.56	.00821
MFGE	3.408	2.91%	3.807	.01552
CCE	2.047	5.43%	2.702	.03534
MINE	1.229	3.18%	1.616	.01987
WRTE	1.835	1.14%	2.550	.00759
FIREE	.5088	1.33%	.6652	.00917
TCPUE	.8488	1.68%	1.092	.01076
SE	.7901	.78%	.9886	.00459
SLGE	2.612	2.21%	3.249	.01355
FGCE	.9849	1.80%	1.158	.01088
CLFA	13.63	1.35%	16.37	.00794
OKAUNR	.2520	6.82%	.3284	.04083
OKIUNR	.1605	5.10%	.2197	.03289
MFGWS	.0538	.82%	.0574	.00408
CCWS	.1921	2.81%	.2398	.01644
MINWS	.3117	3.90%	.3485	.01925
WRTWS	.0873	1.77%	.0986	.00995
FIREWS	.1046	1.71%	.1373	.01134
SERWS	.0751	1.68%	.0923	.00924
FGCWS	.2974	3.87%	.3650	.02091
SLGWS	.1214	2.52%	.1365	.01367
TCPUWS	.1602	2.10%	.1975	.01186
FWS	.5624	1.26%	.7190	.00700
OLY	8.944	3.78%	11.14	.01613
NPY	38.96	6.19%	46.40	.03456
DIRY	19.59	1.86%	24.80	.00988
TRY	16.60	2.07%	19.81	.00913
CSSY	8.784	3.75%	11.83	.01968
Z	.00732	.82%	.0080	.00449
SUT	1.652	1.76%	2.128	.01092
TSIT	8.859	14.06%	10.65	.05509
FIT	32.86	5.41%	41.38	.02796
TPT	1.418	4.97%	1.938	.02921
ABT	.4445	2.26%	.6233	.01563
MFT	1.295	1.52%	1.554	.00899
MVET	.7846	5.28%	.9928	.02980

TABLE II (Continued)

Equation	MAE	MAPE	RMSE	Theil's U
DD	60.22	2.11%	69.24	.01121
TD	75.35	3.36%	108.9	.01872
TL	128.3	2.26%	170.5	.01551
GSP	161.6	1.50%	204.0	.00936
NAE	5.107	.72%	5.769	.00392
RELAUNR	.0494	6.84%	.0616	.04118
WRTWSD	.0947	1.77%	.1042	.00975
FIREWSD	.1129	1.71%	.1424	.01088
TCPUWSD	.1823	2.10%	.2267	.01283
SLGWSD	.1357	2.52%	.1517	.01412
OKOLWS	.00181	3.35%	.00216	.01833
WSY	30.19	.72%	44.08	.00431
PYA	66.99	.91%	87.66	.00515
DPYA	69.50	1.04%	102.3	.00666
PYAD	75.89	.91%	99.18	.00567
DPYAD	78.56	1.04%	115.5	.00731
ST	9.554	2.53%	12.21	.01355
PCNAE	.00790	42.77%	.00917	.1490
D	63.83	.94%	90.34	.00616
P	60.79	.83%	77.89	.00479

proprietors income at 6.19% represents the high. The tax block shows a large range as motor fuels taxes at 1.52% represents the low MAPE error for the block while total state income taxes at 14.06% walks away with the high. Actually, the high for the tax block is federal income taxes at 5.41% if the total state income tax equation is excluded. Many forms of the total state income tax equation were estimated but none were satisfactory in reducing this error. The miscellaneous block shows MAPE errors at between 2.11% for demand deposits and 3.36% for time deposits.

The major aggregates of the OLS model were all simulated with fairly good accuracy. Gross state product showed a simulation MAPE error of 1.50% while non-agricultural employment was much lower at .72%. Wage and salary income simulated at .72% error while personal income was minutely higher at .91%. Total state taxes produced a MAPE error of 2.53%. Most of this latter error is due to the total state income tax equation.

Table III summarizes the MAPE error findings of the OLS model by providing an error distribution for the model. The results show that 85.7% of the equations were simulated with MAPE errors of less than 5%. Roughly half of the equations had MAPE errors of less than 2% while over two-thirds of the model showed MAPE errors of less than 3%. These results are quite comparable to those of some of the other regional econometric models which are presented in Table IV and Table V.

Table IV shows a comparison of MAPE errors over several models for five key variables. These variables are gross regional (state) product, personal income, total employment, manufacturing output, and manufacturing employment. The Oklahoma MAPE errors are all relatively low for these variables except for manufacturing output. In this category, the

TABLE III
OLS MAPE ERROR DISTRIBUTION

Error Distribution	Number	Percent	Cumulative Percent
0 - .99%	9	14.3	14.3
1 - 1.99%	22	34.9	49.2
2 - 2.99%	12	19.0	68.3
3 - 3.99%	9	14.3	82.5
4 - 4.99%	2	3.2	85.7
5% and over	9	14.3	100.0

TABLE IV
COMPARISON OF MAPE ERRORS OVER SELECTED VARIABLES AND REGIONAL MODELS

Variable	Tennessee	Georgia	Missis- sippi	Philadel- phia I	Philadel- phia IV	NE Corridor	Buffalo	Los Angeles	Oklahoma
Gross Regional Product	.96%	2.52%	1.28%	6.32%	.98%	2.05%	1.87%	2.08%	1.50%
Personal Income	.91%	2.06%	.68%	6.69%	1.55%	3.13%	8.42%	1.45%	.91%
Total Employment	1.01%	1.52%	.33%	1.56%	.66%	1.40%	3.39%	.88%	.72% ^a
Manufacturing Output	1.49%	n.a.	3.22%	2.43%	2.19%	2.82%	1.55%	3.07%	4.63%
Manufacturing Employment	2.13%	3.09%	1.96%	2.18%	1.42%	2.65%	3.50%	2.81%	2.91%

Source: Chang (11, p. 60).

^aTotal employment is not estimated in the Oklahoma model. The listed value represents the next highest aggregate, non-agricultural employment.

Oklahoma MAPE error of 4.63% is higher than the range of 1.49% to 3.22% for the other models.

Table V shows that roughly 92%, 89%, 82%, and 68% of the Mississippi, Tennessee, Philadelphia IV, and Georgia models, respectively, have MAPE errors of less than 5%. The Oklahoma model generated approximately 86% of its model under this 5% MAPE error criterion. In a relative sense, the Oklahoma model seems to be slightly better than half and slightly worse than half of the above mentioned models. On the whole, the Oklahoma model showed very good ex post simulation results.

Ex Post Forecasting

A further test of the model can be made by simulating out of the sample period for periods of time in which historical endogenous variables are now available. The only year not in the sample period for which most of the variables are now available is 1976. Therefore, a simulation was performed for the year 1976 and a series of MAPE error statistics were computed to determine the replication abilities of the model.

Table VI presents the predicted values, the actual values, the error, and MAPE statistic for each endogenous variable. The predicted value for 1976 will be the sole entry for variables in which actual values are not yet available.

The MAPE error statistics for 1976 are for the most part very satisfactory. Output MAPE errors range from .58% for the transportation, communication, and public utilities sector to 5.76% for federal government output. The employment sector shows a low of .47% in the state and local government sector and values below 5% for all other sectors except

TABLE V
PERCENTAGE DISTRIBUTION OF MAPE ERRORS FROM SAMPLE PERIOD SIMULATION

Percentage Error	Tennessee	Mississippi	Philadelphia IV	Georgia	Oklahoma
0 - .99	16.67	13.46	6.3	8.9	14.9
1 - 1.99	34.72	32.69	22.0	13.7	34.9
2 - 2.99	22.22	13.46	31.5	18.5	19.0
3 - 3.99	8.33	17.31	11.8	19.4	14.3
4 - 4.99	6.94	15.59	10.2	7.3	3.2
5 or greater	11.11	7.69	18.1	31.6	14.3
TOTAL	100.00	100.00	100.0	100.0	100.0

Source: Chang, (11, p. 62).

TABLE VI
1976 OLS SIMULATION ERRORS

Equation	Predicted	Actual	Error	MAPE
QMFG	2822.50	2722.80	99.7	3.66%
QCC	628.33	652.40	-24.07	3.69%
QWRT	2447.15	2468.00	-20.85	.84%
QFIRE	1860.78	1907.60	-46.82	2.45%
QTCPU	1444.85	1436.50	8.35	.58%
QSER	1440.24	1460.00	-19.76	1.35%
QSLG	1054.72	1046.00	8.72	.83%
QFG	947.46	1005.40	-57.94	5.76%
MFGE	157.67	156.10	1.57	1.01%
CCE	44.099	46.10	- 2.001	4.34%
MINE	39.69	44.40	- 4.71	10.61%
WRTE	217.06	222.20	- 5.14	2.31%
FIREE	47.829	46.60	1.23	2.64%
TCPUE	56.704	57.30	- .596	1.04%
SE	147.06	151.40	- 4.34	2.87%
SLGE	158.94	158.20	.74	.47%
FGCE	48.351	48.80	- .449	.92%
CLFA	1161.08	1159.00	2.08	.18%
OKAUNR	6.276	5.60	.676	12.07%
OKIUNR	2.787			
MFGWS	10.918	11.023	- .105	.95%
CCWS	12.351	11.238	1.113	9.90%
MINWS	15.829	15.627	.202	1.29%
WRTWS	7.941	7.497	.444	5.92%
FIREWS	9.432	9.415	.017	.18%
SERWS	7.845	7.690	.155	2.02%
FGCWS	15.043	14.899	.144	.97%
SLGWS	7.921	7.562	.359	4.75%
TCPUWS	14.004	13.729	.275	2.00%
FWS	102.389	102.219	.170	.17%
OLY	814.12	826.247	-12.127	1.47%
NPY	1045.81	1056.324	-10.514	.99%
DIRY	2373.42	2363.552	9.868	.42%
TRY	2381.83	2393.674	-11.844	.49%
CSSY	633.64	599.56	34.08	5.68%
Z	.8708	.8792	- .008	.96%
SUT	178.258	181.865	- 3.607	1.98%
TSIT	218.504	228.221	- 9.717	4.26%
FIT	1319.12			
TPT	49.615	50.391	- .776	1.54%
ABT	31.885	32.591	- .706	2.17%
MFT	118.389	117.256	1.133	.97%
MVET	29.433	28.400	1.033	3.64%
DD	4917.45	4656.42	261.03	5.61%
TD	6817.43	6879.43	-62.00	.90%
TL	11262.90	11214.01	48.89	.44%

TABLE VI (Continued)

Equation	Predicted	Actual	Error	MAPE
GSP	14169.30	14222.00	-52.70	.37%
NAE	917.401	931.10	-13.699	1.47%
RELAUNR	.8150	.7270	.088	12.10%
WRTWSD	5.962	5.628	.334	5.93%
FIREWSD	7.081	7.068	.013	.18%
TCPUWSD	10.514	10.307	.207	2.01%
SLGWSD	5.947	5.677	.270	4.76%
OKOLWS	.0859	.0881	- .0022	2.50%
WSY	9473.23	9380.88	92.35	.98%
PYA	15826.90	15793.22	33.68	.21%
DPYA	14289.20			
PYAD	11882.00	11856.77	25.23	.21%
DPYAD	10727.70			
ST	849.574	862.214	-12.64	1.47%
PCNAE	.0338	.0492	- .0154	31.30%
D	13583.80			
P	15037.30			

mining. Mining registers a MAPE error of 10.61%. The volatility of this sector does make for high errors in some individual years. The civilian labor force shows a MAPE error of .18% while the unemployment rate soared to a 12.07% error. The farm sector produces the low MAPE error of the wage and salary sector at .17% while contract construction registers the high at 9.90%. Again, construction is a volatile sector that sometimes produces high errors in any given year. The income block shows very low MAPE errors as most equations have errors less than 1%. The tax block also shows low MAPE error statistics. Motor fuels taxes registers the low at .97% while total state income taxes records the high at 4.26%. The financial equations are low except for demand deposits which shows a 5.61% MAPE error.

The key aggregates again show low MAPE statistics for this 1976 simulation. Gross state product shows an error of .37% while non-agricultural employment is somewhat higher at 1.47%. Wage and salary income records a MAPE error of .98% while personal income is very low at .21%. State taxes register a small MAPE error at 1.47%.

An alternative approach to forecasting for 1976 with the OLS model was also attempted. This alternative, suggested by Arthur S. Goldberger (28), attempts to improve on the predictability of equations containing errors that are autocorrelated. The basic addition to each serially correlated equation is a term multiplying the first-order autocorrelation coefficient (ρ) times the error in the last period (e_T). The forecasted endogenous value for the next period is then visualized as:

$$\hat{Y}_{T+1} = \hat{\beta}X_{T+1} + \rho e_T \quad (4.6)$$

where \hat{Y}_{T+1} represents a forecasted endogenous variable in the T+1 period, X_{T+1} indicates an exogenous variable in the T+1 period, and $\hat{\beta}$ denotes the estimated coefficient for the exogenous variable (28, p. 373).

The 16 equations of the model in which serial correlation had been present were altered in this fashion. Estimated ρ values and errors from 1975 observations were used to make the correction. The altered OLS model was then simulated over 1976 again to see if the simulation error for the equations would be reduced. Table VII discloses the result of this additional simulation.⁹

The simulation using the Goldberger alternative produced 19 equations with lower prediction errors, 33 equations with higher prediction errors, and 11 equations that did not change or could not be compared. Of the 16 equations in which serial correlation was initially present, five of those equations showed lower prediction errors while ten equations had larger errors and one equation could not be compared.

Simultaneous Equation Model

The Oklahoma model was also estimated using a simultaneous equation estimator. The method used in this context was an instrumental regression (INST) combined with principal components (PC). The OLS model which was estimated previously was used as the specification form for the simultaneous equation estimator. This use of identical formats facilitates comparisons between the two estimation techniques as to which simulates and forecasts better. The same testing procedures which

⁹ Appendix B presents the results of using this technique for the within sample period simulation of 1961-1975.

TABLE VII

1976 OLS SIMULATION ERRORS: GOLDBERGER ALTERNATIVE

Equation	Predicted	Actual	Error	MAPE
QMFG	2814.38	2722.80	91.58	3.36%
QCC	628.16	652.40	-24.24	3.72%
QWRT	2448.44	2468.00	-19.56	.79%
QFIRE	1861.59	1907.60	-46.01	2.41%
QTCPU	1445.59	1436.50	9.09	.63%
QSER	1440.88	1460.00	-19.12	1.31%
QSLG	1054.86	1046.00	8.86	.85%
QFG	947.46	1005.40	-57.94	5.76%
MFGE	154.47	156.10	-1.63	1.04%
CCE	43.191	46.10	-2.909	6.31%
MINE	39.69	44.40	-4.71	10.61%
WRTE	216.72	222.20	-5.48	2.47%
FIREE	47.833	46.60	1.233	2.65%
TCPUE	55.933	57.30	-1.367	2.39%
SE	147.07	151.40	-4.33	2.86%
SLGE	158.95	158.20	.75	.47%
FCCE	48.351	48.80	-.449	.92%
CLFA	1161.08	1159.00	2.08	.18%
OKAUNR	6.310	5.60	.71	12.68%
OKIUNR	2.895			
MFCWS	10.948	11.023	-.075	.68%
CCWS	12.433	11.238	1.195	10.63%
MINWS	16.085	15.627	.458	2.93%
WRTWS	8.062	7.497	.565	7.54%
FIREWS	9.450	9.415	.035	.37%
SERWS	7.893	7.690	.203	2.64%
FGCWS	15.127	14.899	.228	1.53%
SLGWS	7.919	7.562	.357	4.72%
TCPUWS	14.089	13.729	.36	2.62%
FWS	102.416	102.219	.197	.19%
OLY	823.28	826.247	-2.967	.36%
NPY	1038.06	1056.324	-18.264	1.73%
DIRY	2377.85	2363.552	14.298	.60%
TRY	2382.59	2393.674	-11.084	.46%
CSSY	633.74	599.56	34.18	5.70%
Z	.8731	.8792	-.006	.69%
SUT	178.291	181.865	-3.574	1.97%
TSIT	226.587	228.221	-1.634	.72%
FIT	1306.38			
TPT	49.618	50.391	-.773	1.53%
ABT	31.893	32.591	-.698	2.14%
MFT	118.418	117.256	1.162	.99%
MVET	29.446	28.400	1.046	3.68%

TABLE VII (Continued)

Equation	Predicted	Actual	Error	MAPE
DD	4919.46	4656.42	263.04	5.65%
TD	6822.12	6879.43	-57.31	.83%
TL	11209.90	11214.01	-4.11	.04%
GSP	14164.70	14222.00	-57.30	.40%
NAE	912.206	931.10	-18.894	2.03%
RELAUNR	.8194	.7270	.0924	12.71%
WRTWSD	6.052	5.628	.424	7.53%
FIREWSD	7.094	7.068	.026	.37%
TCPUWSD	10.577	10.307	.27	2.62%
SLGWSD	5.945	5.677	.268	4.72%
OKOLWS	.0869	.0881	-.0012	1.36%
WSY	9474.55	9380.88	93.67	1.00%
PYA	15834.70	15793.22	41.48	.26%
DPYA	14301.70			
PYAD	11887.90	11856.77	31.13	.26%
DPYAD	10737.00			
ST	857.745	862.214	-4.469	.52%
PCNAE	.0280	.0492	-.0212	43.09%
D	13590.00			
P	15041.20			

were used on the OLS model were also used on the simultaneous equation estimation model.

The simultaneous equation model is composed of 63 equations with 45 of the equations showing behavioral relations and 18 representing identities. This model is estimated over the period extending from 1961-1975. The alcoholic beverage tax series prevented estimation from beginning at 1958.

A direct simultaneous equation estimation of the Oklahoma model is initially hampered by the fact that the model contains 18 observations and 45 exogenous variables. When the number of observations is less than or equal to the number of exogenous variables, ". . . then the moment matrix in the least squares estimating procedure will be singular and estimates cannot be found" (26, p. 23).

When confronted with this problem, analysts have traditionally done one of two things. First, then have omitted some of the exogenous variables from the model . . . The second method . . . entails the use of principal components (26, pp. 23-24).

The first method of dealing with this problem seems highly unsatisfactory since specification error is very likely to occur. The most reasonable method of handling this problem would seem to be some sort of usage of principal components.

