ESTIMATION OF A PRODUCTION FUNCTION FOR THE SMALL TOWN AND RURAL DEVELOPMENT POLICY IMPLICATIONS

by

Joseph Samuel Weidman

A Thesis Submitted to the Faculty of the

COMMITTEE ON URBAN PLANNING

In Partial Fulfillment of the Requirements
For the Degree of

MASTER OF SCIENCE

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 7 7

STATEMENT BY AUTHOR

This thesis has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Joseph S Weidman

APPROVAL BY THESIS DIRECTOR

This thesis has been approved on the date shown below:

Robert D. Carpenter

ROBERT D. CARPENTER
Professor of Geography, Regional Development and Urban Planning

March 23, 1977

Date

PREFACE

Work on this study began early in 1972 as part of a term paper assignment for a class in Rural Area Development. Continuing work on this subject was encouraged and supervised by Dr. Harry W. Ayer of the Agricultural Economics Department, and that department provided the funding needed for much of the research and data collection and all of the computer charges. The author wishes to take this opportunity to publicly acknowledge this assistance and to thank the Agricultural Economics Department, and especially Dr. Ayer, for their encouragement.

TABLE OF CONTENTS

TABLE OF CONTENTS--Continued

LIST OF TABLES

ABSTRACT

Cross-sectional data describing capital, labor, geographic, infrastructure and amenity variables associated with a sample of 20 Arizona communities were used to form a production function for the small town. Linear-regression analysis was used to estimate the terms in the production function. The production output was selected to be per capita income of town residents, or alternatively the percentage of town residents above the poverty level. The model predicted over 95% of the observed sample variance in all cases analyzed.

The percentage of the workforce having graduated from high school was found to be the most significant predictor of both per capita income and percentage above poverty. Various terms related to the provision of certain local amenities such as a golf course, a community or state college also were significant and contributed positively to improving local well-being. Other significant terms were related to the community's share of industrial and utility property, the occupational class of the workforce and the local share of state and federal government workers.

The results are compared with current rural development policy and many differences, along with some agreements, are

found. It is concluded that the model needs additional refinement and an expanded data base, but that the approach has merit in re-looking at current development policies.

CHAPTER 1

INTRODUCTION

This study had, as its primary goal, a better understanding of the factors which may contribute to improving the economic vitality of small to medium size Arizona communities. By focusing on differences between these communities, it was hoped that some common factors which are correlated with these differences could be established and a more structured approach to an understanding of possible causes could be taken.

A secondary goal of the study was to develop an analytical framework, or model, which could be used to show the interrelationship of those factors which are associated with the economic vitality of communities and which might then be used to quantitatively predict the result of changes in these factors. As will be discussed later, the rural community was viewed as a producing firm so that the well-established theory of production functions could be applied to its performance. Regression analysis methodology was used on typical cross-sectional "output" and "input" data to determine terms of the production function.

The general analysis method used involves the specification of the classical Cobb-Douglas production function, the terms

of which are derived from a linear-regression analysis of cross-sectional data obtained from a sample of Arizona cities. The Statistical Package for the Social Sciences (SPSS) (Nie et al. 1975) computer program was used for the regression analysis. Cross-sectional data sources included the U. S. Bureau of the Census (1971, 1972) covering housing and population, the Annual Report of the State Superintendent of Public Instruction (Shofstall 1969), Abstract of the Assessment Rolls by the Arizona Department of Revenue (1972), Community Profiles by the Arizona Office of Economic Planning and Development (1972), and various miscellaneous sources including The Arizona Blue Book (Arizona State Library Association 1971), highway maps and tourist brochures.

A recent paper by Ayer and Weidman (1976) presented results of this work from a somewhat different perspective. The Ayer and Weidman paper discussed small town vitality from an economic viewpoint. This theses expands that viewpoint to include some noneconomic factors of interest to urban planners. Also, the focus of this analysis is the individual living in the community, not the community itself. That is, quantities such as per capita income and the percentage of individuals living below the poverty level are used as measures of town vitality--not aggregate measures such as town income or population growth rate.

CHAPTER 2

METHODOLOGY

As mentioned earlier, the unit of analysis is the small to medium size Arizona community. The two measures of community well-being which were chosen were the per capita income of residents (obtained by dividing the total income of all persons received in the form of wages, salaries, dividends, etc., by the number of residents) and the percentage of individuals living above the poverty level (as defined in the U. S. Bureau of the Census 1971). Cross-sectional data were used rather than time series data. Much of the most valuable information about communities is contained in the decennial census and the rapid growth and changing conditions in many communities are such that only the 1970 Census data were considered useful for analytical purposes. Ideally, we would have time series data on a large number of communities to formulate and test our hypothesis. With only cross-sectional data, available at essentially one point in time, the testing of the theory becomes difficult. The data set was not large enough to permit breaking the total set into two sub-sets so that a theory could be developed with one and then tested on the second sub-set. This point will be further

discussed in a later section where the significance of the results will be estimated.

The City as a Producing Unit

It is plausible to view the city as a producing firm. Several authors have already done so (Haurin and Tolley 1972, Tolley 1969, Alonso 1971). Central to the concept is the assumption that leakages are small. That is, the city can be viewed as a producing firm which combines a mixture of capital, labor and entrepreneurial/managerial inputs to produce an output. In terms of the producing firm, the city acts to combine inputs of physical and human resources, under the direction of the managerial/governmental structure, in order to produce an output which is the total income received by its residents.

This study is concerned with understanding communities, not regions. Hence, it must be assumed that the community's resources go only toward the welfare of its residents. If the community is a "bedroom community," an industrial suburb, or has a symbiotic relationship with another city, the analysis may give false results. A related problem exists where a substantial number of residents live on incomes derived from investments in other areas. Wealthy retirement communities are the best example. No way was devised to measure, or to account for, transfer payments in this analysis.

For many nonmetropolitan communities in Arizona and other western states, where separations are large, this assumption of

independence is somewhat valid. But for metropolitan areas, at least, the appropriate unit of analysis should be the entire metropolitan region or urbanized area. As will be discussed later, the data base (i.e., sample size) for the analysis was so small that little attempt was made to carefully preselect cities which were included in the sample that was analyzed. A possibly fruitful avenue for a follow-up study would be to combine some of the smaller communities into sub-county functional economic areas and then to analyze these areas as independent units.

The Cobb-Douglas Production Function

The usual form of this production function, as originally advanced by Douglas (1948), is given by:

$$O = AK^{\alpha}L^{\beta}$$

where O is the output per unit time
K is the capital input per unit time
L is the labor input per unit time
A is a constant having dimension units and containing such factors as managerial skill, and
α and β are the elasticities of capital and labor, respectively.

Returns to scale are given by summing α and β. If $(\alpha+\beta)>1$, the returns to scale are positive and growth in each of the input factors will result in growth in the output.

Transformations Needed to Use Regression Analysis

Regression analysis was chosen as the manner in which the terms in the production function would be derived. This type of

analysis allows only linear forms of the independent variables. That is, the estimated equation is of the form:

$$Y = A_0 + A_1X_1 + A_2X_2 + \ldots A_nX_n + R$$

where Y is called the dependent variable
X's are the "independent" variables
A_0 through A_n are the estimated parameters,
and
R is the residual or error term.

The standard Cobb-Douglas production function is transformed into linear form by taking logarithms (to any base) of both sides so that the classical equation:

$$O = AK^{\alpha}L^{\beta}$$

is transformed into:

$$\log O = \log A + \alpha \log K + \beta \log L.$$

The quantities log A, α and β are now estimated by ordinary least squares regression.

Choice of Output Variables

For this study, per capita income and the number of people above the poverty level were chosen as the output or dependent variables. These variables are related but are not proxies as will be seen later. Per capita income was used as an output variable because it was considered that it best represented a quantifiable measure of people's well-being and because it is a proxy for the value added to the output of a community by use of the input factors. Higher per capita income means higher purchasing power for people (when the cost of living is constant), and an increased ability for them to control their own destiny

and make their own decisions about how to spend their surplus. Other income quantities were considered and discarded.

Among these were:

"family income"--felt to be too sensitive to number of workers in the family and too insensitive to differences in family size

"per capita income of work force population"--felt to be too insensitive to family size requirements

"income of families and unrelated individuals"--a great difference exists in the needs and apparent incomes of families and individuals living alone.

Per capita income, since it is computed by dividing the total value of wages and salaries received by all workers in the city by the total city population, was felt to best measure the economic well-being of the average city resident.

Because per capita income does not explicitly take into account how the income is distributed among the city residents, a second measure of resident well-being, the number of residents above the poverty level, was derived. It was decided to use the non-poor count rather than the number of people living below poverty so that we could always talk about maximizing concepts. In both cases, more of the output variable is desirable.

Statistical Significance Tests

The required statistical significance tests and computations are performed automatically by the SPSS computation procedure and the independent variable inclusion and deletion levels were set in the analysis so that all variables in the regression

equation were significant at the 90% confidence level. A step-wise procedure was used. The most significant term enters the equation first followed by the second most significant term. The procedure ends when none of the remaining terms are significantly related to the remaining variation in the dependent variable.

The alternate technique to a step-wise procedure is to postulate a priori the significant terms in the regression equation, and then, after forcing all terms into the final equation, to examine the ones which are statistically significant. This technique was not adopted because of the large possible number of independent variables. We could easily have "over-specified" the solution since we had more possible variables than sample points to fit with the regression function.

The adopted approach was to use many possible independent variables and see which ones "explain" the observed behavior. This approach has been criticized as "correlation hunting" but no apologies are offered. If there were a well-developed theory of urban growth dynamics applicable to small towns, the a priori specification of independent variables would seem to make sense, with the selection based upon terms other investigators had previously found to be significant.

CHAPTER 3

DATA BASE

As is typical in studies of this type, much time was spent in collecting, analyzing and rejecting various types of data. This section summarizes the data types, their sources and limitations. Additional details, including data sources, are presented in Appendix A.

Selection of Sample

It was decided to use data from all Arizona communities for which full census data were available (over 2,500 population) and for which full property valuation data were available (all counties except Maricopa, Pima and Greenlee). In addition, the specialized areas of Fort Huachuca and Yuma Station were excluded because of their almost total dependence on nearby military bases. The remaining 20 places for which all data were available and which were used in the analysis are listed below.

Benson
Bisbee
Casa Grande
Coolidge
Cottonwood
Douglas
Eloy
Flagstaff
Globe
Holbrook
Kearny
Kingman
Miami
Nogales
Prescott
Safford
Sierra Vista
Willcox
Winslow
Yuma

There is a slight problem in the set since Globe and Miami are very closely interrelated. It would perhaps have been better to combine the data into one sample but this was not done.

Independent Variables Concerned with Available Capital

As one of the major inputs in the production function, we might ideally use the actual cash values of all real and personal private and public property available in each of the sample communities. This capital list would also include all public investment in social overhead capital such as streets, sewage collection and treatment, schools, roads, etc., however the information is generally not available. Instead, six general property classes were formed and the total value of all property in each class was used as an independent variable. These classes are:

RESTOT--total value of "residential" property used for commercial purposes.

COMTOT--total value of "commercial" property used for general commercial purposes.

INDTOT--total value of "industrial" property.

UTLTOT--total value of property owned by utility companies.

TRATOT--total value of property used for transportation purposes.

NRTOT--total assessed value of all natural resources and related property. (For this property class the area over which the property values were summed included the entire elementary school district which contained the city in question.)