"Principal components are a set of linear combinations of the vector of exogenous variables which are mutually orthogonal" (26, p. 24). The intent behind the usage of principal components is to capture within a few principal components a majority of the variance contained within all of the exogenous variables.¹⁰ The information within the exogenous

¹⁰ A deeper discussion of principal component analysis can be found in reference (44).

variables can then be reduced to a dimensionality which is more compatible with the number of observations available for estimation. If the 45 predetermined variables can be reduced to between four and eight principal components, then estimation can take place.

A principal component analysis was, therefore, conducted on a large subset of the exogenous variables. The entire set of exogenous variables could not be used due to the limitations of the available computer packages. A large subset was chosen according to what generally represented the movements of the exogenous variables. The information within the exogenous variables was reduced to a group of six principal components. These six principal components explained over 99% of the variance in the exogenous variables.

The six principal components were then used as instruments in an instrumental regression estimation process (INST). The purpose behind the instrumental regression process is to hopefully purge the (endogenous) explanatory variables of their correlation with the error terms (33, p. 381). The formula for the estimation of the regression coefficients in this type of equation can be denoted as:

$$b = [W'Z(Z'Z)^{-1}Z'W]^{-1} W'Z(Z'Z)^{-1}Z'y$$

where b = vector of regression coefficients,

W = matrix of right-hand variables (other endogenous variables also in the equation),

Z = matrix of instrumental variables, and,

y = left-hand variable (the primary endogenous variable).

In summary, the instrumental regression process is undertaken by first doing a principal component analysis on a subset of the exogenous

variables. Secondly, the six principal components which were extracted from the above process were used as instruments in an instrumental simultaneous equation regression.

All of the equations of the model were found to be identified by use of the order condition. It must be remembered that the order condition for identification is a necessary but not a sufficient condition for identification.

The INST model equations will now be presented in identical order to that which was used for the OLS model equations. Since the model specifications are identical, discussion following each equation will not be presented. Statistics for each equation will include the following: *t* statistics (below each coefficient), coefficients of determination (R^2), standard errors of the model ($S_{y.x}$), and Durbin-Watson statistics (D.W.).

Output

Manufacturing Output (QMFG).

$$QMFG = 3186.41 + 2.81794 \text{ RGDPMA} - 3844.77 \text{ Z} + .783142 \text{ QMFG1} \quad (\text{I.1})$$

$$(1.942) \quad (2.732) \quad (-2.174) \quad (8.829)$$

$$R^2 = .9825 \quad S_{y.x} = 87.9660 \quad \text{D.W.} = 2.6736$$

Contract Construction Output (QCC).

$$QCC = 191.487 + .0450393 \text{ GSP} - 22.475 \text{ MAAA} \quad (\text{I.2})$$

$$(3.834) \quad (4.455) \quad (-2.021)$$

$$R^2 = .8325 \quad S_{y.x} = 23.9568 \quad \text{D.W.} = 2.0818$$

Wholesale and Retail Trade Output (QWRT).

$$QWRT = 10.7738 + .142652 \text{ DPYAD} + .377246 \text{ QWRT1} \quad (\text{I.3})$$

$$(.120) \quad (2.410) \quad (1.279)$$

$$R^2 = .9745 \quad S_{y.x} = 60.1797 \quad D.W. = 1.7466$$

Finance, Insurance, and Real Estate Output (QFIRE).

$$QFIRE = - 1107.75 + .0408056 \text{ DPYAD} + .60272 \text{ POP} + \quad (\text{I.4})$$

$$(-1.261) \quad (1.029) \quad (1.375)$$

$$.487806 \text{ QFIRE}$$

$$(2.059)$$

$$R^2 = .9895 \quad S_{y.x} = 28.1873 \quad D.W. = 2.0714$$

Transportation, Communication, and Public Utilities Output (QTCPU).

$$QTCPU = - 249.962 + .131513 \text{ PYAD} + .0470185 \text{ POP} \quad (\text{I.5})$$

$$(-.420) \quad (7.548) \quad (.161)$$

$$R^2 = .9940 \quad S_{y.x} = 19.9263 \quad D.W. = 2.2358$$

Service Output (QSER).

$$QSER = 126.208 + .110590 \text{ PYAD} \quad (\text{I.6})$$

$$(3.919) \quad (29.983)$$

$$R^2 = .9858 \quad S_{y.x} = 24.4293 \quad D.W. = 1.5909$$

State and Local Government Output (QSLG).

$$QSLG = 66.0837 + .0171357 \text{ DPYAD} + .784658 \text{ QSLG1} \quad (\text{I.7})$$

$$(1.843) \quad (1.660) \quad (5.643)$$

$$R^2 = .9915 \quad S_{y.x} = 11.8858 \quad D.W. = 1.4022$$

Federal Government Output (QFG).

$$QFG = - 1137.98 + 20.17 \text{ RGDPG} + .143924 \text{ QFG1} - 72.0324 \text{ TIME} \quad (\text{I.8})$$

$$(-6.940) \quad (8.091) \quad (1.254) \quad (-8.621)$$

$$R^2 = .9721 \quad S_{y.x} = 19.7273 \quad \text{D.W.} = 2.3080$$

Gross State Product (GSP).

$$\text{GSP} = \text{QMFG} + \text{QCC} + \text{QMIN} + \text{QWRT} + \text{QFIRE} + \text{QTCPU} + \text{QSER} + \text{QSLG} \quad (\text{I.9})$$

$$+ \text{QFG} + \text{QFARM}$$

EmploymentManufacturing Employment (MGFE).

$$\text{MGFE} = 23.2129 + .0147903 \text{ QMFG} + .616954 \text{ MGFE1} \quad (\text{I.10})$$

$$(2.060) \quad (1.713) \quad (2.867)$$

$$R^2 = .9660 \quad S_{y.x} = 4.6755 \quad \text{D.W.} = 1.1025$$

Contract Construction Employment (CCE).

$$\text{CCE} = - 5.46826 + .0587458 \text{ QCC} + .320401 \text{ CCE1} \quad (\text{I.11})$$

$$(-1.099) \quad (3.246) \quad (1.525)$$

$$R^2 = .8713 \quad S_{y.x} = 1.7596 \quad \text{D.W.} = .7697$$

Mining Employment (MINE).

$$\text{MINE} = 4.76371 + .0020348 \text{ QMIN} + .835111 \text{ MINE1} \quad (\text{I.12})$$

$$(.802) \quad (.378) \quad (7.741)$$

$$R^2 = .8350 \quad S_{y.x} = 1.1955 \quad \text{D.W.} = 1.0524$$

Wholesale and Retail Trade Employment (WRTE).

$$\begin{aligned} \text{WRTE} = & 11.4167 + .0254408 \text{ QWRT} - .642487 \text{ WRTWSD} + & (\text{I.13}) \\ & (.648) \quad (3.065) \quad (-.167) \\ & .704586 \text{ WRTE} \\ & (6.148) \end{aligned}$$

$$R^2 = .9971 \quad S_{y.x} = 1.4304 \quad \text{D.W.} = 2.0998$$

Finance, Insurance, and Real Estate Employment (FIREE).

$$\begin{aligned} \text{FIREE} = & 5.48788 + .0270759 \text{ QFIRE} - 1.09115 \text{ FIREWSD} & (\text{I.14}) \\ & (1.349) \quad (18.104) \quad (-1.246) \end{aligned}$$

$$R^2 = .9897 \quad S_{y.x} = .6787 \quad \text{D.W.} = 2.0638$$

Transportation, Communication, and Public Utilities Employment(TCPUE).

$$\begin{aligned} \text{TCPUE} = & 27.8179 + .0144559 \text{ QTCPU} - 1.18322 \text{ TCPUWSD} + & (\text{I.15}) \\ & (2.329) \quad (1.905) \quad (-.742) \\ & .374794 \text{ TCPUE1} \\ & (2.045) \end{aligned}$$

$$R^2 = .9598 \quad S_{y.x} = .8349 \quad \text{D.W.} = 1.0722$$

Service Employment (SE).

$$\begin{aligned} \text{SE} = & - 2.16009 + .0263354 \text{ QSER} + .787139 \text{ SE1} & (\text{I.16}) \\ & (-.948) \quad (1.774) \quad (5.547) \end{aligned}$$

$$R^2 = .9974 \quad S_{y.x} = 1.1690 \quad \text{D.W.} = 1.8906$$

State and Local Government Employment (SLGE).

$$SLGE = - 79.7032 + .333794 QSLG + 9.46778 SLGWSD - \quad (I.17)$$

$$(-1.286) \quad (1.688) \quad (.907)$$

$$1.11136 SLGE1$$

$$(-.831)$$

$$R^2 = .9894 \quad S_{y.x} = 2.4275 \quad D.W. = 2.1817$$

Federal Government Civilian Employment (FGCE).

$$FGCE = 3.05145 + .0302861 QFG + .336754 FGCE1 \quad (I.18)$$

$$(.856) \quad (5.060) \quad (2.889)$$

$$R^2 = .9456 \quad S_{y.x} = 1.2220 \quad D.W. = 1.5496$$

Non-Agricultural Employment (NAE).

$$NAE = MFGE + CCE + MINE + WRTE + FIREE + TCPUE + SE + SLGE \quad (I.19)$$

$$+ FGCE$$

Percentage Change in Non-Agricultural Employment (PCNAE).

$$PCNAE = \frac{NAE - NAE1}{NAE1} \quad (I.20)$$

Civilian Labor Force (CLFA).

$$CLFA = - 490.919 - 7.75209 OKAUNR1 + .45885 POP + \quad (I.21)$$

$$(-1.675) \quad (-1.069) \quad (1.778)$$

$$.379698 CLFA1$$

$$(1.034)$$

$$R^2 = .9635 \quad S_{y.x} = 17.0310 \quad D.W. = 1.9553$$

Unemployment Rate (OKAUNR).

$$\text{OKAUNR} = - 4.94177 + .86716 \text{ USUR} - .000713061 \text{ PYA} + \quad (\text{I.22})$$

$$(-2.347) \quad (10.396) \quad (-2.219)$$

$$1.45344 \text{ MFGWS}$$

$$(2.248)$$

$$R^2 = .9234 \quad S_{y.x} = .3499 \quad \text{D.W.} = 1.5803$$

Relative Unemployment Rate (RELAUNR).

$$\text{RELAUNR} = \frac{\text{OKAUNR}}{\text{USUR}} \quad (\text{I.23})$$

Insured Unemployment Rate (OKIUNR).

$$\text{OKIUNR} = 2.16757 + .675283 \text{ USUR} - .000253374 \text{ PYA} - \quad (\text{I.24})$$

$$(4.772) \quad (9.373) \quad (-11.923)$$

$$16.0568 \text{ PCNAE}$$

$$(-2.416)$$

$$R^2 = .9764 \quad S_{y.x} = .1907 \quad \text{D.W.} = 1.8081$$

Wage and SalaryManufacturing Wages and Salaries (MFGWS).

$$\text{MFGWS} = .350456 + .679448 \text{ USMFGW} - .0318554 \text{ OKIUNR} + \quad (\text{I.25})$$

$$(2.788) \quad (23.883) \quad (-1.957)$$

$$.913343 \text{ FEDMW}$$

$$(5.059)$$

$$R^2 = .9988 \quad S_{y.x} = .0600 \quad \text{D.W.} = .8842$$

Contract Construction Wages and Salaries (CCWS).

$$\text{CCWS} = - 2.09941 + 1.32439 \text{ MFGWS} \quad (\text{I.26})$$

$$(-6.161) \quad (27.344)$$

$$R^2 = .9829 \quad S_{y.x} = .2763 \quad \text{D.W.} = 1.2074$$

Mining Wages and Salaries (MINWS).

$$\text{MINWS} = - 2.6556 + 1.64867 \text{ MFGWS} \quad (\text{I.27})$$

$$(-7.274) \quad (31.773)$$

$$R^2 = .9873 \quad S_{y.x} = .2960 \quad \text{D.W.} = .6012$$

Wholesale and Retail Trade Wages and Salaries (WRTWS).

$$\text{WRTWS} = - .400793 + .758531 \text{ MFGWS} \quad (\text{I.28})$$

$$(-3.473) \quad (46.247)$$

$$R^2 = .9940 \quad S_{y.x} = .0936 \quad \text{D.W.} = .9688$$

Finance, Insurance, and Real Estate Wages and Salaries (FIREWS).

$$\text{FIREWS} = 1.60135 + .900751 \text{ MFGWS} - 2.52629 \text{ RELAUNR} \quad (\text{I.29})$$

$$(3.975) \quad (44.346) \quad (-4.997)$$

$$R^2 = .9940 \quad S_{y.x} = .1158 \quad \text{D.W.} = 2.3331$$

Transportation, Communication, and Public Utilities
Wages and Salaries (TCPUWS).

$$\text{TCPUWS} = - .163818 + 1.52054 \text{ MFGWS} - 3.06459 \text{ RELAUNR} \quad (\text{I.30})$$

$$(-.240) \quad (44.172) \quad (-3.577)$$

$$R^2 = .9939 \quad S_{y.x} = .1963 \quad \text{D.W.} = .9201$$

Service Wages and Salaries (SERWS).

$$\text{SERWS} = - .30614 + .752467 \text{ MFGWS} \quad (\text{I.31})$$

$$(-1.845) \quad (31.905)$$

$$R^2 = .9874 \quad S_{y.x} = .1345 \quad \text{D.W.} = .6695$$

State and Local Government Wages and Salaries (SLGWS).

$$\text{SLGWS} = 1.87034 + .601474 \text{ MFGWS} + .00175239 \text{ ST} - \quad (\text{I.32})$$

$$(3.688) \quad (4.141) \quad (1.302)$$

$$2.52268 \text{ RELAUNR}$$

$$(-5.564)$$

$$R^2 = .9944 \quad S_{y.x} = .1020 \quad \text{D.W.} = 2.5258$$

Federal Government Civilian Wages and Salaries (FGCWS).

$$\text{FGCWS} = .1753 + 1.70477 \text{ MFGWS} - 4.7485 \text{ RELAUNR} \quad (\text{I.33})$$

$$(.140) \quad (27.035) \quad (-3.026)$$

$$R^2 = .9839 \quad S_{y.x} = .3596 \quad \text{D.W.} = 1.3635$$

Farm Wages and Salaries (FWS).

$$\text{FWS} = - .368634 + 1.0183 \text{ HLE} - 17.4646 \text{ OKOLWS} - \quad (\text{I.34})$$

$$(-.140) \quad (15.900) \quad (-.136)$$

$$.345805 \text{ TIME}$$

$$(-1.469)$$

$$R^2 = .9961 \quad S_{y.x} = .8601 \quad \text{D.W.} = 2.9947$$

Relative Labor Cost (Z).

$$Z = \frac{\text{MFGWS}}{\text{USMFGW}} \quad (\text{I.35})$$

Real Wholesale and Retail Trade Wages and Salaries (WRTWSD).

$$\text{WRTWSD} = \frac{\text{WRTWS}}{\text{IPDC}} \quad (\text{I.36})$$

Real Finance, Insurance, and Real Estate Wages and Salaries (FIREWSD).

$$\text{FIREWSD} = \frac{\text{FIREWS}}{\text{IPDC}} \quad (\text{I.37})$$

Real Transportation, Communication, and Public Utilities Wages and Salaries (TCPUWSD).

$$\text{TCPUWSD} = \frac{\text{TCPUWS}}{\text{IPDC}} \quad (\text{I.38})$$

Real State and Local Government Wages and Salaries (SLGWSD).

$$\text{SLGWSD} = \frac{\text{SLGWS}}{\text{IPDC}} \quad (\text{I.39})$$

Personal Income

Other Labor Income (OLY).

$$\text{OLY} = - 182.934 + .14041 \text{ WSY} - 17.6183 \text{ TIME} \quad (\text{I.40})$$

$$(-25.278) \quad (23.143) \quad (-7.212)$$

$$R^2 = .9976 \quad S_{y.x} = 9.3302 \quad \text{D.W.} = 1.4219$$

Non-Farm Proprietors Income (NPY).

$$\text{NPY} = - 677.246 + 1.88326 \text{ NAE} + .602449 \text{ NPY1} - \quad (\text{I.41})$$

$$(-4.599) \quad (4.789) \quad (3.917)$$

$$67.4616 \text{ MAAA}$$

$$(-3.277)$$

$$R^2 = .9733 \quad S_{y.x} = 35.1658 \quad \text{D.W.} = 1.5287$$

Dividend, Rent, and Interest Income (DIRY).

$$\text{DIRY} = - 79.9731 + 9.44696 \text{ UDIRY} + .0416591 \text{ PYA} \quad (\text{I.42})$$

$$(-3.316) \quad (4.053) \quad (1.560)$$

$$R^2 = .9975 \quad S_{y.x} = 25.4948 \quad \text{D.W.} = 1.2268$$

Transfer Payments (TRY).

$$\text{TRY} = - 53.5332 + 11.891 \text{ USTRY} + 169.174 \text{ RELAUNR} \quad (\text{I.43})$$

$$(-.980) \quad (118.603) \quad (2.354)$$

$$R^2 = .9991 \quad S_{y.x} = 16.5361 \quad \text{D.W.} = 1.4503$$

Contributions to Social Insurance (CSSY).

$$\text{CSSY} = - 163.358 + .0713845 \text{ WSY} + 18.7873 \text{ OASI} \quad (\text{I.44})$$

$$(-6.113) \quad (10.708) \quad (1.518)$$

$$R^2 = .9937 \quad S_{y.x} = 12.5503 \quad \text{D.W.} = 1.6878$$

Wage and Salary Income (WSY).

$$\text{WSY} = (\text{MFCE} \times \text{MFCWS}) + (\text{CCE} \times \text{CCWS}) + (\text{MINE} \times \text{MINWS}) + (\text{WRTE} \quad (\text{I.45})$$

$$\times \text{WRTWS}) + (\text{FIREE} \times \text{FIREWS}) + (\text{TCPUE} \times \text{TCPUWS}) + (\text{SE} \times$$

$$\text{SERWS}) + (\text{SLGE} \times \text{SLGWS}) + (\text{FGCE} \times \text{FGCWS}) + \text{FMWS} +$$

$$\text{FWS} + \text{OIWS}$$

Personal Income (PYA).

$$\text{PYA} = \text{WSY} + \text{OLY} + \text{NPY} + \text{DIRY} + \text{TRY} + \text{RAY} + \text{FPI} - \text{CSSY} \quad (\text{I.46})$$

Real Personal Income (PYAD).

$$\text{PYAD} = \frac{\text{PYA}}{\text{IPDC}} \quad (\text{I.47})$$

Fiscal Year Personal Income (P).

$$P = \frac{PYA + PYA1}{2} \quad (I.48)$$

Disposable Personal Income (DPYA).

$$DPYA = PYA - FIT - TSIT \quad (I.49)$$

Real Disposable Personal Income (DPYAD).

$$DPYAD = \frac{DPYA}{IPDC} \quad (I.50)$$

Fiscal Year Disposable Personal Income (D).

$$D = \frac{DPYA + DPYA1}{2} \quad (I.51)$$

Ratio of Other Labor Income to Total Wage and Salary Income (OKOLWS).

$$OKOLWS = \frac{OLY}{WSY} \quad (I.52)$$

Tax Revenue

Sales and Use Taxes (SUT).

$$SUT = - 57.5202 + .00435013 D + 136.496 I \quad (I.53)$$

$$(-2.263) \quad (1.434) \quad (2.599)$$

$$R^2 = .9963 \quad S_{y.x} = 2.1213 \quad D.W. = 1.5006$$

Total State Income Taxes (TSIT).

$$TSIT = - 51.0643 + .0174303 P \quad (I.54)$$

$$(-6.788) \quad (18.802)$$

$$R^2 = .9646 \quad S_{y.x} = 10.1072 \quad D.W. = .9071$$

Tobacco Products Taxes (TPT).

$$\begin{aligned}
 \text{TPT} = & -138.642 + .0863843 \text{ CRI} - .0000771406 \text{ P} + & (\text{I.55}) \\
 & (-1.744) \quad (3.811) \quad (-.071) \\
 & .0607605 \text{ PO} \\
 & (1.700)
 \end{aligned}$$

$$R^2 = .9708 \quad S_{y.x} = 2.1301 \quad \text{D.W.} = 2.5627$$

Alcoholic Beverage Taxes (ABT).

$$\begin{aligned}
 \text{ABT} = & 10.2151 - .266858 \text{ ABWPI} + .18095 \text{ ABRI} + .00220109 \text{ P} & (\text{I.56}) \\
 & (2.063) \quad (-4.454) \quad (8.771) \quad (9.379)
 \end{aligned}$$

$$R^2 = .9911 \quad S_{y.x} = .7391 \quad \text{D.W.} = 1.5324$$

Motor Fuels Taxes (MFT).

$$\begin{aligned}
 \text{MFT} = & -75.7115 + .0063017 \text{ P} - .20333 \text{ GPI} + .0536284 \text{ PO} & (\text{I.57}) \\
 & (-.850) \quad (3.470) \quad (-3.848) \quad (1.376)
 \end{aligned}$$

$$R^2 = .9912 \quad S_{y.x} = 1.9377 \quad \text{D.W.} = 1.5911$$

Motor Vehicle Excise Taxes (MVET).

$$\begin{aligned}
 \text{MVET} = & 26.763 - .362978 \text{ MVWPI} + .00371134 \text{ P} & (\text{I.58}) \\
 & (4.566) \quad (-4.602) \quad (9.865)
 \end{aligned}$$

$$R^2 = .9759 \quad S_{y.x} = 1.0268 \quad \text{D.W.} = 1.8700$$

Total State Taxes (ST).

$$\text{ST} = \text{SUT} + \text{TSIT} + \text{TPT} + \text{ABT} + \text{MFT} + \text{MVET} + \text{OTHER} + \text{GPT} \quad (\text{I.59})$$

Federal Income Taxes (FIT) .

$$\text{FIT} = - 2552.73 + .0461517 \text{ PYA} + 1.1386 \text{ POP} \quad (\text{I.60})$$

$$(-1.709) \quad (2.076) \quad (1.729)$$

$$R^2 = .9752 \quad S_{y.x} = 45.2696 \quad \text{D.W.} = 1.0591$$

MiscellaneousDemand Deposits (DD) .

$$\text{DD} = 1305.09 + .257275 \text{ PYA} - 55.1068 \text{ MAAA} \quad (\text{I.61})$$

$$(14.647) \quad (14.940) \quad (-1.713)$$

$$R^2 = .9915 \quad S_{y.x} = 71.9829 \quad \text{D.W.} = 1.6237$$

Time Deposits (TD) .

$$\text{TD} = - 1332.24 + .601331 \text{ PYA} - 161.91 \text{ MAAA} \quad (\text{I.62})$$

$$(-10.007) \quad (23.371) \quad (-3.369)$$

$$R^2 = .9963 \quad S_{y.x} = 107.547 \quad \text{D.W.} = 2.3707$$

Total Loans (TL) .

$$\text{TL} = - 772.519 + .820887 \text{ PYA} - 157.429 \text{ FRDR} \quad (\text{I.63})$$

$$(-5.255) \quad (31.887) \quad (-2.773)$$

$$R^2 = .9965 \quad S_{y.x} = 151.963 \quad \text{D.W.} = 1.3620$$

Model Results

There are only a few basic differences between the estimation results of the OLS and INST models. The most significant difference between the two arises in the state and local government employment

equation. The INST model shows the wage and salary variable (SLGWSD) becoming large and positive while the lagged endogenous variable (SLGE1) turns negative. Both of these variables do seem to be highly insignificant though. Another difference appears in the farm wage and salary equation where the variable representing the ratio of other labor income to total wage and salary income (OKOLWS) becomes negative. This variable was found to be insignificant in both models. A final difference appears in the tobacco products equation. In this equation, the primary tax base variable, fiscal year personal income, become much smaller and insignificant.

All of the other equations show basically the same results in the two models. Coefficients have similar magnitudes and identical signs. Of the basic differences discussed above, the most crucial is in the state and local government employment equation. The results of the INST model form for this equation are totally contrary to the specified theory. With findings of this sort, it would not be surprising to find this equation in particular and the INST model in general to perform less satisfactory in simulation and forecasting. A large error in one equation can easily be passed on to many other equations and the whole model can become contaminated.

Ex Post Simulation

The INST model was simulated over the 1961-1975 time period in identical fashion to that of the OLS model. Table VIII presents the results of this simulation. The same four error statistics are presented for analysis of the simulation. Again, primary emphasis is placed upon the MAPE error statistic.

TABLE VIII

OKLAHOMA ECONOMETRIC STATE MODEL INST SIMULATION
ERROR STATISTICS, 1961-1975

Equation	MAE	MAPE	RMSE	Theil's U
QMFG	58.88	3.69%	80.18	.02193
QCC	19.74	3.65%	25.27	.02352
QWRT	64.48	3.68%	77.26	.02175
QFIRE	20.78	1.55%	27.41	.009789
QTCPU	19.83	1.94%	25.37	.01242
QSER	24.47	2.36%	29.34	.01346
QSLG	10.53	1.27%	13.74	.008268
QFG	11.82	1.08%	16.39	.007659
MFGE	3.953	3.34%	4.440	.01813
CCE	2.068	5.62%	2.535	.03325
MINE	1.259	3.26%	1.637	.02013
WRTE	2.212	1.37%	2.811	.008397
FIREE	.4890	1.33%	.5574	.007695
TCPUE	.7003	1.38%	.8909	.008784
SE	.7902	.78%	.9567	.004436
SLGE	13.39	11.05%	15.98	.06649
FGCE	1.058	1.97%	1.198	.01124
CLFA	11.47	1.12%	14.55	.007074
OKAUNR	.2540	6.93%	.3383	.04190
OKIUNR	.5167	18.85%	.5915	.08672
MFGWS	.0497	.73%	.05918	.004206
CCWS	.2027	2.93%	.2471	.01697
MINWS	.2401	2.74%	.3116	.01727
WRTWS	.0835	1.68%	.1104	.01116
FIREWS	.1100	1.84%	.1447	.01195
SERWS	.08699	1.93%	.1034	.01035
FGCSW	.3635	4.63%	.4318	.02476
SLGWS	.1274	2.72%	.1549	.01550
TCPUWS	.2153	2.76%	.2390	.01438
FWS	.5659	1.28%	.7528	.007333
OLY	18.28	5.52%	24.27	.03536
NPY	31.28	5.15%	36.25	.02728
DIRY	20.42	1.88%	24.90	.00993
TRY	16.84	2.13%	19.60	.009029
CSSY	13.58	6.08%	16.74	.02794
Z	.006484	.73%	.00746	.004188
SUT	1.576	1.72%	2.027	.01041
TSIT	8.037	12.51%	10.33	.05404
FIT	31.51	5.18%	39.05	.02635
TPT	1.262	4.74%	1.824	.02751
ABT	.4753	2.41%	.6593	.01654
MFT	1.141	1.32%	1.390	.008047
MVET	.8278	5.71%	1.018	.03063
DD	57.58	2.00%	69.02	.01119
TD	84.87	4.10%	112.5	.01941

TABLE VIII (Continued)

Equation	MAE	MAPE	RMSE	Theil's U
TL	98.03	2.07%	139.4	.01271
GSP	152.3	1.44%	183.8	.008453
NAE	12.53	1.61%	15.12	.01029
RELAUNR	.05007	6.93%	.06394	.04265
WRTWSD	.08978	1.68%	.1110	.01041
FIREWSD	.1220	1.84%	.1607	.01227
TCPUWSD	.2398	2.76%	.2674	.01515
SLGWSD	.1452	2.72%	.1765	.01642
OKOLWS	.002172	3.72%	.00259	.02197
WSY	108.4	2.02%	140.6	.01377
PYA	135.0	1.59%	169.3	.009969
DPYA	139.4	1.84%	171.8	.01122
PYAD	140.9	1.59%	167.3	.009576
DPYAD	146.4	1.84%	172.7	.01096
ST	9.753	2.53%	11.94	.01327
PCNAE	.03127	136.33%	.0361	.4550
D	68.71	1.12%	93.67	.006394
P	62.70	.94%	80.62	.004962

The output block of the INST model displays a range of MAPE errors from 1.08% for federal government output to 3.69% for the manufacturing sector. This block shows a slight improvement over the OLS model in terms of reducing simulation error. Manufacturing shows the largest improvement with an approximate 1% reduction taking place in the MAPE error.

The employment block presents a MAPE error range of .78% for service employment to 11.05% for state and local government employment. As was expected, the contrary findings of the INST estimation of the state and local government sector led to large simulation errors. In general, most of the employment equations show inferior MAPE error results to those of the OLS model. The civilian labor force equation displays a 1.12% MAPE error while both unemployment rate equations present much higher errors at 6.93% for the regular rate and 18.85% for the insured rate. The insured unemployment rate is alarmingly higher than that for the OLS model.

The wage and salary block shows a low MAPE error of .73% for manufacturing and a high of 4.63% for the federal government sector. On the whole, this block shows slightly higher MAPE errors than those present in the OLS model.

The personal income behavioral equations display a range of MAPE errors of 1.88% for dividend, interest, and rent income to 6.08% for contributions to social insurance. Generally, this sector also shows inferior results to those found in the other model. Other labor income jumps over 1.5% in added simulation error and contributions to social insurance show an increase of over 2%.

The tax block demonstrates slightly improved results over those of the OLS model. The range of MAPE errors extends from 1.32% for motor fuels taxes to 12.51% for total state income taxes. Total state income taxes shows an improvement of 1.5% in reduced MAPE error.

The miscellaneous block shows similar results to those of the OLS model. All three of the financial equations display MAPE errors of approximately 2% to 4%. Time deposits rise in simulation error while demand deposits and total loans show slight reduction.

Except for gross state product, the major aggregates of the INST model show greatly added simulation error to that found in the OLS model. GSP error falls slightly from 1.50% for the OLS model to 1.44% for the INST model. Non-agricultural employment more than doubles in simulation error up to 1.61% while wage and salary income almost triples in error from .72% to 2.02%. Personal income rises in the INST model to 1.59% MAPE error as compared to .91% in the OLS model. Total state taxes display no change in the two models.

In general, it would have to be concluded that the OLS model tended to present lower simulation error statistics. The major aggregates along with three of the five major blocks of equations disclosed higher simulation errors in the INST model as compared to the OLS model. In addition, Table IX presents a distribution of MAPE errors found in the INST model. Comparisons of different categories as represented by cumulative percentage in the last column also display the slight superiority of the OLS model over that of the INST model.

TABLE IX
INST MAPE ERROR DISTRIBUTION

Error Distribution	Number	Percent	Cumulative Percent
0 - .99%	4	6.3%	6.3%
1 - 1.99%	25	39.7%	46.0%
2 - 2.99%	13	20.6%	66.7%
3 - 3.99%	6	9.5%	76.2%
4 - 4.99%	3	4.8%	81.0%
5% and over	12	19.0%	100.0%

Ex Post Forecasting

An ex post forecast for the year 1976 was also made in the attempt to further test the INST model. Table X presents the results of this test. Included within the table are predicted values, actual values, actual errors, and the MAPE statistic for most of the endogenous variables for the year 1976. Only the predicted values will be presented for variables in which actual values are not yet available.

The output block shows similar results in the INST model to those of the OLS model. About half of the equations show improvement in forecasting error while the other half produce higher errors. The interval of MAPE errors for this block extends from .02% for transportation, communication, and public utilities to 6.20% for the federal government sector.

The employment block produces a range of MAPE errors from .36% for contract construction to 10.61% for mining. In general, this sector shows similar results to those of the OLS model. The civilian labor force denotes a MAPE error of .17% while the unemployment rate shows a 12.62% error. Both of these are similar to the OLS model.

The wage and salary block shows improvement in half of the equations and greater error in the other half as compared to the OLS model. The MAPE error range runs from .44% for farming and the federal government to 9.54% for contract construction. The construction error here is slightly lower than that in the OLS model.

Personal income also shows similar results in the two models. A low MAPE error of .11% is found for dividend, interest, and rent income while contributions to social insurance registers the high at 3.05%.

TABLE X
1976 INST SIMULATION ERRORS

Equation	Predicted	Actual	Error	MAPE
QMFG	2834.04	2722.80	111.24	4.09%
QCC	640.479	652.40	-11.921	1.83%
QWRT	2442.23	2468.00	-25.77	1.04%
QFIRE	1869.69	1907.60	-37.91	1.99%
QTCPU	1436.72	1436.50	.22	.02%
QSER	1435.18	1460.00	-24.82	1.70%
QSLG	1050.85	1046.00	4.85	.46%
QFG	943.105	1005.40	-62.295	6.20%
MFGE	157.364	156.10	1.264	.81%
CCE	45.9344	46.10	-.1656	.36%
MINE	39.6947	44.40	-4.7053	10.61%
WRTE	217.302	222.20	-4.898	2.20%
FIREE	48.4664	46.60	1.8664	4.01%
TCPUE	56.9176	57.30	-.3824	.67%
SE	146.701	151.40	-4.699	3.10%
SLGE	156.960	158.20	-1.24	.78%
FGCE	48.3847	48.80	-.4153	.85%
CLFA	1161.00	1159.00	2.00	.17%
OKAUNR	6.30684	5.60	.70684	12.62%
OKIUNR	2.82390			
MFGWS	10.8801	11.023	-.1429	1.30%
CCWS	12.3101	11.238	1.0721	9.54%
MINWS	15.2821	15.627	-.3449	2.21%
WRTWS	7.85210	7.497	.3551	4.74%
FIREWS	9.33238	9.415	-.08262	.88%
SERWS	7.88077	7.690	.19077	2.48%
FGCWS	14.8340	14.899	-.065	.44%
SLGWS	7.82248	7.562	.26048	3.44%
TCPUWS	13.8697	13.729	.1407	1.02%
FWS	102.667	102.219	.448	.44%
OLY	802.741	826.247	-23.506	2.84%
NPY	1060.82	1056.324	4.496	.43%
DIRY	2366.08	2363.552	2.528	.11%
TRY	2378.00	2393.674	-15.674	.65%
CSSY	617.850	599.56	18.29	3.05%
Z	.86777	.8792	-.01143	1.30%
SUT	178.684	181.865	-3.181	1.75%
TSIT	210.508	228.221	-17.713	7.76%
FIT	1324.26			
TPT	49.1742	50.391	-1.2168	2.41%
ABT	31.8167	32.591	-.7743	2.38%
MFT	118.327	117.256	1.071	.91%
MVET	29.3183	28.400	.9183	3.23%

TABLE X (Continued)

Equation	Predicted	Actual	Error	MAPE
DD	4896.73	4656.42	240.31	5.16%
TD	6783.43	6879.43	-96.00	1.40%
TL	11303.7	11214.01	89.69	.80%
GSP	14175.6	14222.00	-46.40	.33%
NAE	917.725	931.10	-13.375	1.44%
RELAUNR	.819071	.7270	.092071	12.66%
WRTWSD	5.89497	5.628	.26697	4.74%
FIREWSD	7.0063	7.068	-.0617	.87%
TCPUWSD	10.4127	10.307	.1057	1.03%
SLGWSD	5.87274	5.677	.19574	3.45%
OKOLWS	.0853611	.0881	-.0027389	3.11%
WSY	9404.09	9380.88	23.21	.25%
PYA	15766.0	15793.22	-27.22	.17%
DPYA	14231.2			
PYAD	11836.3	11856.77	-20.47	.17%
DPYAD	10684.1			
ST	841.321	862.214	-20.893	2.42%
PCNAE	.0341727	.0492	-.0150273	30.54%
D	13554.8			
P	15006.8			

The tax block displays MAPE errors of from .91% for motor fuels taxes to 7.76% for total state income taxes. This total state income tax figure is 3.5% higher than that in the OLS model.

The miscellaneous block also presents similar figures between the two models. Total loans registers the low MAPE error at .80% while demand deposits displays the high at 5.16%.

The major aggregates primarily show slight reductions in simulation errors in the INST model. Although total state taxes rise by almost 1% in error, gross state product, non-agricultural employment, and personal income all record minutely smaller errors. These latter variables show reduced errors of about .04%. Total wage and salary income does show a larger reduction of error. This wage aggregate displays a reduction in error of .73%.