Although no data on the value of publicly-owned property which comprises a large portion of the social overhead capital

are available, it was decided to use the number of federal, state and local government employees as a proxy for this class of capital. Accordingly, three additional capital-related variables were defined:

FGTOT--total number of federal government employees in city

SGTOT--total number of state government employees in city

LGTOT--total number of local government employees in city

Since the populations of the communities in the data set differed by over a factor of 10, all of these capital variables were normalized by dividing by city population (PEOPLE) so that communities of different sizes could be more fairly compared. The corresponding population-normalized variables were RESPER, COMPER, INDPER, UTLPER, TRAPER, NRPER, FGPER, SGPER, AND LGPER.

Independent Variables Concerned with Available Labor Supply

The variables of interest here concern the composition of the work force in terms of skill levels. It is assumed that the presently employed workers in each city are working at their maximum skill levels (underemployment is ignored). Hence, the occupational breakdown of the total work force will indicate the absolute or relative percentage of workers trained as managers, for example, as compared to service workers. The following work force composition classes were used as labor-related independent variables:

GROUP1--number of workers in professional, technical management and administrative positions.

GROUP2--number of workers in craftsmen, foremen and related positions.

GROUP3--number of workers who are operatives, including transport equipment operatives.

GROUP4--number of workers in clerical, sales, service and related occupations.

GROUP5--all other workers--includes laborers, farmers, farm managers and farm foremen and private household workers.

Another important characteristic of the available labor force is its educational level. There are a number of ways to express this variable and, since earlier analyses showed educational level was highly correlated with income, it was decided that these data should not be highly aggregated but should express the various educational levels that existed in each community. Accordingly, the following education-related variables describing the work force--population over 25 years of age--were used in the analysis:

EDUCAT--the mean education for all persons over 25 years of age. (This variable was later discarded.)

COLGRA--the number of college graduates in the over-25 year old population.

SOMCOL--the number of over-25 year old population with over one year of college.

HIGHSC--the number of over-25 year old population who have graduated from high school.

As was the case with the capital-related variables, all the above variables (except EDUCAT) were normalized by dividing

by the appropriate population. The corresponding normalized variables were these: PERGR1, PERGP2, PERGP3, PERGP4, PERGP5, LABORATE (total workers per capita), PCTCOL, PCTSOM AND PCTHS.

Geographic Variables

This group of variables is included so that the effect of an unfavorable geographic position on personal income can be estimated. Presumably, if two cities are equal in terms of the capital and labor variables, the city with the longer transportation routes to its markets, or with extremes of climate, will suffer a disadvantage. Three new independent variables are created to examine this possibility:

ALTITUDE--in the author's opinion, an altitude of about 5,000 feet is ideal for Arizona. Cities below this level are too hot in the summer--cooling costs will increase--while cities above this altitude may have transportation problems in winter. This is admittedly a great oversimplification and implies that manufacturing is the important generator of wealth in the city. If seasonal tourism is more important, this altitude difference from 5,000 feet may be a positive factor for at least part of the year. However, ALTITUDE was processed by the computer to form the equation:

CLIMATE = (city's altitude in feet - 5,000)^2, so that a positive quantity was generated which became increasingly large as the departure from 5,000 feet increased.

ISOLATE--this is a measure of the city's separation from its major market center--presumably the nearest SMSA unless a much more careful analysis is done. ISOLATE is the measured distance in road miles from each city to the nearest SMSA. For most Arizona cities, this is Tucson or Phoenix, For a few cities, Las Vegas is closer.

SEPARATE--the separation, in miles, between the city in question and the nearest larger city. This is a measure of the market area of the city since even small isolated cities may capture trade from a wide market area if no larger center captures the population. The other side of this coin, of course, is a realization that, for some of these cities, population densities are so low that this capture effect is insignificant for city growth.

Infrastructure Variables

This group of variables relates to the mechanism by which the city functions as a producing unit. What we are seeking to establish is whether the city has the structural setup to make it an attractive place to live and, presumably, to start new businesses or to expand existing ones. A city, or place, with a municipal water and sewer system, a city manager, an adopted comprehensive plan, adequate housing, including a supply of rental units, and adequate transportation facilities would seem to have the prerequisites for economic expansion. The actual factors used in the analysis are listed below.

NOSEWERS--this factor is the percentage of residents not on public sewers. The remainder use septic tanks or other disposal means.

INCORP--unincorporated places suffer a great disadvantage in attracting most kinds of new business because they do not have established municipal services or a city manager, or similar full-time official, to keep things running smoothly. This independent variable is entered as a "dummy" variable with a value of "1" if the place is incorporated and a "0" if it is not. These dummy variable terms are not logarithmically transformed in order to avoid undefined mathematical operations. They were used in the analysis as "0's" or "1's" and the coefficients of any significant terms noted. Interpretation of these results is slightly different than for the continuous variables.

PLANNING--communities with an adopted and current comprehensive plan are better able to attract new business because they can show some evidence of knowing where they are heading and how the surrounding land may develop. It was not possible to poll each city to determine the existence and age of its plan, so a proxy quantity was determined by assuming that every city with a Planning and/or Zoning Commission had an ongoing planning process or was at least trying to control its future. This variable was also entered as a dummy variable with "1" denoting the existence of some planning function and "0" denoting no planning function existed.

The general condition of the housing stock was measured by means of two variables:

NEWHOUSE--the percentage of all housing built since 1950.

VACANCY-- the percentage of year-around housing units that are vacant.

Transportation facilities are measured by means of four dummy variables:

INTERSTA--places near an interstate highway have superior access to the nation's highway network and hence enjoy an advantage over less fortunately located places. This variable is scored as "1" if a community is within five miles of an interstate interchange, otherwise the score is "0."

AIR1-- basic airport facilities only--paved runway.

AIR2-- second level airport services; all of above facilities plus lighted runway(s), gasoline service and beacon.

AIR3-- third level airport services; all of above plus hanger facilities, FAA radio and mechanics available.

All three of these airport variables were coded as dummy variables, scored as "0" if the service was not available and as "1" if it was.