In summary, the basic findings of the ex post forecasting test of the INST model are very similar to those of the OLS model. Very similar error statistics are found with both models. No clear-cut advantage can be placed with either model from these 1976 prediction tests.

Comparisons of OLS and INST Models

After identical estimation and testing of each model, a slight advantage or degree of superiority would have to be placed with the OLS model. Although ex post forecasting indicated no superiority with either model, the results of the ex post simulation indicated that the OLS model provided lower simulation errors. The two models used identical specification formats and the general estimation results showed similar coefficients in sign and magnitude for all but three equations. These comparisons between the OLS and INST models must be conditionalized

though. The INST model was produced by using the specification format that was found best for the OLS model. It might be possible that if a multitude of specification formats were estimated for the INST model, a form producing superior error statistics to those of the OLS model might be found. In addition, the OLS model was corrected for serial correlation while the INST model was not. Hence the findings of this study are conditional upon the structure in which the comparisons were made.

Drawing upon the conditional findings just discussed, the OLS model will be the model used later for providing a future forecast for Oklahoma. In addition, the OLS model will undergo a multiplier analysis in the next section.

Multiplier and Impact Elasticity Analysis

The OLS model was further tested by use of a multiplier and impact elasticity analysis. This test involves shocking the entire set of equations with a change in one or more of the exogenous variables. The actual process can be explained in the following manner.

Thus, one calculates a 'control solution' involving the analyst's 'best guess' as to the future course of the exogenous variables and predicts first the endogenous variables, $y_T^c, y_{T+1}^c, y_{T+2}^c, \dots, y_{T+K}^c$ and then a 'perturbed solution' in which one or more exogenous variables is shocked by the amount ∂ , $y_T^p, y_{T+1}^p, y_{T+2}^p, \dots, y_{T+K}^p$. One is able to calculate a 'dynamic multiplier' of the form $\frac{y_{T+K}^p - y_{T+K}^c}{\partial}$ (24, p. 150).

The multiplier analysis is dependent upon the units in which the variables are estimated. It is, therefore, sometimes wiser to calculate impact elasticities in which units make no difference. The impact

elasticity relates the induced percentage change of an endogenous variable from a certain percentage change in the exogenous variables. In this test of the OLS model, elasticities were estimated as opposed to pure multipliers.

A control solution for the OLS model was made for a period of six years beginning in 1976. This control solution is based upon assumptions made for the exogenous variables that will be explained in the next chapter.¹¹ This control solution provided values for the endogenous variables for six periods.

A perturbed solution was estimated by shocking the model one time only in the initial period with 1% increases in six national variables. These variables represented real gross domestic product for manufacturing; real gross domestic product for government; U. S. dividends, interest, and rent income; U. S. transfer payments; U. S. manufacturing wages; and the federal minimum wage. These shocks initially entered the output, income, and wage and salary blocks but then flowed through the entire model as shocks were passed along.

Table XI presents a time dimension of impact elasticities. The elasticities for all six periods are included in the table for a selected number of the variables.

The output block shows strong initial effects for all the sectors except state and local government. The federal government sector shows the largest initial increase of 3.6%. The sectors of this block are generally characterized by large initial effects followed by a large

¹¹The control solution mentioned in the multiplier analysis is actually the forecast made for Oklahoma for the period of 1976-1981. This forecast will be presented in the next chapter.

TABLE XI
A TIME DIMENSION OF IMPACT ELASTICITIES FOR THE OLS MODEL
DUE TO 1% CHANGES IN SIX NATIONAL VARIABLES^a
(PERCENTS)

Equation	1	2	3	4	5	6
QMFG	.30823	.23676	.18067	.13759	.1048	.08013
QCC	.62913	.17893	.11358	.08832	.07347	.06264
QWRT	.69795	.36389	.21842	.15261	.12036	.10169
QFIRE	.57932	.21546	.11795	.08718	.07282	.06444
QTCPU	1.16482	.17654	.13562	.11649	.10251	.09144
QSER	1.02482	.15733	.12219	.10545	.09331	.08404
QSLG	.1716	.16381	.15317	.14135	.13041	.11931
QFG	3.60109	.38278	.03811	.00382	.00042	0
MFGE	.11479	.13846	.12839	.10875	.08769	.07008
CCE	.4084	.25392	.15939	.11127	.08541	.06978
MINE	0	0	0	0	0	0
WRTE	.08154	.16398	.18148	.17531	.16102	.14533
FIREE	.42087	.21285	.11327	.08188	.06892	.06117
TCPUE	.12397	.10835	.0881	.0752	.06719	.06144
SE	.19584	.19047	.17958	.1675	.15538	.14299
SLGE	.02453	.09744	.13505	.14811	.1485	.14249
FGCE	2.23161	.99426	.34938	.11573	.03813	.01261
CLFA	0	-.0102	-.00083	.00655	.01122	.01492
OKAUNR	.40602	-.35234	-.36122	-.36392	-.3422	-.3357
MFGWS	.97181	0	0	.0007	.00065	.0006
CCWS	1.13104	-.00074	0	.0006	0	.00051
MINWS	1.18515	-.00057	0	.00046	.00042	.00078
WRTWS	1.02741	-.00465	0	0	0	.00082
FIREWS	.9504	.05565	.05219	.04855	.04176	.03646
SERWS	.99834	-.00035	-.0001	.00098	0	0
FGCWS	1.14071	.04295	.04049	.03709	.03164	.02814
SLGWS	.80908	.13458	.07176	.0651	.05626	.04929
TCPUWS	1.10181	.03265	.03043	.02866	.0243	.02142
FWS	.01758	.00277	.00173	.00162	.00076	.00071
OLY	2.01149	.34659	.2392	.19096	.16105	.1398
NPY	.31554	.42134	.43429	.41467	.37994	.3375
DIRY	1.05543	.02686	.02089	.018	.01637	.01515
TRY	.99083	-.01824	-.01756	-.01662	-.01399	-.01238
CSSY	1.44483	.2559	.18207	.14966	.12862	.11407
Z	-.02537	-.00034	-.00011	.00034	.00034	.00046
SUT	.23954	.25379	.06114	.05261	.0463	.04195
TSIT	.74598	.761	.18026	.14367	.12316	.10791
FIT	.85284	.13218	.1036	.09003	.08061	.07318
TPT	.16608	.18429	.04783	.04139	.03816	.03563
ABT	.61345	.65944	.16411	.13633	.12086	.10906
MFT	.56339	.61255	.15311	.12743	.11397	.10282
MVET	1.04849	1.0751	.2538	.20147	.17177	.14996

TABLE XI (Continued)

Equation	1	2	3	4	5	6
DD	.93381	.14426	.11408	.09976	.08935	.08057
TD	1.56671	.22677	.17014	.14248	.12308	.10778
TL	1.27587	.19043	.14619	.12483	.10912	.09689
GSP	.76291	.21763	.13561	.10449	.0865	.0735
NAE	.24144	.18974	.15419	.13463	.11965	.10735
CCWSD	1.13072	-.00052	0	0	0	.0009
WRTWSD	1.02734	-.00049	-.00015	.0003	.00044	.00043
FIREWSD	.95044	.05541	.05211	.04849	.04169	.03678
TCPUWSD	1.10142	.04311	.03039	.02846	.02456	.02147
SLGWSD	.80915	.13458	.07178	.06461	.05581	.04953
WSY	1.25279	.2256	.16311	.13568	.11814	.10529
PYA	1.1234	.171	.13266	.1138	.10023	.09019
DPYA	1.15471	.16556	.13451	.11535	.10175	.09106
PYAD	1.12438	.1715	.13248	.11402	.10032	.09008
DPYAD	1.15402	.16528	.13416	.11505	.10176	.0915
ST	.38972	.41038	.10058	.08164	.07197	.06363
D	.60734	.63119	.14909	.12439	.10796	.09655
P	.59119	.61957	.15068	.12247	.10709	.09526

^aThe six national variables were real gross domestic product for manufacturing; real gross domestic product for government; U. S. dividends, interest, and rent income; U. S. transfer payments; U. S. manufacturing wages; and the federal minimum wage.

drop in the second period (in terms of increases) and a tailing off from then on. Manufacturing and the state and local government sector, in contrast, show even drops in effects over the six periods.

The employment block produces smaller overall changes for the sectors than those found in the output block. This is probably due to the fact that employment is dependent upon output and, therefore, receives a lesser shock than that felt in the output block. The federal government sector does receive the largest initial increase of 2.23%. Most of the variables of this sector show even diminishing rates of increase over the six periods. State and local government reacts somewhat differently in that it shows a tiny initial increase of .02% but then increases up to about .15% in the fifth period. In addition, the mining sector showed no response to these changes. The civilian labor force shows no real reaction to this set of national changes. The unemployment rate initially rises but then shows negative changes for the rest of the period.

The wage and salary block is characterized by large initial increases in the first period and no significant changes from then on. All of the sectors except farming display increases of around 1% in the first period. This is followed by insignificant positive or negative movements over the rest of the time dimension. The farm sector shows no significant changes from these national shocks.

The income block shows a variety of results from this elasticity analysis. Other labor income and contributions to social insurance show large first period increases (2.01% and 1.44%, respectively). The second period increases are much smaller and continue to diminish at an even rate through the six periods. Dividend, interest, and rent

income begins with a large increase (1.055%) but the remaining period effects are negligible. Transfer payments start with a large increase (.99%) and follow this with negative although minute movements from then on. Non-farm proprietors income initially starts with a more modest increase (.32%) and continues to rise until the latter periods.

The tax sectors generally all follow a similar pattern. Most of the sources of taxes show large initial increases followed by a slightly higher increase in the second period. The third period brings much smaller increases and the remainder shows a tailing off in the increases. Federal income taxes display a large initial increase and much smaller increases throughout the remainder of the analysis.

The miscellaneous block shows increases of from .93% for demand deposits to 1.57% for time deposits in the first period followed by greatly diminishing increases in the second and remaining periods.

The major aggregates react in a similar fashion to what the equations in their respective blocks do. Gross state product, wage and salary income, and personal income show large initial increases (.76%, 1.25%, and 1.12%, respectively) followed by much smaller increases in the second period and thereafter. Non-agricultural employment initially rises .24% and then increases at smooth diminishing rates in the remaining periods. In contrast, total state taxes rise .38% in the first period, .41% in the second period, and then increase by much smaller amounts in the remaining periods.

The multiplier and impact elasticity analysis produced fairly satisfactory results. One percent increases in the six national variables generally led to large increases in the first period and much smaller increases thereafter for the endogenous variables. The wage and salary

block equations did deviate slightly from this behavior. Large first period increases were generally followed by no response in the remaining period for this block.

CHAPTER V

FORECAST FOR OKLAHOMA

Introduction

In this chapter, the OLS model which was estimated and tested in the previous chapter will be used to generate a sample forecast of expected economic activity for Oklahoma. The forecast will cover the six year period of 1976-1981. In order to generate a series of forecasts, it is necessary to have on hand a series of assumptions concerning the national and state exogenous variables.

National exogenous variable assumptions will be taken from a forecast provided by the Wharton Annual and Industry Model. This forecast was released on June 17, 1977. State exogenous variable assumptions will be provided from "guesstimates" on likely occurrences for these state variables.

Since forecasts are dependent upon many various assumptions, there is a possibility that large errors may enter the forecasting process. Norman Glickman (24) had noted:

Such forecasts are subject to at least three classes of errors, . . .

- a) There may be biased estimators or sampling errors, and
- b) The error term in the forecast period may not equal zero.

Error categories a) and b) are known as 'model errors.'
Finally,

- c) The values of the exogenous variables may be incorrectly forecasted; this is often done with the use of national models (p. 155).

There is a great chance that some sort of error may enter into the forecast. If national variables are wrongly predicted, then the regional model results can be expected to be faulty. Similarly, the state exogenous variables can be inaccurately predicted and this may lead to forecasting error. These potential errors are almost impossible to remove. The model builder and the forecaster along with any potential user of the forecasts must be aware of these possible errors. Any forecast is contingent upon the assumptions that are used in making that forecast.

Some potential errors within a forecast can be removed by the use of constant adjustments. These adjustments take place on the constant or intercept term of an equation. Constant adjustments are usually made by correcting the intercept term for the error in prediction made in the last year or two of simulation. Constant adjustments are usually deemed appropriate when a certain equation has repeatedly over or underestimated a particular variable. In addition, constant adjustments can be used to incorporate definitional changes in variables, account for changed behavior in variables, and to incorporate new information (11, pp. 69-70).

The OLS model, which was used to forecast in this chapter, did not use any constant adjustments. Hardly any of the equations show any repeated tendency to overestimate or underestimate the endogenous variables in the later periods. Also, it was not felt that any other conditions or information needed to be allowed for. As with any assumption, this may be fallacious. The forecasts provided with this model are those attained without tampering in any manner with the estimated model.

Forecast Assumptions

The national exogenous variables derive their assumptions for future behavior from a forecast provided by the Wharton Annual and Industry Model. This forecast was developed under the designation of "Control Solution Plus Carter Energy Program". It was distributed in June of 1977 and it involves adding the Administration's energy program to the forecasted baseline solution for the U. S.

The basic provisions of this Wharton forecast involve the economic stimulus package that was passed by Congress along with President Carter's suggestions on an appropriate energy package.

The various components of the complete Carter package can be grouped as follows:

1. The well-head equalization tax to start in 1978.
2. Industrial use taxes on oil and natural gas to start in 1979.
3. Utility use taxes on oil and natural gas to start in 1983.
4. Tax incentives to induce residential conservation investment such as insulation and solar energy to start in 1978.
5. Tax incentives to induce industry to convert to coal fired boilers and to install electricity cogeneration equipment to start in 1978.
6. Tax incentives to induce utilities to reduce oil and gas fired generation equipment--rebate to occur starting in 1983 against all qualified investment made after April 20, 1977.
7. Standby gasoline tax to start in 1979.
8. Gas-guzzler tax and rebate for efficient autos to start with 1979 model year (66, pp. 10-12).

In general, the predictions of real growth and inflation rates for the U. S. in the next ten years indicate that inclusion of the Carter energy program will lead to lower real growth and higher inflation.

Table XII contains the predicted behavior of the national exogenous variables which are used in the Oklahoma model. The values used for 1976 are the actual ones that occurred. These values are identical or very close to the ones supplied by the Wharton forecast for 1976 with one exception. This exception is real gross domestic product for U. S. manufacturing (RGDPMA). The value supplied by the Wharton forecast was somewhat lower than that which actually occurred.

TABLE XII
NATIONAL EXOGENOUS VARIABLES ASSUMPTIONS

Variable	Actual 1976	1977	1978	1979	1980	1981
RGDPMA	304.9	303.4	331.1	346.0	351.7	362.3
RGDPG	164.0	165.8	170.4	174.5	177.5	180.9
MAAA	8.43	8.06	8.34	8.62	8.71	8.43
USUR	7.7	7.0	5.9	5.6	5.9	6.1
USMFGW	12.538	13.691	14.937	16.274	17.640	19.035
FEDMW	2.30	2.30	2.65	2.90	3.10	3.35
UDIRY	189.401	213.112	238.626	262.447	281.836	302.335
USTRY	192.832	209.4	227.2	245.5	268.9	291.4
OASI	5.85	5.85	6.05	6.05	6.05	6.30
FRDR	5.5	5.3	5.8	6.1	5.8	5.9
IPDC	1.332	1.405	1.484	1.577	1.684	1.782
I	1.2985	1.3685	1.4445	1.5305	1.6305	1.733

The exogenous assumptions for the state variables were generally determined by noting the growth rate of the variables over the last few years. Table XIII presents the assumed behavior of these variables through 1981. Again, actual 1976 values were used to initiate the forecast.

TABLE XIII
STATE EXOGENOUS VARIABLES ASSUMPTIONS

Variable	Actual 1976	1977	1978	1979	1980	1981
POP	2766.0	2799.0	2832.0	2865.0	2898.0	2931.0
PO	2740.5	2782.5	2815.5	2848.5	2881.5	2914.5
TIME	19	20	21	22	23	24
HLE	109.1	115.17 ^a	122.76	130.85	139.47	148.66
CRI	260	260 ^a	260	260	260	260
ABRI	140.5	140.5 ^a	140.5	140.5	140.5	140.5
ABWPI	138.1	142.0 ^a	146.6	151.4	156.3	161.4
GPI	233.6	255.8 ^a	276.0	297.8	320.4	345.7
MVWPI	146.4	152.5	158.9	165.5	172.4	179.6
QMIN	791.3	791.3	791.3	791.3	791.3	791.3
QFARM	732.0	767.8	805.3	844.7	886.0	929.3
FMWS	347.197	359.06	371.33	384.02	397.15	410.73
OIWS	20.333	21.995	23.793	25.738	27.842	30.117
RAY	175.1	203.07	235.50	273.11	316.73	367.32
FPI	197.0	368.2 ^a	404.03	443.34	486.48	533.81
OTHER	72.176	78.741	85.903	93.717	102.242	111.542
GPT	151.316	170.494	192.103	216.451	243.884	274.794

^aActual values for 1977.

State population estimates were made by noting the actual 1976 values along with what was estimated by certain state agencies for 1980.

Equal increments of growth were then allotted for each year. Mining output was held at a constant amount over the next few years. The reason this was done involved the fact that recently mining output has been growing slightly but general behavior over the last nine years shows a negative growth rate. It was decided, therefore, to hold mining output at a constant figure. The cigarette rate index and the alcoholic beverage rate index were held at their constant amounts which were initiated in 1969 and 1972, respectively. All of the other state variables were generated by using growth rates of the respective variables over the last nine years.

Forecast Results

Tables XIV and XV present the results of the forecast from 1976-1981 for the key Oklahoma endogenous variables.¹ As the tables indicate, positive growth was evident for most of the variables. Federal government output and employment, mining employment, relative labor costs, and the unemployment rate were variables that showed downward movements. Each block of equations will be discussed, in general, as to their results.

The output block indicated positive growth for all of the estimated sectors except the federal government area. The transportation, communication, and public utilities sector registered the largest increase of 32.46% for this 1976-1981 period. The federal government sector, with a decline of 2.1% for this time period, reported the low. There were some ups and downs recorded during the forecast period for federal government

¹Appendix C presents the results of forecasting from 1976-1981 using the Goldberger suggestion of altering serially correlated equations.

TABLE XIV

OKLAHOMA FORECASTS OF SELECTED VARIABLES ASSUMING CONTROL SOLUTION PLUS CARTER ENERGY PROGRAM

Equation	1976	1977	1978	1979	1980	1981
QMFG	2822.50	2948.05	3093.92	3248.62	3396.71	3543.99
QCC	628.328	655.547	677.905	698.562	718.607	747.019
QWRT	2447.15	2561.18	2714.78	2863.43	2990.99	3127.12
QFIRE	1860.78	1953.95	2060.12	2156.44	2238.35	2327.45
QTCPU	1444.85	1546.35	1658.97	1751.13	1824.05	1913.80
QSER	1440.24	1525.43	1620.38	1697.39	1757.52	1832.41
QSLG	1054.72	1092.67	1135.92	1181.45	1226.86	1273.96
QFG	947.462	907.031	926.053	940.465	930.951	927.532
MFGE	157.672	163.941	169.783	175.617	181.311	186.924
CCE	44.0989	45.7213	47.3043	48.7960	50.2280	52.0206
MINE	39.6897	39.5128	39.3640	39.2388	39.1336	39.0450
WRTE	217.064	225.023	233.624	242.991	252.763	262.847
FIREE	47.8285	50.0344	52.4367	54.7137	56.7257	58.8428
TCPUE	56.7035	58.2370	59.4764	60.6365	61.7577	62.9852
SE	147.057	153.825	161.479	169.543	177.628	186.021
SLGE	158.937	165.213	171.785	178.919	186.526	194.389
FGCE	48.3507	46.5974	46.6247	47.0894	46.9408	46.7839
CLFA	1161.08	1175.63	1195.98	1221.34	1247.76	1272.98
OKAUNR	6.27553	5.53437	4.70071	4.46251	4.64640	4.68871
MFGWS	10.9177	11.7542	12.9358	14.0948	15.2282	16.4238
CCWS	12.3514	13.4531	15.0092	16.5356	18.0284	19.6029
MINWS	15.8291	17.3092	19.3999	21.4508	23.4564	25.5719
WRTWS	7.94131	8.58485	9.49388	10.3856	11.2576	12.1774
FIREWS	9.43177	10.2424	11.3043	12.3579	13.4078	14.5337
SERWS	7.84498	8.46278	9.33547	10.1915	11.0287	11.9117
FGCWS	15.0432	16.5278	18.5195	20.4884	22.4385	24.5189
SLGWS	7.92129	8.63378	9.53075	10.4450	11.3747	12.3749
TCPUWS	14.0042	15.3123	17.0868	18.8380	20.5677	22.4086
FWS	102.389	108.150	115.455	123.238	131.533	140.392
OLY	814.120	946.642	1132.94	1330.08	1533.60	1759.62

TABLE XIV (Continued)

Equation	1976	1977	1978	1979	1980	1981
NPY	1045.81	1170.06	1289.44	1408.34	1537.05	1694.77
DIRY	2373.42	2680.47	3015.03	3332.88	3603.88	3892.22
TRY	2381.83	2575.56	2789.24	3007.87	3285.79	3551.33
CSSY	633.638	716.669	830.982	950.138	1072.86	1209.68
Z	.870772	.858535	.866022	.866094	.863279	.862821
SUT	178.258	195.036	214.259	235.655	259.138	283.636
TSIT	218.504	250.851	289.567	331.994	375.906	422.548
FIT	1319.12	1460.13	1621.62	1788.23	1959.96	2145.39
TPT	49.6145	52.3629	55.1941	58.2127	61.3063	64.5378
ABT	31.8848	34.7262	38.1450	41.9557	45.9180	50.1543
MFT	118.389	127.663	139.765	153.021	166.702	180.886
MVET	29.4327	33.6246	38.9273	44.8693	50.9976	57.5476
DD	4917.45	5441.32	6013.03	6605.64	7229.66	7930.02
TD	6817.43	8047.72	9368.55	10738.0	12186.6	13824.4
TL	11262.9	12865.6	14637.5	16501.3	18510.1	20641.1
GSP	14169.3	14749.3	15484.7	16173.5	16761.3	17413.9
NAE	917.401	948.105	981.877	1017.54	1053.01	1089.86
RELAUNR	.815004	.790625	.796731	.796877	.787525	.768642
WRTWSD	5.96195	6.11022	6.39750	6.58567	6.68503	6.83356
FIREWSD	7.08091	7.28997	7.61746	7.83637	7.96187	8.15584
TCPUWSD	10.5137	10.8985	11.5140	11.9455	12.2136	12.5750
SLGWSD	5.94692	6.14504	6.42234	6.62335	6.75459	6.94437
WSY	9437.23	10549.6	12015.7	13560.5	15151.3	16905.3
PYA	15826.9	17776.9	20050.9	22405.9	24842.0	27494.7
DPYA	14289.2	16065.9	18139.7	20285.7	22506.1	24926.8
PYAD	11882.0	12652.6	13511.4	14207.9	14751.8	15429.1
DPYAD	10727.7	11434.8	12223.5	12863.5	13364.7	13988.1
ST	849.574	943.499	1053.86	1175.88	1306.09	1445.65
D	13583.8	15177.6	17102.8	19212.7	21395.9	23716.4
P	15037.3	16801.9	18913.9	21228.4	23623.9	26168.3

TABLE XV

FORECASTED INCREASES OF SELECTED VARIABLES FROM 1976-1981

Equation	Increase
QMFG	25.56%
QCC	18.89%
QWRT	27.79%
QFIRE	25.08%
QTCPU	32.46%
QSER	27.23%
QSLG	20.79%
QFG	-2.10%
MFGE	18.55%
CCE	17.96%
MINE	-1.62%
WRTE	21.09%
FIREE	23.03%
TCPUE	11.08%
SE	26.50%
SLGE	22.31%
FGCE	-3.24%
CLFA	9.64%
OKAUNR	-25.29%
MFGWS	50.43%
CCWS	58.71%
MINWS	61.55%
WRTWS	53.34%
FIREWS	54.09%
SERWS	51.84%
FGCWS	62.99%
SLCWS	56.22%
TCPUWS	60.01%
FWS	37.12%
OLY	116.14%
NPY	62.05%
DIRY	63.99%
TRY	49.10%
CSSY	90.91%
Z	-.91%
SUT	59.12%
TSIT	93.38%
FIT	62.64%
TPT	30.08%
ABT	57.30%
MFT	52.79%
MVET	95.52%

TABLE XV (Continued)

Equation	Increase
DD	61.26%
TD	102.78%
TL	83.27%
GSP	22.90%
NAE	18.80%
RELAUNR	-5.69%
WRTWSD	14.62%
FIREWSD	15.18%
TCPUWSD	19.61%
SLGWSD	16.77%
WSY	78.45%
PYA	73.72%
DPYA	74.45%
PYAD	29.85%
DPYAD	30.39%
ST	70.16%
D	74.59%
P	74.02%

output. The two largest sectors, manufacturing and wholesale and retail trade, listed increases of 25.56% and 27.79%, respectively. The manufacturing sector recorded the highest real output figure for 1981 at just over 3.5 billion. Most of the sectors of this block of equations reported average growth rates per year of between 4% and 6.5%. The key aggregate of this block, gross state product, measured an increase of 22.09% over this period. In average terms, this increase amounted to 4.6% per year.

The employment block reported increases for all sectors except mining and federal government civilian employment. The largest increase over this 1976-1981 period occurred in the service sector. This sector registered a 26.5% increase during this forecast period. This amounted to a per year average growth rate of 5.3%. The largest decline occurred in the federal government sector where -3.24% growth was recorded during this time. The largest employer at the beginning and end of the forecast period was the wholesale and retail trade sector. This area forecasted an employment of 262.8 thousand workers for 1981 and an increase of 21.09% during the forecast period. Non-agricultural employment, as a whole, listed an increase of 172 thousand workers during this 1976-1981 period. This aggregate increased 18.8% over the forecast period or at a per year growth rate of 3.76%. The civilian labor force increased 9.64% during this time and the unemployment rate dipped from 6.27% to 4.68%.

Nominal average annual wages reported large increases during the forecast period. Mining continued to be the highest paying sector in average annual wage terms. Wages increased from 15.8 thousand dollars to 25.5 thousand dollars in the mining sector. This amounts to an

increase of 61.55% over the 1976-1981 forecast period. The civilian federal government sector registered the highest increase of 62.99% during this term. In 1981, wages for this government sector were forecasted to be 24.5 thousand dollars. Most of the other sectors recorded increases of between 50% and 60%. The farm sector did list the lowest increase of 37.12% during the 1976-1981 period. This figure is in terms of total disbursements as opposed to average annual figures for the other sectors. Total wage and salary income reported an increase of 78.45% over this forecast period. Much of these wage increases are due to inflationary movements. The four real wage and salary variables estimated in the model indicate that only about 25% to 30% of the wage increases are real wage increases. The actual real wage increases amount to about 3% per year.

The income block showed large increases during this time period. Other labor income reported the largest increase of 116% over this 1976-1981 period. Transfer payments measured the lowest increase of 49%. Dividends, interest, and rent income still provides the largest source of income in 1981 other than total wages and salaries. Dividend, interest, and rent income rises to a forecasted amount of almost 3.9 billion dollars for 1981. Total wage and salary income measures 16.9 billion dollars in this year. The composite personal income total increased 73.32% during the forecast period. Personal income rose from a forecast of 15.8 billion dollars in 1976 to almost 27.5 billion dollars in 1981. This growth amounted to an average of 14.7% per year. Again, most of this increase is due to inflation.

The tax block also listed large forecasted increases from 1976 to 1981. The largest percentage increases occurred in motor vehicle

excise taxes and total state income taxes. These sectors reported increases of 95.52% and 93.38%, respectively, during the forecast period. Tobacco products taxes forecasted the lowest increase of 30.08% during this period. The other state tax sources all reported increases of between 50% and 60%. Federal income taxes from Oklahoma increased 62.24% for the forecasted time. Total state income taxes are forecasted to provide the largest individual state tax source in 1981 of 422.5 million dollars. Total state taxes are forecasted to rise from almost 850 million dollars in 1976 to 1.445 billion dollars in 1981. This amounts to an increase of 70.16% during this time.

The miscellaneous block lists increases of 61.26% for demand deposits, 102.78% for time deposits, and 83.27% for total loans. Demand deposits are forecasted to rise to 7.9 billion dollars in 1981 while time deposits will increase to 13.8 billion dollars. Total loans are predicted to move to just over 20.6 billion dollars at the end of the forecast period.

CHAPTER VI

ALTERNATIVE FORECASTING TECHNIQUE

Introduction

There are usually alternative methods or techniques for accomplishing any task. It is no different in the field of forecasting as there are several techniques available for projecting future economic activity. These methods range from the most naive extrapolative techniques to the complex usage of tools such as econometric models. Since the main intent of this manuscript is to develop an econometric model for Oklahoma which has the capability of forecasting future activity within the state, it might be of interest to see how this complex forecaster compares (in the ability to predict) with that of some other alternative forecasting tool.

A forecasting method which has shown great promise since its introduction in the early 1970's is the Box-Jenkins "time-series analysis" (9). This technique consists of a very sophisticated extrapolative usage and involves building a time-series model for each variable one desires to predict. Interactions with other variables are not used in building time-series models as was so in the econometric model. In a nutshell, ". . . the time-series model accounts for patterns in the past movements of a particular variable, and uses that information to

predict future movements of the variable" (49, p. 418). Time series models use only their own historical data for model construction.¹

It will be the purpose of this chapter to construct time-series models for four of the main aggregate variables of the econometric model.² These key endogenous variables are personal income (PYA), wage and salary income (WSY), non-agricultural employment (NAE), and gross state product (GSP). The constructed time-series models will then be compared with the econometric model as to the within-sample simulation errors produced by the two. This should give some indication as to how well the econometric model simulates as opposed to another forecasting technique. In addition, forecasts will be determined for the 1976-1981 time period using the time-series technique. Two sources of forecasts will then be available for pondering the future of Oklahoma.

A final item to be presented in this chapter will be a determination of a composite forecast. The composite forecast will put together the information provided by the econometric model and the time-series model for the four previously mentioned variables. This analysis will attempt to determine whether the time-series model can add some crucial information to the econometric model forecast which is missing. The composite forecasting approach is an attempt to provide a better joint forecast than can be provided by either the econometric model or the time-series model singularly.

¹Much of this chapter draws upon the time-series discussion of Charles R. Nelson (46) and Robert S. Pindyck and Daniel L. Rubinfeld (49).

²Time-series models usually require as a minimum at least 50 observations to acquire satisfactory fits. This study has only 18 observations available for the modeling process. Therefore, it would not be expected that the models attained here would be as good as those under more favorable conditions.

Box-Jenkins Time-Series Technique

The Box-Jenkins time-series modeling approach can be characterized as an extrapolative technique.³ It is, though, much more complex, sophisticated, and different than the usual deterministic extrapolation method.

The difference arises because time-series analysis presumes that the series to be forecasted has been generated by a stochastic (or random) process, with a structure than can be characterized and described. In other words, a time-series model provides a description of the random nature of the (stochastic) process that generated the sample of observations under study. The description is given not in terms of a cause-and-effect relationship (as would be the case in a regression model), but rather in terms of the way that randomness is embodied in the process (49, p. 421).

In essence, the time-series modeling approach assumes that each value of a given time series (y_1, y_2, \dots, y_T) is obtained randomly from a probability distribution (49, p. 431). It is then the intent of the technique to capture and identify the properties of this randomness.

An important question that arises when consideration of the usage of this technique is made involves whether the time series to be modeled is stationary or nonstationary. The Box-Jenkins method is much easier to apply when the underlying stochastic process that produced a given time series is stationary or invariant with respect to time. A nonstationary series is very difficult to model and can only be constructed using this Box-Jenkins approach when it can be transformed via differencing into a stationary series. Series in this latter category

³This section will only outline the basic points of the Box-Jenkins technique. A deeper discussion can be found in the authors' original book (9) or in Nelson (46) or Pindyck and Rubinfeld (49).

which can be successfully transformed are referred to as homogeneous nonstationary processes.

As was mentioned earlier, the Box-Jenkins time-series modeling approach does not involve the usage of other explanatory variables in the construction of a time-series model. A series is, instead, related ". . . to its own past values and to a weighted sum of current and lagged random disturbances" (49, p. 452). The specification is usually linear in form. The series may also be related to only its past values or only its current and lagged random disturbances. These possibilities provide the three major types of time-series models: moving average models, autoregressive models, and mixed autoregressive-moving average models. These three model forms along with a deviation off of the mixed model will be presented at this point.

Moving Average Models

A moving average model is generated by a weighted sum of current and past random disturbances. A process dating back q periods for the past disturbances may be denoted as $MA(q)$ and written as:

$$Y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (6.1)$$

where μ = mean of series,

ε_t = random disturbance in period t ,

ε_{t-i} = random disturbance dating back i periods, and

θ_i = moving average parameter for disturbance dating back i periods (49, p. 453).

The random disturbance terms are believed to be produced by a white noise process with mean of zero and variance σ^2 . Covariances between disturbances are assumed to be equal to zero.

Autoregressive Models

An autoregressive model is generated by a weighted average of past observations and a current random disturbance. A process involving lags of p periods may be denoted as $AR(p)$ and written as:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \delta + \varepsilon_t \quad (6.2)$$

where δ = constant trend term, and

ϕ_i = autoregressive parameter for i^{th} previous period (49, p. 458).

The random disturbance is also expected to be generated by a white noise process.

Mixed Autoregressive-Moving Average Models

A mixed model is assumed to be determined by p past observations as well as a current random disturbance and q lagged random disturbances. This process can be represented as $ARMA(p, q)$ and written as:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \delta + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (6.3)$$

where all variables and parameters are as explained before (49, p. 465).

White noise processes are also assumed for the random disturbances.

Homogeneous Nonstationary Processes

It was mentioned previously that certain nonstationary time series which could be transformed into stationary time series by the use of differencing can be modeled using the Box-Jenkins approach. Series which can achieve this result were labeled as homogenous nonstationary

processes. Differencing a series one or more times may sometimes be necessary in order to achieve this end. A homogenous nonstationary time series or process may be denoted as ARIMA(p, d, q) where d stands for the number of times the series was differenced in order to achieve stationarity. This model may be visualized as:

$$\phi(B)\Delta^d y_t = \delta + \theta(B)\varepsilon_t \quad (6.4)$$

where $\phi(B)$ and $\theta(B)$ are backshift operators while Δ^d indicates the use of differencing d times to achieve stationarity (49, p. 470). The autoregressive backshift operator $\phi(B)$ can be written as:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad (6.5)$$

while the moving average backshift operator $\theta(B)$ can be visualized as:

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad (6.6)$$

(49, p. 470). A backshift operator is just a simple notation form defined as:

$$By_t = y_{t-1} \quad (6.7)$$

and

$$B\varepsilon_t = \varepsilon_{t-1} \quad (6.8)$$

and similarly for all of the longer lags.

The integrated models of this sort may also be devised for only autoregressive effects or only moving average characteristics. An autoregressive model of order p that must be differenced d times in order to attain stationarity may be indicated as ARI(p, d). Similarly,

a moving average model with q lags which has to be differenced d times to become stationary can be denoted as $IMA(d, q)$. Many economic time series fall into this category of nonstationary series and must be differenced in order to achieve stationarity.

Model Construction

The process of constructing a time-series model and using it to make forecasts actually involves four separate phases of operation. The initial process in a time-series model development is the identification stage. This is followed by a parameter estimation stage and then a set of diagnostic checks to see if the model adequately describes the data. Finally, if a model passes all checks it may be used for forecasting.

The identification phase of a time-series model deals with determining what processes are responsible for the make-up of a given time series. Specifically, it must be determined whether a series is stationary or not, whether it can be made stationary by differencing if it is originally nonstationary, and whether moving average, autoregressive, or mixed influences explain the movement of the time series. Identification is arrived at through the usage of autocorrelation and partial autocorrelation functions.³ These functions for a given time series can be compared with those of several theoretical time series in order to deduce the appropriate structure for a model. Moving average, autoregressive, and mixed models all have different combinations of properties as pertain to their autocorrelation and partial autocorrelation functions.