Amenity Variables

Klaassen (1968), in his provocative argument for the importance of "amenity" factors in stimulating area economic growth, suggested that factors such as housing supply, level of education services, medical care facilities, recreation opportunities, shopping facilities and physical planning, when supported by sound local financing, are important determinants in new plant

location decisions, and hence have a direct effect on the level of local income through an increase in the demand for local labor. (Klaassen argued that healthy, happy people are also better workers.) Some of these factors, namely housing and physical planning, have already been included in the model. In this section, we will attempt to include the remaining factors.

Level of educational facilities. The Annual Report of the State Superintendent of Public Instruction (Shofstall 1969) contains the desired information. Three numbers were extracted for the high school district in which each of our sample points was located: average daily membership, number of teachers and grand total expenditures. It was reasoned that the two ratios: teachers per pupil (denoted as TEACHPER) and dollars (of grand total expenditures) per pupil (denoted as BUCKSPER), would provide "input-level" measures on the quality of local education.

In addition to these quantitative measures of the educational system of the town, it was reasoned that the inclusion of a dummy variable, called COLLEGE, was important. This variable was scored as "1" if the town contained a community college or a branch of a state college. Towns with colleges should certainly be more attractive to potential investors and new plants than towns not so favored.

Medical care facilities. Two measures of medical care facilities were used in the analysis: hospital beds per person (BEDSPER) and doctors per person (DOCSPER). It is recognized

that these medical facilities may be designed to serve a larger population than that of the city in which they are located, but the availability of a large hospital and a large number of doctors would seem to be a positive factor in stimulating economic growth, following Klaassen's (1968) argument.

Recreation facilities. It was reasoned that almost all cities now have some form of municipal park and swimming pool. However, not all have golf courses available, private or public; therefore cities with golf courses should show some benefits in terms of attracting industries. This variable was coded as a dummy variable since it is not clear that the benefit of an 18-hole course is twice that of a 9-hole course. If a community had a golf course, the variable GOLF was scored as a "1," otherwise the score was "0."

Communications facilities. In addition to Klaassen's factors, presented above, the existence of a weekly and/or a daily newspaper, and a local radio station in the community were considered as important factors. The influence of these local communications media cannot be overestimated in shaping public opinion, especially since they are usually "pro-growth" and town boosters.

Three dummy variables, defined as indicated below, were used to qualitatively measure the communication infrastructure in the town. The variables are:

DAILYPAP--existence of a daily newspaper published in the town.

WEEKPAP--existence of a weekly, or more frequent, paper (but not a daily paper) published in the town.

RADIO--existence of one or more radio stations in the town.

It was first decided to include TV stations, but they are so infrequent in all except the very largest cities, that little contribution was expected to be made to a study of differences in smaller towns and cities.

A Scale Factor Variable

The variable PEOPLE was already mentioned as a normalization factor on such variables as those associated with the capital and labor supply. The variable itself was used as an independent variable because "growth pole" theory is still popular (Hoover 1967). Traditionally, the larger cities have been considered better candidates for economic development than smaller communities.

Summary of Data Base

Table 1 summarizes the various types of variables included in the analysis, their computer tag names, what they are intended to measure and the data source.

Discarded Variables

At various stages of the study, variables were included in the analysis and later discarded, usually because a "better"

Table 1. Summary of Variables Used in Analysis

Tag Name	Type	Intended to Measure	Data Source*
YPCAP	Dependent variable	Per capita income; well-being of people	1
NONPOOR	Dependent variable	Percentage above poverty level; alternative to YPCAP	1
PEOPLE	Independent--a scale variable	Size of city, and for normalization purposes	1
RESPER	Independent--a property variable	Per capita residential property used for commercial purposes	2
COMPER	Property variable	Per capita value of commercial property	2
INDPER	Property variable	Per capita value of industrial property	2
UTLPER	Property variable	Per capita value of utility property	2
TRAPER	Property variable	Per capita value of transportation (rail and airline) property	2
NRPER	Property variable	Per capita value of natural resources and mining property	2
FGPER	Proxy for federal property in city	Federal government employees per capita	1
SGPER	Proxy for state property in city	State government employees per capita	1
LGPER	Proxy for local government property	Local government employees per capita	1

Table 1, Summary of Variabes Used in Analysis, Continued

Tag Name	Type	Intended to Measure	Data Source*
PERGP1	Labor skill level variable	Percent of workforce in management and administration	1
PERGP2	Labor skill level variable	Percent of workforce who are craftsmen or foremen	1
PERGP3	Labor skill level variable	Percent of workforce who are operatives	1
PERGP4	Labor skill level variable	Percent of workforce in clerical, sales or service	1
PERGP5	Labor skill level variable	Percent of workforce in farming and laborers	1
LABORATE	Labor scale factor	Workers per capita	1
PCTCOL	Education-related variable	Percent of workforce who are college graduates	1
PCTSOM	Education-related variable	Percent of workforce who have had some college education	1
PCTHS	Education-related variable	Percent of workforce who are high school graduates	1
CLIMATE	Geography-related variable	Square of altitude departure from 5,000 feet	3
ISOLATE	Geography-related variable	Distance in road miles to nearest SMSA	4
SEPARATE	Geography-related variable	Distance in road miles to next larger city--a proxy for market area	4

Table 1, Summary of Variables Used in Analysis, Continued

Tag Name	Type	Intended to Measure	Data Source*
NOSEWERS	Infrastructure variable	Percentage of residents not on public sewers	5
INCORP	Infrastructure variable	Dummy variable = "1" if place is incorporated	1
PLANNING	Infrastructure variable	Dummy variable = "1" if a planning and zoning commission exists	6
NEWHOUSE	Infrastructure variable (housing-related)	Percentage of all housing built since 1950	5
VACANCY	Intrastructure variable (housing-related)	Percentage of year-round housing units which are vacant	5
INTERSTA	Infrastructure variable (transportation-related)	Dummy variable = "1" if city is within 5 miles of an interstate interchange	4
AIR1	Infrastructure variable (transportation-related)	Dummy variable = "1" if city had airport with paved runways	7
AIR2	Infrastructure variable (transportation-related)	Dummy variable - "1" if city had second level airport services	7
AIR3	Infrastructure variable (transportation-related)	Dummy variable - "1" if city had first class airport	7
TEACHPER	Amenity variable (education-related)	Teachers per pupil	8