³Nelson (46) may be consulted on exactly how this is accomplished.

After the basic properties of a time series are identified, estimation of the model parameters can take place. Estimation takes place by minimizing the sum of squares of the residuals of a conditional log likelihood function. The sum of squares of the residuals are nonlinear in form and must be linearized around initial guesses for the moving average and autoregressive parameters (θ 's and ϕ 's). This nonlinear estimation technique uses the first two terms of the Taylor series expansion to accomplish this feat (49, p. 483). Iteration takes place until the change in parameter values from one iteration to the next reaches some convergence criterion. Initial guesses for parameter starting values can be obtained from use of several theoretical derivations including the Yule-Walker equations.

Once a series has been identified and estimated, it is appropriate to conduct some sort of diagnostic check upon the model. The check will determine whether the model specification is correct. This check is accomplished by examining the residuals of the estimated model. The residuals, if the model has been correctly specified, would be expected to possess the same approximate properties that the random error terms of the actual processes contain. These properties are a normal distribution and independence among the error terms. Therefore, the sample autocorrelation function of the residuals for a correctly specified model should contain autocorrelations of approximately zero for displacements greater than or equal to one lag (49, p. 490). The sample autocorrelations for the residuals can be extracted and a chi-square test statistic suggested by Box and Pierce (8) can be calculated and compared with the chi-square tables to determine whether the model is appropriately specified. The suggested statistic, R , is composed of the first K

residual autocorrelations $\hat{r}_1, \hat{r}_2, \dots, \hat{r}_k$ and can be computed in the following manner:

$$R = T \sum_{k=1}^K \hat{r}_k^2 \quad (6.9)$$

where T = number of observations in the time series (49, p. 491). The R statistic is assumed to have a chi-square distribution with $(K-p-q)$ degrees of freedom.

After a model has been identified, estimated, and checked, it can be used to make future predictions for the time series. The best or optimum forecast at a given time origin is the conditional expectation of the future observation (9). This optimum provides the minimum mean square forecast error. Computer packages are available to compute these forecasts as well as provide the information for identification, estimation, and diagnostic checking.

Time-Series Models for Four Variables

Four key endogenous variables were modeled using the Box-Jenkins technique. These variables were gross state product, non-agricultural employment, wage and salary income, and personal income. After identification and estimation had taken place, these time-series models were used to forecast within the sample from 1961-1975. MAPE error statistics were computed so that the time-series models could be compared with the econometric model's results for these variables. For all of the variables except personal income, more than one model form was found to be acceptable. The model form with the lowest MAPE error was later chosen as the one to use for making forecasts for that particular

variable. Each variable and its estimated time-series model(s) will now be presented. Statistical significance at the 5% level for the estimated parameter will be denoted by an asterisk (*). In addition, residual error sums of squares (R.S.S.), residual mean square errors (R.M.S.), residual standard errors (R.S.E.), and the chi-square statistic for the diagnostic check with the appropriate degrees of freedom $\chi^2_{()}$ will be presented with each equation.

Gross State Product

The identification phase of the model construction indicated that gross state product might be estimated in one of the following four forms: ARIMA (1,0,0), ARIMA (2,1,0), ARIMA (2,1,1), or ARIMA (1,1,1).⁴ Of these four possibilities, only two were found to possess acceptable time-series properties and pass the diagnostic checks.

ARIMA (2,1,0).

$$(1 - .16539 B - .15234 B^2) \Delta y_t = 277.25^* + \epsilon_t \quad (6.10)$$

$$\text{R.S.S.} = 803370.0 \quad \text{R.M.S.} = 66948.0 \quad \text{R.S.E.} = 258.74$$

$$\chi^2_{(5)} = 3.9635$$

ARIMA (2,1,1).

$$(1 - .25888 B + .05468 B^2) \Delta y_t = 314.72 + (1 + .15346 B)\epsilon_t \quad (6.11)$$

$$\text{R.S.S.} = 814670.0 \quad \text{R.M.S.} = 74061.0 \quad \text{R.S.E.} = 272.14$$

$$\chi^2_{(4)} = 3.8256$$

⁴Models will be presented in the full ARIMA form. It is true that models such as ARIMA (1,0,0) could be denoted as AR(1) while the ARIMA (2,1,0) model could be written as ARI(2,1). The full form will be used for consistency purposes.

The within sample forecasting results for 1961-1975 produced MAPE errors of 3.40% for the ARIMA (2,1,0) model and 4.24% for the ARIMA (2,1,1) model. Therefore, it was decided that the ARIMA (2,1,0) model would be used for generating forecasts for gross state product.

Non-Agricultural Employment

Five different forms of rht non-agricultural employment variable were estimated using this time-series format. The model forms were ARIMA (1,0,0), ARIMA (2,0,0), ARIMA (0,1,1), ARIMA (0,1,2), and ARIMA (1,1,1). Of this list, three were found to possess acceptable results and pass diagnostic checks.

ARIMA (0,1,1).

$$\Delta y_t = 18.52^* + (1 + .78672 B^*)\epsilon_t \quad (6.12)$$

$$R.S.S. = 1068.4 \quad R.M.S. = 71.226 \quad R.S.E. = 8.4396$$

$$\chi^2_{(6)} = 1.4112$$

ARIMA (0,1,2).

$$\Delta y_t = 18.321^* + (1 + .70075 B + .1857 B^2)\epsilon_t \quad (6.13)$$

$$R.S.S. = 1045.2 \quad R.M.S. = 74.655 \quad R.S.E. = 8.6403$$

$$\chi^2_{(5)} = 1.3697$$

ARIMA (1,1,1).

$$(1 - .14046 B) \Delta y_t = 16.63^* + (1 + .71549 B^*)\epsilon_t \quad (6.14)$$

$$R.S.S. = 1043.3 \quad R.M.S. = 80.254 \quad R.S.E. = 8.9585$$

$$\chi^2_{(5)} = 1.6543$$

All three of the time-series models produced fairly low forecasting errors for the period 1961-1975. The ARIMA (0,1,1) model registered a MAPE error of 2.19%. The ARIMA (0,1,2) model recorded a MAPE error of 2.34%. The lowest MAPE error of 1.84% was found in the ARIMA (1,1,1) model. The latter model, ARIMA (1,1,1), will be used for forecasting activity in the non-agricultural employment area.

Wage and Salary Income

The identification phase of the time-series modeling attempt located three potential modeling forms. These forms were ARIMA (1,0,0), ARIMA (1,1,0), and ARIMA (2,1,0). The ARIMA (1,0,0) modeling attempt failed in the estimation phase while the other two forms were successful and passed diagnostic checks. In addition, the ARIMA (2,1,0) model without a trend term was also found to pass all checks.

ARIMA (1,1,0).

$$(1 - .97652 B^*) \Delta y_t = 45.855 + \epsilon_t \quad (6.15)$$

$$R.S.S. = 135210.0 \quad R.M.S. = 9657.6 \quad R.S.E. = 98.273$$

$$\chi^2_{(6)} = 4.1299$$

ARIMA (2,1,0).

$$(1 - .97974 B^* + .11017 B^2) \Delta y_t = 56.85 + \epsilon_t \quad (6.16)$$

$$R.S.S. = 125450.0 \quad R.M.S. = 10454.0 \quad R.S.E. = 102.25$$

$$\chi^2_{(5)} = 5.3566$$

ARIMA (2,1,0).

$$(1 - 1.1017 B^* + .23922 B^2) \Delta y_t = \epsilon_t \quad (6.17)$$

$$R.S.S. = 137350.0 \quad R.M.S. = 10565.0 \quad R.S.E. = 102.79$$

$$\chi^2_{(5)} = 5.5692$$

The ARIMA (1,1,0) model showed the best within-sample forecasting of the three models. This model registered a MAPE error of 5.49% over the 1961-1975 period. The ARIMA (2,1,0) models (with and without a trend term) recorded MAPE errors of 11.30% and 15.94%. None of the models replicated economic activity to any great degree of success. The ARIMA (1,1,0) model will be used for forecasting.

Personal Income

Three forms of the personal income equation were identified as possible modeling molds. These forms were ARIMA (1,0,0), ARIMA (1,1,0), and ARIMA (1,1,1). Unfortunately, only the ARIMA (1,1,0) model surfaced as an acceptable alternative. The other forms either failed diagnostic checks or violated certain estimation properties.

ARIMA (1,1,0).

$$(1 - .93341 B^*) \Delta y_t = 109.21 + \epsilon_t \quad (6.18)$$

$$R.S.S. = 652060.0 \quad R.M.S. = 46575.0 \quad R.S.E. = 215.81$$

$$\chi^2_{(6)} = 4.206$$

The within-sample forecast error for this model was 19.20%. This error is very high and indicates that forecasts for personal income using this technique will probably not be very reliable. The limited number

of observations for the modeling process probably hindered this estimation process greatly. With no other alternative, this model will have to be used for forecasting.

Forecasts Using the Time-Series Models

The ARIMA time-series model forecasts for the 1976-1981 time period are presented in Table XVI for the four estimated variables. Table XVII denotes the percentage increases that occur for the four variables over this 1976-1981 forecast period. The ARIMA forecasts show that gross state product will increase by almost two billion dollars during this time. Gross state product is forecasted to grow by 13.85% over this period. Non-agricultural employment is forecasted to grow 10.60% during this time up to just below one million workers. Wage and salary income shows the largest percentage increase of the four variables. This key aggregate will grow 46.19% during the forecast period and rise to over 13 billion dollars in 1981. Personal income is forecasted to grow 44.61% from 1976 to 1981. Personal income will measure over 22.5 billion dollars in 1981 according to the forecast.

Composite Forecasting

Sometimes a better predictor or forecaster can be developed by combining information from more than one forecasting technique. Such a process of combining information enters into the realm of composite prediction. In this section, one such linear composite prediction is attempted by combining predictions from both the econometric model and the time-series model. This act of combining econometric model forecasts with those of time-series models has been done before in other studies (14) (30) (46).

TABLE XVI

ARIMA MODEL FORECASTS FOR FOUR VARIABLES, 1976-1981

Variable	1976	1977	1978	1979	1980	1981
GSP	14148.46	14540.49	14932.43	15324.32	15716.20	16108.07
NAE	895.843	913.658	932.789	952.106	971.449	990.794
WSY	9151.473	9943.727	10763.22	11609.32	12481.40	13378.84
PYA	15582.13	16936.94	18310.73	19702.22	21110.25	22533.70

TABLE XVII

PERCENTAGE INCREASES FROM 1976-1981 OF THE FOUR FORECASTED
VARIABLES FROM THE ARIMA PROCESS

Variable	Increase
GSP	13.85%
NAE	10.60%
WSY	46.19%
PYA	44.61%

A linear composite prediction can be viewed in the following form:

$$A_t = \beta_1(\text{ECON})_t + \beta_2(\text{ARIMA})_t + \varepsilon_t \quad (6.19)$$

where A_t = actual value of a variable for period t ,
 $(\text{ECON})_t$ = value predicted from the econometric model for period t ,
 $(\text{ARIMA})_t$ = value predicted from the time-series model for period t ,
 β_1 and β_2 = fixed coefficients, and
 ε_t = composite prediction error (46, p. 212).

The coefficients β_1 and β_2 are estimated by using ordinary least squares and this process provides the minimum mean-square-error linear composite prediction for the sample period (46, p. 212). If the econometric model and time-series model predictions are individually unbiased, then (6.19) may be rewritten simply as:

$$A_t = \beta(\text{ECON})_t + (1 - \beta)(\text{ARIMA})_t + \varepsilon_t \quad (6.20)$$

where β = a fixed coefficient and all other variables are explained as before (46, p. 212).

If the time-series model (since it draws on a smaller subset of information) cannot contribute any new information to the composite forecast, then the estimates of β_1 and β_2 would approximate unity and zero, respectively. If the times-series model and the econometric model both individually contribute some information to the composite forecast, then the coefficients will both be positive.

The four variables which were modeled in the time-series format in the previous section will be used to generate composite forecasts in this section. Ordinary least squares will be used to estimate β_1 and β_2

from within the 1961-1975 sample period. The composite models presented in (6.19) and (6.20) along with these same equations incorporating intercept terms will be estimated. Serial correlation problems will be corrected by use of the Cochrane-Orcutt technique (13). After estimation, these equations will be used to make composite forecasts for the period 1976-1981. Since actual 1976 values are now available for these four variables, MAPE errors will be estimated to compare the forecasting abilities of the econometric model, the time-series model, and the composite model.

Estimation of Composite Models

Table XVIII presents the results of the composite model estimation procedure for gross state product, non-agricultural employment, wage and salary income, and personal income. The table discloses the estimated values of β_1 and β_2 along with the standard error of the model, the Durbin-Watson statistic, and the weight allotted to the ARIMA prediction when equation (6.20) is estimated.⁵ This latter constrained estimate provides a means of comparison between the unconstrained and the constrained regression. If the individual predictions are essentially unbiased, then the constrained regression estimates should differ only minutely from those of the unconstrained regression.

All four of the variables produced large, significant coefficients for the econometric model predictions and small, insignificant, and sometimes negative coefficients for the time-series predictions. In addition, all four variables required a serial correlation correction.

⁵ Equations incorporating intercept terms were found to produce insignificant coefficients for this variable and are, therefore, not reported.

The gross state product variable registered a value of .967 for the econometric model prediction coefficient and a figure of .032 for the time-series prediction coefficient. The constrained regression produced a similar figure of .030 for the time-series contribution. Non-agricultural employment generated a value of over one for the econometric model contribution and listed a negative contribution for the time-series prediction. The constrained contribution was also negative. Personal income produced very similar results to those of non-agricultural employment. The econometric model coefficient was greater than one, the time-series coefficient was negative, and the constrained time-series coefficient was also negative. The wage and salary variable showed the largest contribution of all for the time-series predictions. The variable recorded a value of .107 for this prediction contribution. The coefficient was insignificant, though. The econometric model coefficient was a significant .884 while the constrained time-series coefficient was positive but much smaller than the unconstrained.

TABLE XVIII
COMPOSITE ESTIMATION RESULTS

Variable	$\hat{\beta}_1$	$\hat{\beta}_2$	Standard Error of Model	D.W.	Weights Given to ARIMA Under $\hat{\beta}_1 + \hat{\beta}_2 = 1$
GSP**	.967*	.032	133.4271	.9180	.030
NAE**	1.046*	-.048	5.7842	1.6201	-.019
WSY**	.884*	.107	32.4552	1.7361	.016
PYA**	1.092	-.080	73.2044	1.8198	-.024

* Denotes significance at 5% level.

** Denotes serial correlation correction.

The results seem to indicate that only in the wage and salary income equation does the time-series prediction add any significant information. The coefficient on the constrained regression for the time-series predictor does put a damper on any strong interpretation of the above result. The other three variables, gross state product, non-agricultural employment, and personal income, seemed to be unaffected by the time-series predictor. No new information was included by this time-series technique for these variables.

Forecasting with Composite Models

Even though the estimation results showed very little contribution from the time-series predictions, the composite models were used to generate forecasts for the period 1976-1981. Table XIX presents the results of these forecasts. The same basic upward movements that have characterized all forecasts of these four variables are visible here. As would be expected, the magnitudes of the composite forecasts closely resemble those of the econometric model. Table XX relates the percentage increases that are forecasted for the four variables during the period 1976-1981. Since the composite models do place great value upon the econometric model forecasts, the percentage increases of the composite models more closely resemble those of the econometric model.

A comparison can be made for 1976 of the forecasting accuracy of the three forecasters presented in this manuscript. Table XXI presents the MAPE errors covering the four variables and the three forecasting techniques for 1976. Results show that the econometric model produced the lowest forecasting error for gross state product, non-agricultural employment, and personal income. In three of the four cases, the

TABLE XIX

COMPOSITE MODEL FORECASTS FOR FOUR VARIABLES, 1976-1981

Variable	1976	1977	1978	1979	1980	1981
GSP	14154.56	14727.87	15451.54	16130.15	16711.10	17354.70
NAE	916.60	947.86	982.27	1018.65	1054.82	1092.44
WSY	9353.54	10389.83	11773.54	13229.59	14729.26	16375.82
PYA	16036.40	18057.42	20430.73	22891.07	25438.64	28221.52

TABLE XX

PERCENTAGE INCREASES FROM 1976-1981 OF THE FOUR FORECASTED
VARIABLES FROM THE COMPOSITE PROCESS

Variable	Increase
GSP	22.61%
NAE	19.18%
WSY	75.08%
PYA	75.98%

TABLE XXI

COMPARISON OF 1976 MAPE ERRORS ACROSS FORECASTING TECHNIQUES

Variable	Econometric Model	ARIMA Model	Composite Model
GSP	.37%*	.52%	.47%
NAE	1.47%*	3.79%	1.56%
WSY	.98%	2.35%	.29%*
PYA	.21%*	1.34%	1.54%

* Denotes lowest error for variable.

time-series model registered the highest MAPE error of the three techniques.

The 1976 forecasting error results would seem to indicate that the econometric model is superior to the time-series and composite modeling techniques. These results have to be accepted with a grain of salt since only one year was available for comparing forecasting accuracy. In addition, the limited number of observations available for the time-series modeling attempt surely hindered the results of this technique and, therefore, those of the composite model.

CHAPTER VII

SUMMARY AND CONCLUSIONS

Summary and Conclusions

The major intent of this study was to develop a functioning econometric model for the State of Oklahoma. By functioning, it was meant that the model should be capable of providing future forecasts for the state and be able to simulate policy impact analysis. This goal was accomplished in this manuscript as the model was constructed, tested, and used to provide a sample forecast through 1981. In addition, the construction phase allowed for the inclusion of several policy variables so that policy impacts can be studied within this framework.

The early chapters of this study built a foundation for the construction of an Oklahoma model by discussing the trends and history of regional econometric model building. The pioneering works of Lawrence Klein and Norman Glickman were used as a guideline for comparing the past and current tendencies in this area of research. Several state and regional models were discussed as to their structures and simulation results. From these past attempts, it was decided to pattern the Oklahoma model after those of Philadelphia, Mississippi, Georgia, and Tennessee. These models have tried to maximize their theoretical foundations within the constraints of data accessibilities. Attempts at building Keynesian subsystems without regard to certain data needs does

not seem to be very popular or successful on the regional level at this time.