Table 1, Summary of Variables Used in Analysis, Continued

Tag Name	Type	Intended to Measure	Data Source*
BUCKSPER	Amenity variable (education-related)	Grand total expenditures per pupil	8
COLLEGE	Amenity variable (education-related)	Dummy variable - "1" if city had a community college or branch of a state college	3
BEDSPER	Amenity variable (related to medical care level)	Hospital beds per person	8
DOCPSER	Amenity variable (related to medical care level)	Doctors per person	9
GOLF	Amenity variable (recreation-related)	Dummy variable = "1" if community has a 9-hole golf course	10
DAILYPAP	Amenity variable (communications-oriented)	Dummy variable - "1" if a daily newspaper is published	11
WEEKPAP	Amenity variable (communications-oriented)	Dummy variable - "1" if a weekly newspaper is published	11
RADIO	Amenity variable (communications-oriented)	Dummy variable - "1" if one or more local radio stations exist	12

*1 (U. S. Bureau of the Census 1971).

2 (Arizona Department of Revenue 1972).

3 (Arizona State Library Association 1971).

4 (Arizona Department of Highways 1971).

5 (U. S. Bureau of the Census 1972).

6 (League of Arizona Cities and Towns 1972).

7 (Arizona Department of Aeronautics 1971).

8 (Shofstall 1969).

9 (Arizona Health Planning Authority 1971).

10 (Arizona Office of Economic Planning and Development 1971).

11 (Arizona Newspapers Association 1970).

12 (Arizona Broadcasters Association 1970).

variable had been obtained. Table 2 lists the major discarded variables and summarizes the reasons for discarding.

In addition to these variables, various weighting factors were, at one stage in the analysis, applied to labor-skill level variables to account for the widely differing average wages received by these workers groups. This was felt to be an unnecessary complication since the regression analysis itself will determine the best weighting factors to apply to the labor-skill level variables.

Table 2. Discarded Variables

Tag Name	Measured	Why Discarded
EDUCAT	Mean years of schooling of population over 25 years old	Disaggregated to provide more information on relative importance of high school and college education
AGE	Date of incorporation of community	No economic justification for inclusion
AUTPER	Per capita value of automotive-related property in community	Summed into general commercial property class
AMUPER	Per capita value of amusement-related commercial property in community	Summed into general commercial property
MEDPER	Per capita value of medical-related property in community	Summed into general commercial property and also replaced by BEDSPER and DOCSPER
MINERS	Per capita number of miners in community	Used NRPER as a better measure of natural resources in community
MFGEMP	Per capita number of manufacturing employees in community	Used INDPER as a better measure; also used labor force characteristics
MAXT MINT	Maximum temperature in July; minimum temperature in January	Replaced by CLIMATE as a better measure of climatic effects
DENSITY	Population density of community	Too many communities have annexed large open areas for number to be meaningful
POPINC	Percentage population increase from 1960 to 1970	Considered as a possible dependent variable but discarded in favor of YPCAP

Table 2, Continued

Tag Name	Measured	Why Discarded
NETVAL	Net assessed valuation of all property in community	Disaggregated into property classes
TAXRAT	Tax rate in community	A proxy for level of services; it seemed dubious, at best
PUBADM	Number of public administrators in community	A proxy for level of public services; seemed very doubtful
IMPRES	Weighted sum of significant mineral resources, tourist attractions, large government installations and major universities	Seemed to be too much like adding "apples and oranges"; some of the terms were treated separately
SOCINS	Social institutions--weighted sum of banks, hospitals, community colleges, non-major state universities, newspapers and a courthouse	As above, a very subjective variable
TRANSP	Combined sum of highway, rail and air transportation facilities	A very subjective variable; the property class variables cover these same facilities

CHAPTER 4

RESULTS OF ANALYSIS

This section presents the results of the regression analysis in which equations for YPCAP and NONPOOR, the dependent variables, were derived in terms of the various capital, labor, amenity and infrastructure variables as discussed previously. All statistically significant terms are included in the equations.

Regression Equation for YPCAP

The equation obtained in logarithmic form for YPCAP is given by:

$$\begin{aligned} LOG(YPCAP) = {} & 2.7720 + .5213\ LOG(PCTHS) \\ & - .0279\ LOG(SGPER) + .2003\ LOG(UTLPER) \\ & - .4962\ LOG(PERGP4) + .0382\ COLLEGE \\ & + .0505\ LOG(FGPER) + .0353\ GOLF \end{aligned}$$

All of the coefficients are significantly different from 0 at the 95% confidence level except that of SGPER which is significant at the 90% confidence level. In this equation, all of the variables are in logarithmic form. By taking antilogs of both sides of the equation, we get:

$$YCAP = \$591.56\ \frac{(PCTHS)^{.5213} \times (COLLEGE)^{.0382} \times (UTLPER)^{.2003} \times}{(PERG4)^{.4962} \times (SGPER)^{.0279}}$$
$$\frac{(FGPER)^{.0505} \times (GOLF)^{.0353}}{}$$

The percentage of the sample variance which is "explained" by the given equation is 97.7%. That is, 97.7% of the inter-community variance in YPCAP is removed when all of the above terms are used to predict YPCAP.

The summary table presented as part of the SPSS output shows that the term PCTHS accounts for about 79% of the observed variance; SGPER and COLLEGE account for about 5%; UTLPER and PERGP4 account for another 3% each. FGPER and GOLF account for about 3% together.