The third chapter of this manuscript laid down the theoretical specification forms for the blocks of equations in the model. The output and income blocks were modeled as dependent upon their particular national and state influences. Also, pertinent variables for individual equations were allowed for. The employment block of equations was specified according to a theoretical derivation of a CES production function assuming profit maximization. Employment in a sector was modeled as dependent upon sectoral output and real wages. The wage and salary block of equations was determined according to a theory of "intra-area wage rollout" in which the manufacturing sector plays a prominent role. Since tax revenues are determined by a tax base and a tax rate, the tax revenue block was constructed around these variables. Price indices and population effects were also used to describe activity in this block. The miscellaneous block of equations, which covered a small financial sector, was constructed around local activity variables and alternative financial interest rates.

The actual estimation and testing of the Oklahoma model was accomplished in the fourth chapter. Two methods of estimation were used to construct the model. One method involved a single equation estimator (OLS) while the other took account of simultaneous effects (INST). These models were both tested in simulation runs over the sample period of 1961-1975. The conditional results showed that, in general, the OLS model generated lower error statistics than the INST model. All comparisons were pointed out to be conditional upon the fact that the INST model was estimated from the specification form found best for the OLS

model. These two models were also tested in an out of sample simulation test for 1976. The simulation errors for this year were very similar for both models. A decision was made at this point that the OLS model had shown slightly superior replication abilities and would, therefore, be used for forecasting purposes. The OLS model was further submitted to a slight alteration test to see if better 1976 forecasts could be generated if equations showing serial correlation were allowed for. Serial correlation corrections had been made previously in the OLS estimation process. Although 1976 forecasts did not show better results from employing this alternative, a within sample simulation and a forecast were also generated using this suggestion. A final multiplier test was made on the OLS model to disclose the effects of national changes on the state variables.

Comparisons with other state models showed that the Oklahoma model had generated very similar error statistics. Approximately 86% of the endogenous variables in the OLS model produced MAPE error statistics below 5%. Tennessee and Mississippi showed slightly better results while Philadelphia and Georgia were somewhat worse. Individual variable comparisons also showed the Oklahoma model to be quite comparable to the other regional models.

A forecast was made for the Oklahoma economy incorporating national assumptions involving President Carter's controversial energy bill. This forecast ran from 1976 through the year 1981. In general, this forecast painted a rosy picture for the State of Oklahoma. Positive growth was projected for all but a few sectors. Mining employment, federal government civilian employment, and federal government output were the only sectors showing downward movements in this forecast.

Gross state product was projected to increase over 22% during this time while non-agricultural employment was forecasted to increase almost 19%. Wage and salary income and personal income were found to increase 70% during this time but much of this growth was attributed to inflationary movements. Total state taxes were forecasted to climb to almost 1.5 billion dollars in 1981 or to show an increase of over 70% during this forecast period. The unemployment rate was projected to show a small downward movement from the start in 1976.

After estimation, testing, and forecasting of the Oklahoma model had been successfully completed, it was felt that it might be interesting to compare the forecasting abilities of the econometric model with those of some alternative forecasting procedure. The Box-Jenkins time-series technique seemed to be the perfect tool to use for the comparison. Four variables from the econometric model were selected to be modeled in the time-series mold. These variables were gross state product, non-agricultural employment, wage and salary income, and personal income. With the exception of personal income, more than one form of each equation was selected for testing purposes.

Replication tests over the period of 1961-1975 were made on the several forms of the four variables. MAPE error statistics were used to find the one form for each equation which best reproduced economic activity. The MAPE errors for these four variables were generally much higher in this time-series format than they were when modeled by the econometric model. A set of forecasts from 1976-1981 were then generated using these time-series models for the four selected variables. The forecasts did tend to differ from those in the econometric model. A situation which severely hindered the usage of the time-series

technique in this instance was the limited number of observations available for model development.

A composite forecasting model was then developed which combined the forecasts of the econometric model with those of the time-series model. The intent behind this procedure was to combine information from both forecasting tools and, therefore, produce a better joint forecast than what can be achieved individually. The composite models of all four variables tended to show that the time-series model had very little to add to the forecasting abilities of the econometric model. Only the wage and salary income equation showed any minor influence or information being added by inclusion of the time-series forecaster. Not surprisingly, the composite model forecasts tended to resemble those of the econometric model.

Since actual 1976 data was available, the econometric model, the time-series model, and the composite model were compared as to their forecasting accuracy for the four selected variables. In three of the four cases, the econometric model showed the lowest forecasting error. Only in the wage and salary income equation did the composite model record the lowest error. The time-series model generated the highest forecast error in most of the cases.

Favorable results were achieved in the estimation, testing, and forecasting procedures of this study. An econometric model for the State of Oklahoma was developed and found to compare quite favorably to models of other regions in terms of simulation errors. Comparison with an alternative forecasting technique also tended to solidify the impression that the Oklahoma econometric model is a good, sound replicator of economic activity. As a first step in long-range econometric

model building for the state, this project would appear to have provided a good "jumping off" place.

Suggestions for Future Research

There are always certain limitations to any study as well as suggestions for future improvements. To deal with the limitations first, it appears that there are three areas within the present model that are in need of some additional work. These areas involve the unemployment rate equations, the exogenous mining output sector, and the total state income tax equation.

The unemployment rate variable used in this study contained a series revision in the early 1970's. An adjustment was made in the original series with help from a few overlapping years to try and put all observations on a common framework. This ad hoc attempt for developing a consistent unemployment rate series was probably not completely successful. Simulation errors for this variable did tend to be fairly high. It would be most helpful if the agencies responsible for establishing these series would place all past observations on a similar footing or foundation when they decide to measure a variable in a different manner. A more consistent unemployment rate series would probably improve the forecasting ability of this variable.

The mining output sector of the econometric model was made exogenous due to the inability to adequately model this sector. Several variables were tried but none proved successful. The availability of certain rare explanatory variables that are only pertinent to this sector could provide for a successful modeling of this sector. Aid from certain "specialists" in this area could prove quite helpful.

The total state income tax equation is another area that could benefit from future research and study. Many specifications for this variable were tried in hopes of reducing the large simulation error it recorded. Several dummy variables as well as nonlinear specifications were attempted with little success. Log-linear specifications for the entire tax sector might prove fruitful to this equation. Such a specification would, though, be more costly to the simulation process.

If some aid could be provided to the above mentioned sectors then the Oklahoma econometric model would be a superior forecasting model. These limitations, although bothersome, do not really seem to present any serious problems.

There are five major areas of extension in which the econometric model could be added to. These areas are:

1. A greater breakdown in the manufacturing sector,
2. A more detailed development of the employment block from the individual sectors to the civilian labor force determination,
3. A more extensive tax revenue sector,
4. The creation of a government expenditures block, and
5. The addition of more equations in the miscellaneous block.

The manufacturing sector could be disaggregated into several of the key two-digit SIC classifications in order to more adequately explain movements in this large area. This greater breakdown could be provided for in the output, employment, and wage and salary blocks. Manufacturing is the largest source of real output in the state and an industry in this position is usually composed of varied activities within its ranks (38). An explanation of this varied activity would most assuredly provide a better forecasting model and one of greater interest to many industrialists.

The employment block as it now stands is built up from the individual sectors to the development of a non-agricultural employment aggregate. From this point there is a larger jump to the civilian labor force level with no concrete substance inbetween. It might be more beneficial and complete if a consistent farm employment variable could be included and a total employment variable created. From this point, various alternatives would be available for making the final link from total employment up to the civilian labor force.

The tax revenue block is always an important concern to policy-makers on the state level. Any improvements that can be made to this block are always viewed in a very favorable light. Explanations of certain revenue sources that were made exogenous in this model as well as the estimation of other general operating revenues via behavioral equations would add strength to this tax revenue block. As was mentioned earlier, additional work on the total state income tax equation would also help this block greatly. A successful division of state income taxes into a corporate and individual partition would also be very advantageous. Error for this equation could be drastically reduced if such a split were made possible in a consistent fashion.

In line with extensions in the tax revenue block would be the inclusion of a state government expenditures section. Since Oklahoma is required by law to balance its state budget, it might be very fruitful to have a revenue and expenditures section for making policy decisions. School expenditures might be a very important area of investigation in this light.

The catchall miscellaneous block could be vastly expanded to the benefit of the Oklahoma model. Additional work in the financial sector

might be of great interest to the state banking association. More detail could be provided for in both the deposits and loans categories. A demographic sector providing for explanations of population tendencies and migration movements within the state might be of great interest to many students of regional economics as well as urban planners. A manufacturing investment equation could add to the predictive abilities of the manufacturing output sector. A miscellaneous block would benefit from any activity area that could be successfully modeled.

An econometric model can be a great aid to regional forecasting and policymaking. To provide this service, it must be constructed upon a sound, theoretical foundation while still being flexible enough to adapt to the data limitations present on the state and regional level. To be able to successfully forecast into the future, an econometric model must be continuously updated and changed in order to incorporate the new information which becomes available to it. As such, it must itself be involved in a dynamic process if it is to function successfully. It is hoped that this has been a huge first step in an ongoing process towards providing such a service for the State of Oklahoma.

BIBLIOGRAPHY

- (1) Adams, F. Gerard, Carl G. Brooking, and Norman J. Glickman. "On the Specification and Simulation of a Regional Econometric Model: A Model of Mississippi." Review of Economics and Statistics, Vol. 57, No. 3, 1975, pp. 286-298.
- (2) Adams, F. Gerard and Michael K. Evans. Econometric Forecasting with Wharton-EFU Model: 5-Years. Philadelphia: University of Pennsylvania, Department of Economics, Discussion Paper No. 78, 1968.
- (3) Anderson, Robert J., Jr. "A Note on Economic Base Studies and Regional Econometric Forecasting Models." Journal of Regional Science, Vol. 10, No. 3, 1970, pp. 325-333.
- (4) Ankrust, O. "Prim I: A Model of the Price and Income Distribution Mechanism of an Open Economy." Review of Income and Wealth, Vol. 16, No. 1, 1970, pp. 51-78.
- (5) Ball, R. J. "The Significance of Simultaneous Methods of Parameter Estimation in Econometric Models." Applied Statistics, Vol. 11, No. 3, 1963, pp. 14-25.
- (6) Bell, Frederick W. "An Econometric Forecasting Model for a Region." Journal of Regional Science, Vol. 7, No. 2, 1967, pp. 109-127.
- (7) Borts, George H. "The Equalization of Returns and Regional Economic Growth." American Economic Review, Vol. 50, No. 3, 1960, pp. 319-347.
- (8) Box, G. E. P. and D. A. Pierce. "Distribution of Residual Autocorrelations in Autoregressive-Integrated-Moving Average Time Series Models." Journal of the American Statistical Association, Vol. 65, No. 332, 1970, pp. 1509-1526.
- (9) Box, G. E. P. and Gwilym M. Jenkins. Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day, Inc., 1970.
- (10) Burton, R. and J. Dyckman. A Quarterly Econometric Forecasting Model for the State of California. Berkeley: Center for Planning and Development, Research Institute of Urban and Regional Development, University of California, 1965.
- (11) Chang, Hui S. Tennessee Econometric Model: Phase I. Knoxville: Center for Business and Economic Research, College of Business Administration, University of Tennessee, March, 1976.

- (12) Chu, Shih-fan and Benjamin M. Wofford. An Econometric Model for the State of Nevada. Reno: Bureau of Business and Economic Research, University of Nevada, Research Report No. 12, December, 1974.
- (13) Cochrane, D. and G. H. Orcutt. "Application of Least Squares Regressions to Relationships Containing Autocorrelated Error Terms." Journal of the American Statistical Association, Vol. 44, No. 245, 1949, pp. 32-61.
- (14) Cooper, J. Phillip and Charles R. Nelson. "The Ex Ante Prediction Performance of the St. Louis and FRB-MIT-PENN Econometric Models and Some Results on Composite Predictors." Journal of Money, Credit, and Banking, Vol. 7, No. 1, 1975, pp. 1-32.
- (15) Cooper, Ronald L. "The Predictive Performance of Quarterly Econometric Models of the United States." Econometric Models of Cyclical Behavior, Ed. Bert G. Hickman. Studies in Income and Wealth, Vol. 2, No. 36, 1972, pp. 813-926.
- (16) Crow, Robert Thomas. "A Nationally-Linked Regional Econometric Model." Journal of Regional Science, Vol. 13, No. 2, 1973, pp. 187-204.
- (17) Duesenberry, J. S., G. Fromm, L. R. Klein, and E. Kuh. The Brookings Quarterly Econometric Model of the U. S. Chicago: Rand-McNally Co., 1965.
- (18) Durbin, J. "Testing for Serial Correlation in Least-Squares Regression When Some of the Regressors are Lagged Dependent Variables." Econometrica, Vol. 38, No. 3, 1970. pp. 410-421.
- (19) Engle, Robert F. "Issues in the Specification of an Econometric Model of Metropolitan Growth." Journal of Urban Economics, Vol. 1, No. 2, 1974, pp. 250-267.
- (20) Evans, Michael K. Macroeconomic Activity. New York: Harper and Row, 1969.
- (21) Evans, Michael K., Yoel Haitovsky, and George I. Treyz. "An Analysis of the Forecasting Properties of U. S. Econometric Models." Econometric Models of Cyclical Behavior. Ed. Bert G. Hickman. Studies in Income and Wealth, Vol. 2, No. 36, 1972, pp. 949-1139.
- (22) Evans, Michael K. and Lawrence R. Klein. The Wharton Economic Forecasting Model. Philadelphia: University of Pennsylvania, Economics Research Unit, 1968.
- (23) Fromm, G. and L. R. Klein. "The Brookings-Social Science Research Council Quarterly Econometric Model of the U. S.: Model Properties." American Economic Review, Vol. 55, No. 2, 1965, pp. 348-369.

- (24) Glickman, Norman J. Econometric Analysis of Regional Systems: Explorations in Model Building and Policy Analysis. New York: Academic Press, Inc., 1977.
- (25) _____. "Son of the Specification of Regional Econometric Models." Papers of the Regional Science Association, Vol. 32, 1974, pp. 155-177.
- (26) _____. "An Econometric Forecasting Model for the Philadelphia Region." Journal of Regional Science, Vol. 11, No. 1, 1971, pp. 15-32. ✓
- (27) _____. "An Econometric Model of the Philadelphia Region." (Ph.D. dissertation, University of Pennsylvania, 1969.)
- (28) Goldberger, Arthur S. "Best Linear Unbiased Prediction in the Generalized Linear Regression Model." Journal of the American Statistical Association, Vol. 57, No. 298, 1962, pp. 369-375.
- (29) _____. Impact Multipliers and Dynamic Properties of the Klein-Goldberger Model. Amsterdam: North-Holland, 1959.
- (30) Granger, C. W. J. and Paul Newbold. Forecasting Economic Time Series. New York: Academic Press, 1977.
- (31) Grimes, A. Ray, Jr. "A Satellite Econometric Model for Georgia." Mimeo. Athens, Georgia: School of Business, University of Georgia, 1973.
- (32) Henderson, James M. and Richard E. Quandt. Microeconomic Theory. 2nd ed. New York: McGraw-Hill, 1971.
- (33) Johnston, J. Econometric Methods. 2nd ed. New York: McGraw-Hill, 1972.
- (34) Kendrick, J. W. and C. M. Jaycox. "The Concept and Estimation of Gross State Product." Southern Economic Journal, Vol. 32, No. 2, 1965, pp. 153-168.
- (35) Klein, Lawrence R. "The Specification of Regional Econometric Models." Papers of the Regional Science Association, Vol. 23, 1969, pp. 105-115.
- (36) Klein, Lawrence R. and Norman J. Glickman. "Econometric Model-Building at the Regional Level." Regional Science and Urban Economics, Vol. 7, 1977, pp. 3-23.
- (37) Klein, Lawrence R. and Arthur S. Goldberger. An Econometric Model of the United States, 1929-1952. Amsterdam: North-Holland, 1955. ✓

- (38) Lage, Gerald M. and Timothy C. Ireland. "Oklahoma Gross State Product: 1950-1975." Review of Regional Economics and Business, Vol. 3, No. 1, 1978, pp. 3-12.
- (39) Landon, John H., William R. Latham, III, and Kenneth A. Lewis. "An Improved Quarterly Forecasting Model of a Small Region." Mimeo. Newark, Delaware: University of Delaware, May, 1976.
- (40) Lee, C. F. and J. B. Legler. "An Application of the Box-Jenkins ARIMA Model to Regional Econometric Forecasting: Georgia as a Case Study." Paper presented at the annual meeting of the American Statistical Association, Atlanta, Georgia, August, 1975.
- (41) L'Esperance, W. L., G. Nestel, and D. Fromm. "Gross State Product and an Econometric Model of a State." Journal of the American Statistical Association, Vol. 64, No. 327, 1969, pp. 787-807.
- ✓ (42) Liew, Chong K. and Dae K. Kahng. "The Oklahoma Econometric Model I." Oklahoma Business Bulletin, Vol. 39, No. 7, 1971, pp. 7-24.
- (43) Morgan, James N. and Edward H. Robb. "A State and Regional Econometric Model for Missouri." Business and Government Outlook, Vol. 1, No. 3, 1976.
- (44) Morrison, Donald F. Multivariate Statistical Analysis. 2nd ed. New York: McGraw-Hill, 1976.
- (45) Naylor, T. H. The Design of Computer Simulation Experiments. Durham, North Carolina: Duke University Press, 1969.
- (46) Nelson, Charles R. Applied Time Series Analysis for Managerial Forecasting. San Francisco: Holden-Day, Inc., 1973.
- (47) _____. "Predictive Performance of FRB-MIT-PENN Models of the U. S." American Economic Review, Vol. 62, No. 5, 1972, pp. 902-917.
- (48) Oklahoma Employment Security Commission. Oklahoma Employment Statistics. Oklahoma State Employment Service, Research and Planning Division, Oklahoma City, Oklahoma, 1976.
- (49) Pindyck, Robert S. and Daniel L. Rubinfeld. Econometric Models and Economic Forecasts. New York: McGraw-Hill, 1976.
- (50) Preston, R. S. The Wharton Annual and Industry Forecasting Model. Philadelphia: Department of Economics, University of Pennsylvania, Studies in Quantitative Economics, No. 7, 1972.
- (51) Ratajczak, Donald. "Data Limitations and Alternative Methodology in Estimating Regional Econometric Models." Review of Regional Studies, Vol. 4, No. 2, 1974, pp. 51-64.