A list of the communities, the actual YPCAP, the predicted YPCAP and the residual error is given in Table 3. The standard deviation of the residual error terms is $88 which is 3.5% of the mean of the predicted income.

The two dummy variable terms COLLEGE and GOLF must be interpreted somewhat differently than the continuous variable terms. Since the anti-log of 1 is 10 and the anti-log of 0 is 1, the "product" version of the equation for YPCAP shows that the "value" to a community of a community college or a branch of a state university is $(10)^{.0382}$ or 9.2%, while the "value" of a golf course is $(10)^{.0353}$ or 8.5%.

The implications of these results for rural development planning will be discussed in the next chapter.

Table 3. Results from Regression Analysis for YPCAP

City	Actual YPCAP	Predicted YPCAP	Error
Benson	$2450	$2513	$-63
Bisbee	2877	2858	19
Casa Grande	2436	2403	33
Coolidge	2385	2358	27
Cottonwood	2296	2415	-119
Douglas	2160	2222	-62
Eloy	1732	1718	14
Flagstaff	2801	2826	-25
Globe	2690	2627	63
Holbrook	2332	2417	-85
Kearny	3065	3095	-30
Kingman	3001	2847	153
Miami	2189	2177	12
Nogales	1782	1756	26
Prescott	2687	2798	-111
Safford	2179	2083	94
Sierra Vista	3706	3605	101
Willcox	2427	2458	-31
Winslow	2665	2694	-29
Yuma	2884	2845	39

Regression Equation for NONPOOR

A similar regression equation was obtained for NONPOOR with the result:

$$\text{NONPOOR} = \frac{(\text{PCTHS})^{.4598} \times (\text{CLIMATE})^{.0088}}{(\text{INDPER})^{.03592} \times (\text{TEACHPER})^{.2531} \times (\text{SGPER})^{.0184}}$$

All coefficients are significant at the 95% confidence level.

Almost 98% of the sample variance is "explained" by these five variables. The term PCTHS accounts for 72% of the variance; the term INDPER accounts for over 17.5% of the remaining variance; TEACHPER then explains an additional 4%; CLIMATE an additional 2% and SGPER just over 1%.

Table 4 displays the predicted and actual values of NONPOOR and the residual errors. The standard deviation of the residual error terms is 1.54% which is 1.9% of the mean value of NONPOOR.

It is interesting to note that there are many different terms in the equations for YPCAP and NONPOOR. These two quantities should be expected to be closely related, however, the only two terms shared are PCTHS and SGPER. Different "amenity" variables are in each equation--COLLEGE and GOLF in the YPCAP equation and CLIMATE and TEACHPER in the NONPOOR equation--as well as different "property" variables. It should also be noted that in general we always want "more" of PCTHS and "less" of SGPER. That is, for increased levels of YPCAP and NONPOOR, we want to

Table 4. Results from Regression Analysis for NONPOOR

City	Actual NONPOOR	Predicted NONPOOR	Error
Benson	85.2%	84.4%	0.8%
Bisbee	88.3	88.2	0.1
Casa Grande	78.1	75.5	2.6
Coolidge	76.4	77.7	-1.3
Cottonwood	79.1	80.5	-1.4
Douglas	77.8	79.0	-0.2
Eloy	60.4	61.0	-0.6
Flagstaff	87.8	88.3	-0.5
Globe	89.2	86.8	2.4
Holbrook	84.4	83.7	0.7
Kearny	99.5	102.0	-2.5
Kingman	90.4	88.6	1.8
Miami	86.7	85.2	1.5
Nogales	70.1	70.4	-0.3
Prescott	81.2	82.5	-1.3
Safford	84.1	83.5	0.6
Sierra Vista	92.5	93.0	-0.5
Willcox	80.2	80.0	0.2
Winslow	82.2	82.7	-0.5
Yuma	87.3	87.8	-0.5

increase the percentage of the over-25 population who are high school graduates and decrease the percentage of state government employees in the community. However, more federal government employees per capita helps to increase YPCAP.

The list of terms not in either equation is long. Some notable omissions are:

PEOPLE--no population size effect was noted.

RESPER, COMPER, TRAPER, NRPER, LGPER--most of the property-related capital variables were not significant, but this is possibly due to the high degree of inter-correlation between "independent" variables.

LABORATE, PERGP1, PERGP2, PERGP3, PERGP5--most of the labor-related variables did not enter the final equations. Again, this is possibly due to the high degree of inter-correlation.

PCTCOL and PCTSOM--only the percentage of high school graduates was found to be significant.

ISOLATE and SEPARATE--these classical geographical factors were not significant.

Only three of the many amenity variables were found to be significant.

CHAPTER 5

DISCUSSION OF RESULTS

It is tempting to impute a "cause and effect" relationship between the dependent and independent variables. However, without additional information, it is impossible to do this. Several of the observed relationships have been noticed by other investigators, and a close examination of the values of some of the independent variables in the sample communities reveals the relationship between the variables in other cases.

Overall Results--Form of Equations

The high values of R^2 indicate that almost all of the variance between communities is "explained" by the terms in the two equations. Only 2.3% of the sample variance in YPCAP is unexplained; only 2.2% of the sample variance in NONPOOR is not accounted for. These are very high figures and when it is realized that the YPCAP ranges from a high of $3,706 (Sierra Vista) to a low of $1,732 (Eloy) and that NONPOOR varies from 99.5% (Kearny) to a low of 60.4% (Eloy), the "explanatory power" of the equations is realized.

However, when the individual terms in each equation are examined and compared with the theory postulated earlier, a

number of discrepancies result. It was theorized that all of the independent variables except, perhaps, for PERGP4, PERGP5, CLIMATE, ISOLATE, and NOSEWERS could contribute positively to increasing both per capita income and the number of people above the poverty level. The results of the regression analyses indicated that the term SGPER, the per capita number of state government employees, had a negative effect on YPCAP while the equation for NONPOOR had a total of four "wrong" signs. One of the "wrong" signs in NONPOOR was held by the term INDPER, the per capita value of industrial property in the community.

In an effort to better understand what these equations were "explaining," the original data for each independent variable were examined and all values more than one standard deviation from the mean value were noted. These results were illuminating and are discussed below for each independent variable.

The "Education" Term--PCTHS

Wilbur Thompson conducted a cross-sectional multiple regression analysis of 135 Standard Metropolitan Statistical Areas which was summarized in the appendix to Thompson's (1968) article "Internal and External Factors in the Development of Urban Economies." The analysis used Census of Population and Census of Manufacturing data from the period 1940 to 1960. The dependent variables "explained" in the analyses were median family income,

percent of families earning less than $3,000 per year and a measure of family income inequality.

> The level of and distribution of education-median school years completed by persons twenty-five years old and over, percent of population with less than six years of education, and educational inequality . . . entered the income equations as independent variables so frequently that these local educational characteristics were set up as dependent variables to be explained (Thompson 1968, p. 64).

Thompson gave his ideas on how and why educational levels influence family income. They basically follow the ideas presented in this thesis, but Thompson felt these quantities also measure "occupational skill levels and the general cultural environment of the locality," an idea similar to Klaassen's (1968) discussions in Chapter 2 of his book.

Perloff and Wingo (1968, p. 8) in the introduction to Issues in Urban Economics discussed Thompson's work and stated "it is the totality of the supply factors in a given situation, the relative 'cumulative advantage' of the area, that becomes critical for growth." Very likely the significance of the education-related term (PCTHS) in both equations is due to a number of factors, as suggested above. The high amount of intercorrelation of this term with all other education-related terms implies that, to a certain degree, the terms are interchangeable and the analyst could pick the one he wanted to focus interest on. The Ayer and Weidman (1976) paper on page 81 uses the education variable as a measure of the "quality of human capital." In their (1976, p. 83) discussion of the importance of education,

the problem of out-migration of the better educated young people to urban areas is compared to the lack of an education-oriented thrust for rural development in the Rural Development Act of 1972. "In short, historical migration of the educated to urban areas does not relieve that fact that education has much to do with the productive capacity of rural towns (Ayer and Weidman 1976, p. 83)."

The "State Government" Term--SGPER

This is the second term which appeared in both equations, however its contribution to "explaining" YPCAP or NONPOOR was small in both cases. In an attempt to discover what this term was "explaining," a list of communities with high values of SGPER was prepared and it was discovered that two of the three cities with high values of SGPER, that is, large ratios of state government employees per capita, had colleges in the community (Coolidge and Flagstaff) and thus might be presumed to have an unusually large number of state employees as staff. The two lowest ratios of state employees per capita were in Bisbee and Sierra Vista, both of which receive "wealth" from other sources. Thus SGPER appears to be a correction term, primarily serving to reduce the predicted per capita income and number of NONPOOR in Flagstaff and Coolidge and to increase these predicted values in Bisbee and Sierra Vista. Since Sierra Vista has the highest per capita income in the sample, it is possible that SGPER really serves to help boost the prediction for the community.

The Property-Related Terms-- UTLPER and INDPER

The presence of UTLPER, the per capita value of utility property in the community, helped increase YPCAP. Safford and Douglas had low values of UTLPER, around $280,000 per person, while Holbrook and Sierra Vista had high values, between $800,000 and almost $1,000,000 per person. It is noted in the tabulations of per capita income that Sierra Vista has the highest income while Safford and Douglas have among the lowest incomes. The significance of a term such as UTLPER in the equation might be expected since it implies a high private sector investment in utility systems such as telephones, gas and electricity. Workers in these companies are usually well paid with good job security, hence the positive contribution of this term can be rationalized on several points.

The term INDPER, the per capita value of industrial property, has a negative effect on NONPOOR. It also accounts for over 17% of the variance in NONPOOR, so its effect is highly significant. An examination of the values of INDPER for each case showed that INDPER was very low in Miami, Kearny and Bisbee. It varied from about $2,000 to $7,000 per person and was very high in Casa Grande and Flagstaff, from over $200,000 to over $350,000 per person. Since the wealth of Miami, Kearny and Bisbee is concentrated in natural resources (mining), the industrial sector is underdeveloped in these communities. However, most families were living above the poverty level with Kearny having about

99.5% of its people living above poverty. The percentage above poverty in Casa Grande was only 78.1% and the error between the predicted and actual percentage of NONPOOR was 2.6%, the largest error in the sample. Thus, it should not be inferred that a high value of industrial property per capita causes many people to live below the poverty limit. Other factors are operating here and the small sample size tends to allow special cases in one or two towns to be "explained" by one independent variable. It should be noted, however, that having a high value of industrial property per capita was not a help, either in raising per capita income or in getting people above the poverty level. This is possibly explained by the low wages paid in 1969 (the base year for income) to industrial workers in some low-skill, non-unionized plants.

The Labor Skill Level Term--PERGP4

PERGP4 is the percentage of a community's workforce which is employed in clerical, sales or service occupations. Many of these jobs are traditionally low paid and it is not surprising that towns with high ratios of PERGP4 tend to have lower per capita incomes. Holbrook, Flagstaff and Prescott all have high numbers for this variable, 42 to 43% of the workforce is in this category, while Eloy and Kearny have the lowest ratios (22 to 25%). The first three cities are all tourist-oriented which tends to increase the availability of service and sales jobs. Kearny, a mining community, and Eloy, a farming-oriented

community, have little to attract the tourist and also many of the local residents do their shopping in nearby larger towns. Three percent of the inter-town variance in per capita income is accounted for by PERGP4, and, while the total effect is not large, there is a definite sign that towns with an already high dependence on tourism may not really improve income conditions for residents by increasing the proportion of their total work-force which is engaged in these jobs.