- (52) Richardson Harry W. Regional Economics. New York: Praeger Publishers, Inc., 1969.
- (53) Richter, Charles. "Some Limitations of Regional Econometric Models." The Annals of Regional Science, Vol. 6, No. 1, 1972, pp. 28-34.
- ✓ (54) Roberts, B. F. and G. Wittels. An Econometric Model of California. Mimeo. Berkeley: School of Business, University of California, 1972.
- (55) Sandmeyer, Robert L., Dale Wasson, and Rudy I. Greer. A Study of Oklahoma State Taxes. Stillwater, Oklahoma: Office of Business and Economic Research, College of Business Administration, Oklahoma State University, 1978.
- (56) Schwartz, Charles F. and Robert E. Graham, Jr. Personal Income by States Since 1929. Washington: U. S. Government Printing Office, 1956.
- (57) Smith, Paul E. "An Econometric Growth Model of the United States." American Economic Review, Vol. 53, No. 4, 1963, pp. 682-693.
- (58) Suits, Daniel B. "Forecasting and Analysis with an Econometric Model." American Economic Review, Vol. 52, No. 1, 1962, pp. 104-132.
- (59) Theil, H. Economic Forecasts and Policy. Amsterdam: North-Holland, 1961.
- (60) Theil, H. and J. C. G. Boot. "The Final Form of Econometric Equation Systems." Readings in Economic Statistics and Econometrics. Ed. Arnold Zellner. Boston: Little, Brown, and Company, 1968, pp. 611-630.
- (61) Thompson, Wilbur R. A Preface to Urban Economics. Baltimore: The Johns Hopkins Press, 1975.
- (62) United States Department of Commerce. Annual Survey of Manufacturers. Social and Economic Statistics Administration, Bureau of the Census, Washington, D. C.
- (63) _____. Survey of Current Business. Bureau of Economic Analysis, Washington, D. C.
- (64) United States Federal Deposit Insurance Corporation. Assets and Liabilities: Commercial and Mutual Savings Banks. Washington, D. C.
- (65) Wachter, M. "A Labor Supply Model for Secondary Workers." Review of Economics and Statistics, Vol. 54, No. 2, 1972, pp. 141-151.

- (66) Wharton Econometric Forecasting Associates, Inc. Wharton E. F. A.
Annual and Industry Forecasting Model Review of the Post-
Meeting Control Solution. Philadelphia, Pennsylvania,
June 17, 1977.

APPENDIX

APPENDIX A

DATA SOURCES

- BEA - Intermediate Tables, United States Department of Commerce, Bureau of Economic Analysis, Regional Economics Information System, Washington, D. C.
- DRAY - Data from unpublished econometric model of Oklahoma farm sector by Daryll Ray, Department of Agricultural Economics, Oklahoma State University.
- EE - United States Department of Labor. Employment and Earnings, United States, 1909-1975. Bureau of Labor Statistics, Washington, D. C., Bulletin 1312-10.
- ERP - Economic Report of the President. United States Government, Transmitted to the Congress, January, 1977, Washington, D. C.
- FDIC - Assets and Liabilities: Commercial and Mutual Savings Banks. See reference (24).
- IRS - United States Department of the Treasury. Individual Income Tax Returns. Internal Revenue Service, Washington, D. C.
- KLOS - Brinker, Paul A. and Joseph J. Klos. Poverty, Manpower, and Social Security. Austin, Texas: Austin Press, 1976.
- LI - Gross state product study denoted in reference (40).
- OEPS - Employment Security Program Statistics. Oklahoma Employment Security Commission, Oklahoma State Employment Service, October, 1976.
- OES - Oklahoma Employment Statistics. See reference (50).
- OTC - Data from Dale Wasson of the Oklahoma Tax Commission.
- OTS - Oklahoma tax study denoted in reference (57).
- SCB - Survey of Current Business. See reference (61).
- SSME - 1974 Social Security and Medicare Explained. Commerce Clearing House, New York, 1974.

Variable	Source
ABRI	OTC
ABT	OTS, Table III-2
ABWPI	OTC
CCE	OES, Table B-1.04
CCE1	OES, Table B-1.04
CCWS	BEA, Printouts; OES, Table B-1.04
CLFA	OES, Table A-1.01
CLFA1	OES, Table A-1.01
CRI	OTC
CSSY	BEA, Printouts
D	BEA, Printouts
DD	FDIC, Table 7
DIRY	BEA, Printouts
DPYA	BEA, Printouts
DPYAD	BEA, Printouts
FEDMW	KLOS, p. 499
FGCE	OES, Table B-1.44
FGCE1	OES, Table B-1.44
FGCWS	BEA, Printouts; OES, Table B-1.44
FIREE	OES, Table B-1.37
FIREWS	BEA, Printouts; OES, Table B-1.37
FIREWSD	BEA, Printouts; OES, Table B-1.37; SCB, July Table 7.11
FIT	IRS
FMWS	BEA, Printouts
FPI	DRAY
FRDR	ERP, Table B-65
FWS	BEA, Printouts
GPI	OTC
GPT	OTS, Table III-2
GSP	LI, Table 2
HLE	DRAY
I	SCB, July Table 7.11

Variable	Source
IPDC	SCB, July Table 7.11
MAAA	ERP, Table B-65
MFGE	OES, Table B-1.05
MFGE1	OES, Table B-1.05
MFGWS	BEA, Printouts; OES, Table B-1.05
MFT	OTS, Table III-2
MINE	OES, Table B-1.02
MINE1	OES, Table B-1.02
MINWS	BEA, Printouts; OES, Table B-1.02
MVET	OTS, Table III-2
MVWPI	OTC
NAE	OES, Table B-1.01
NPY	BEA, Printouts
NPY1	BEA, Printouts
OASI	SSME, p. 14
OIWS	BEA, Printouts
OKAUNR	OES, Table A-1.03
OKAUNR1	OES, Table A-1.03
OKIUNR	OEPS, Table A-9
OKOLWS	SCB
OLY	BEA, Printouts
OTHER	OTS, Table III-2
P	BEA, Printouts
PCNAE	OES, Table B-1.01
PO	SCB, August Table 49
POP	SCB, August Table 49
PYA	BEA, Printouts
PYAD	BEA, Printouts; SCB, July Table 7.11
QCC	LI, Table 2
QFARM	LI, Table 2
QFG	LI, Table 2
QFG1	LI, Table 2
QFIRE	LI, Table 2

Variable	Source
QFIRE1	LI, Table 2
QMFG	LI, Table 2
QMFG1	LI, Table 2
QMIN	LI, Table 2
QSER	LI, Table 2
QSLG	LI, Table 2
QSLG1	LI, Table 2
QTCPU	LI, Table 2
QWRT	LI, Table 2
QWRT1	LI, Table 2
RAY	BEA, Printouts
RELAUNR	OES, Table A-1.03; ERP, Table B-29
RGDPG	SCB, July Table 6.2
RGDPMA	SCB, July Table 6.2
SE	OES, Table B-1.39
SE1	OES, Table B-1.39
SERWS	BEA, Printouts; OES, Table B-1.39
SLGE	OES, Table B-1.45
SLGE1	OES, Table B-1.45
SLGWS	BEA, Printouts; OES, Table B-1.45
SLGWSD	BEA, Printouts; OES, Table B-1.45; SCB, July Table 7.11
ST	OTS, Table III-2

Variable	Source
SUT	OTS, Table III-2
TCPUE	OES, Table B-1.26
TCPUE1	OES, Table B-1.26
TCPUWS	BEA, Printouts; OES, Table B-1.26
TCPUWSD	BEA, Printouts; OES, Table B-1.26; SCB, July Table 7.11
TD	FDIC, Table 7
TIME	ERP
TL	FDIC, Table 7
TPT	OTS, Table III-2
TRY	BEA, Printouts
TSIT	OTS, Table III-2
UDIRY	BEA, Printouts
USMFGW	BEA, Printouts; EE, p. 38
USTRY	BEA, Printouts
USUR	ERP, Table B-29
WRTE	OES, Table B-1.30
WRTE1	OES, Table B-1.30
WRTWS	BEA, Printouts; OES, Table B-1.30
WRTWSD	BEA, Printouts; OES, Table B-1.30; SCB, July Table 7.11
WSY	BEA, Printouts
Z	BEA, Printouts; OES, Table B-1.05; EE, p. 38

APPENDIX B

GOLDBERGER ALTERNATIVE FOR 1961-1975 OLS SIMULATION

Equation	MAE	MAPE	RMSE	Theil's U
QMFG	67.93	4.16%	87.74	.02387
QCC	20.20	3.69%	26.33	.02442
QWRT	54.77	3.15%	66.15	.01858
QFIRE	23.42	1.71%	29.05	.01037
QTCPU	14.11	1.37%	17.52	.00855
QSER	16.96	1.72%	23.21	.01062
QSLG	12.81	1.54%	16.45	.00993
QFG	13.68	1.26%	17.56	.00821
MFGE	2.786	2.36%	3.347	.01362
CCE	1.666	4.45%	2.081	.02717
MINE	1.229	3.18%	1.616	.01987
WRTE	1.856	1.15%	2.488	.00741
FIREE	.4956	1.31%	.6356	.00876
TCPUE	.7240	1.43%	.9257	.00912
SE	.8401	.82%	1.052	.00488
SLGE	2.575	2.18%	3.156	.01317
FGCE	.9849	1.80%	1.158	.01088
CLFA	13.66	1.35%	16.42	.00797
OKAUNR	.2359	6.25%	.3115	.03877
OKINUR	.1762	5.67%	.2323	.03477
MFGWS	.0413	.63%	.0456	.00324
CCWS	.1620	2.33%	.2037	.01398
MINWS	.2395	3.10%	.2869	.01588
WRTWS	.0702	1.37%	.0892	.00900
FIREWS	.1044	1.70%	.1377	.01138
SERWS	.0749	1.61%	.0877	.00879
FGCWS	.2509	3.22%	.3082	.01766
SLGWS	.1062	2.19%	.1221	.01222
TCPUWS	.1278	1.65%	.1563	.00939
FWS	.5582	1.26%	.7190	.00700
OLY	7.438	2.83%	9.349	.01355
NPY	37.54	5.96%	44.08	.03278
DIRY	16.79	1.52%	21.84	.00870
TRY	15.69	1.93%	18.83	.00868
CSSY	8.440	3.47%	11.67	.01941
Z	.00562	.63%	.00644	.00362
SUT	1.632	1.74%	2.052	.01053
TSIT	7.875	11.62%	9.309	.04841
FIT	25.61	4.27%	31.81	.02148
TPT	1.412	4.95%	1.934	.02915
ABT	.4413	2.24%	.6027	.01511
MFT	1.293	1.53%	1.594	.00922
MVET	.7590	5.18%	.9539	.02863

Equation	MAE	MAPE	RMSE	Theil's U
DD	60.26	2.10%	68.21	.01104
TD	67.31	3.10%	98.46	.01692
TL	114.7	2.04%	153.1	.01391
GSP	144.2	1.34%	185.9	.00854
NAE	5.119	.72%	6.023	.00409
RELAUNR	.0453	6.27%	.0571	.03817
WRTWSD	.0740	1.37%	.0891	.00835
FIREWSD	.1126	1.70%	.1429	.01092
TCPUWSD	.1443	1.65%	.1780	.01008
SLGWSD	.1180	2.19%	.1351	.01258
OKOLWS	.0014	2.50%	.00172	.01460
WSY	27.12	.61%	35.91	.00351
PYA	59.90	.80%	80.40	.00472
DPYA	72.58	1.08%	93.23	.00607
PYAD	67.15	.80%	89.21	.00510
DPYAD	81.44	1.08%	104.1	.00659
ST	8.751	2.17%	10.68	.01186
PCNAE	.00781	41.71%	.00913	.14780
D	54.74	.78%	74.16	.00506
P	50.44	.66%	66.12	.00407

APPENDIX C

1976-1981 FORECAST WITH
GOLDBERGER ALTERNATIVE

Equation	1976	1977	1978	1979	1980	1981
QMFG	2814.37	2937.31	3083.20	3239.00	3388.51	3537.19
QCC	628.158	654.799	676.976	697.650	717.785	746.295
QWRT	2448.44	2559.13	2710.61	2858.58	2986.26	3122.77
QFIRE	1861.59	1952.53	2057.66	2153.83	2235.95	2325.31
QTCPU	1445.59	1543.51	1655.43	1747.83	1821.18	1911.28
QSER	1440.89	1522.94	1617.28	1694.50	1755.00	1830.20
QSLG	1054.86	1092.51	1135.44	1180.70	1225.94	1272.92
QFG	947.462	907.031	926.053	940.465	930.951	927.532
MFGE	154.468	161.216	167.964	174.471	180.581	186.440
CCE	43.1904	44.8385	46.6396	48.3367	49.9205	51.8154
MINE	39.6897	39.5128	39.3640	39.2388	39.1336	39.0450
WRTE	216.717	224.530	233.064	242.411	252.193	262.304
FIREE	47.8332	49.9942	52.3757	54.6506	56.6676	58.7909
TCPUE	55.9333	57.6744	59.1462	60.4462	61.6406	62.9050
SE	147.069	153.787	161.386	169.407	177.463	185.836
SLGE	158.954	165.211	171.752	178.849	186.418	194.249
FGCE	48.3507	46.5974	46.6247	47.0894	46.9408	46.7839
CLFA	1161.08	1175.47	1195.64	1220.88	1247.25	1272.44
OKAUNR	6.30955	5.57684	4.73949	4.49541	4.67431	4.71321
MFGWS	10.9475	11.7711	12.9450	14.0997	15.2308	16.4251
CCWS	12.4331	13.4946	15.0301	16.5461	18.0335	19.6055
MINWS	16.0849	17.4830	19.5182	21.5317	23.5121	25.6105
WRTWS	8.06166	8.64620	9.52496	10.4013	11.2655	12.1813
FIREWS	9.44974	10.2454	11.2992	12.3504	13.4004	14.5267
SERWS	7.89352	8.49475	9.35656	10.2056	11.0382	11.9183
FGCWS	15.1272	16.5557	18.5231	20.4832	22.4312	24.5110
SLGWS	7.91906	8.63214	9.52381	10.4361	11.3661	12.3666
TCPUWS	14.0885	15.3532	17.1034	18.8429	20.5676	22.4058
FWS	102.416	108.158	115.456	123.237	131.532	140.391
OLY	823.287	947.479	1130.86	1327.16	1530.53	1756.56

Equation	1976	1977	1978	1979	1980	1981
NPY	1038.06	1157.72	1275.75	1395.10	1524.99	1684.06
DIRY	2377.85	2681.27	3014.61	3332.10	3603.05	3891.41
TRY	2382.59	2576.60	2790.37	3008.88	3286.61	3552.02
CSSY	633.741	714.833	828.760	948.018	1070.92	1207.87
Z	.873144	.859772	.866641	.866398	.863426	.862891
SUT	178.291	195.001	214.098	235.467	258.955	283.464
TSIT	226.587	254.468	290.730	332.112	375.596	422.072
FIT	1306.38	1450.91	1615.05	1783.47	1956.37	2142.50
TPT	49.6181	52.3519	55.1602	58.1743	61.2695	64.5036
ABT	31.8934	34.7001	38.0644	41.8645	45.8304	50.0731
MFT	118.418	127.574	139.490	152.710	166.404	180.609
MVET	29.4463	33.5833	38.8001	44.7254	50.8594	57.4193
DD	4919.47	5433.16	6002.28	6594.98	7219.77	7920.83
TD	6822.14	8028.72	9343.54	10713.2	12163.5	13803.1
TL	11209.9	12828.0	14601.4	16467.4	18479.0	20612.3
GSP	14164.7	14728.9	15459.2	16148.6	16738.9	17394.1
NAE	912.205	943.362	978.317	1014.90	1050.96	1088.17
RELAUNR	.819422	.796691	.803304	.802752	.792256	.772658
WRTWSD	6.05230	6.15389	6.41844	6.59561	6.68971	6.83578
FIREWSD	7.09441	7.29208	7.61404	7.83158	7.95751	8.15189
TCPUWSD	10.5770	10.9275	11.5252	11.9486	12.2135	12.5734
SLGWSD	5.94524	6.14388	6.41766	6.61771	6.74946	6.93974
WSY	9474.57	10525.8	11986.9	13532.9	15126.2	16881.8
PYA	15834.7	17745.3	20009.3	22364.6	24803.6	27459.1
DPYA	14301.7	16039.9	18103.5	20249.0	22471.7	24894.5
PYAD	11887.9	12630.1	13483.3	14181.7	14729.0	15409.2
DPYAD	10737.1	11416.3	12199.1	12840.2	13344.2	13970.0
ST	857.745	946.913	1054.35	1175.22	1305.04	1444.48
D	13590.1	15170.8	17071.7	19176.2	21360.3	23683.1
P	15041.2	16790.0	18877.3	21186.9	23584.1	26131.4

VITA 2

Timothy Charles Ireland

Candidate for the Degree of

Doctor of Philosophy

Thesis: AN ECONOMETRIC MODEL FOR THE STATE OF OKLAHOMA

Major Field: Economics

Biographical:

Personal Data: Born in Enid, Oklahoma, August 16, 1952, the son of Dr. and Mrs. John B. Ireland.

Education: Graduated from Enid High School, Enid, Oklahoma, in May, 1970; received the Bachelor of Science degree in Mathematics from Phillips University in May, 1974; received the Master of Science degree in Economics from Oklahoma State University in May, 1976; completed the requirements for the Doctor of Philosophy degree in Economics at Oklahoma State University in December, 1978.

Professional Experience: Received Graduate Fellowship in the Department of Economics, Oklahoma State University, September, 1974, through August, 1977; Research Associate for the Oklahoma Econometric Modeling Project, sponsored by the Kerr Foundation, Inc., September, 1977, through August, 1978; part-time instructor in the Department of Economics, September, 1977, through January, 1978.

Professional Organizations: Phi Kappa Phi, Beta Gamma Sigma, Omicron Delta Epsilon, Southern Economic Association.