The "Federal Government" Term--FGPER

Like its counterpart SGPER, FGPER, the percentage of the town's population that works for the federal government, was included primarily as a surrogate for the per capita value of federally-owned property in the community. Since the assessment rolls do not list federal or state property because it is not taxed, the per capita number of federal and state employees was used instead. FGPER was only useful in helping account for variations in YPCAP where its presence removed only about 1.5% of the variance in YPCAP. Towns with high values of FGPER were Sierra Vista, Yuma and Benson, all located near important military bases. Kearny, Eloy and Cottonwood had low values of FGPER since there are no important federal employers nearby. Values of FGPER range from a low of two federal employees per thousand population in Kearny to a high of 188 in Sierra Vista. The average value of FGPER is about 15 federal employees per thousand population.

The "Amenity" Terms--COLLEGE, GOLF AND TEACHPER

COLLEGE and GOLF both had positive effects on YPCAP. As mentioned earlier, the presence of a college added 9.2% to YPCAP and the presence of a golf course added 8.5% to YPCAP. The communities having a college in 1970 were Coolidge, Douglas, Flagstaff, Kingman, Prescott and Yuma. The communities without a golf course were Eloy, Holbrook and Sierra Vista. It is likely that, in the case of golf courses, every community wants to have a course as soon as their finances permit. Towns with low values of YPCAP, such as Eloy, probably do not feel their residents are ready to pay for a course as yet. Since 1970, Holbrook has obtained a municipal golf course which it is struggling to maintain.

The presence of a college in a community undoubtedly has a positive benefit since it provides new jobs, attracts some students from other counties or states and provides residents with such cultural amenities as college sports, a few plays and concerts and a better library than would otherwise be the case.

The term TEACHPER, the teacher to pupil ratio in the town's schools, was introduced in an attempt to measure the quality of local education. Because of state and federal standards for education and because of active teachers' groups and parent-teacher associations, the range of values for TEACHPER is very small, varying from a high of 6.53 teachers per 100 pupils in Sierra Vista to a low of 4.25 teachers per 100 pupils in Yuma.

Coolidge also had a good teacher/pupil ratio while Nogales, Safford, Winslow and Flagstaff all had ratios nearly as low as did Yuma. The inclusion of TEACHPER in the equation for NONPOOR accounted for about 4% of the inter-town variance. TEACHPER did not enter the equation for YPCAP.

According to Klaassen's (1968) concept of amenity factors, higher values of TEACHPER should result in higher values of YPCAP and NONPOOR since industry will tend to locate in communities with a superior educational system. However, this idea was not substantiated in the present work. TEACHPER entered the equation for NONPOOR with a negative sign and the towns with the highest values of TEACHPER ranked second (Sierra Vista) and eighteenth (Coolidge) out of 20 in the NONPOOR list. Nogales, which had a low value for TEACHPER, had a low value for NONPOOR--nineteenth out of 20. It is not clear just what this term is "explaining." It may be that several factors are at work simultaneously. Rich communities such as Sierra Vista may have more teachers because they can afford them. Coolidge may have more teachers because the college can supply teachers cheaply, for practice teaching or aides, perhaps. Poorer communities such as Nogales may not be able to have as many teachers as they want. The companion term to TEACHPER was BUCKSPER, the per student educational budget. This term did not enter either equation and was also not highly correlated with TEACHPER.

The "Geography" Term--CLIMATE

It was reasoned that the climate of a community will have some effect on its economic vitality, especially if manufacturing is important. However CLIMATE, the square of the altitude departure from 5,000 feet, only entered the equation for NONPOOR where it accounted for about 2% of the variance. The sign of this term was positive which was not expected. The low altitude cities of Yuma, Casa Grande and Coolidge had the largest values for CLIMATE and the cities of Holbrook and Winslow, since they are near 5,000 feet altitude, had the lowest values. Cities on the upper and lower ends of the range of this variable are scattered throughout the NONPOOR distribution. Again, no real "explanation" for this variable was discovered.

Numerical Significance of Results

The exponents of the terms in a Cobb-Doulgas production function are the elasticities of the factors of production (Ayer and Weidman 1976, p. 82). Accordingly, the effect of a 10% increase in each of the terms in the equation for YPCAP is as follows:

PCTHS--increased by 10%--increases YPCAP by 5.2%

UTLPER--increased by 10%--increases YPCAP by 2.0%

FGPER--increased by 10%--increases YPCAP by 0.5%

A golf course increases YPCAP by 8.5% and a community college or a state college increases YPCAP by 9.2%.

The equation for NONPOOR showed that:

PCTHS--increased by 10%--increases NONPOOR by 4.6%

The other terms, with negative coefficients, are subject to interpretation, as discussed previously.

Comparison of Results with Present Rural Development Policies

At the present time five federal agencies supply the bulk of money for rural development. These are: Economic Development Administration, Four Corners Regional Commission, Environmental Protection Agency, the Department of Housing and Urban Development and the Farmers Home Administration. Of these agencies, the Farmers Home Administration is primarily engaged in subsidizing low cost homes for rural families. There is no attempt made by the FmHA to stimulate economic development through home construction loans or grants. They simply respond to existing needs. In view of the results of this analysis, this policy is sound. Recall that none of the housing-related amenity variables proved significant in increasing YPCAP or NONPOOR.

The EPA supplies money for studying water quality problems and for improving or installing wastewater collection and treatment facilities. No attempt is made by EPA to directly link grants to economic development, although they are sensitive to land use policies and controls and to the need to adequately plan for, and to accommodate, economic and population growth. The results of this study showed no correlation between economic

development and the percentage of houses connected to sewers. However it should be recalled that 1970 Census data were used and the effect of EPA policies and actions is only now being felt. It would not be surprising to find a different story if 1980 Census data were used.

HUD, through its Community Development Block Grant program, awards grants to communities for a wide range of projects. Many projects involve improving the quality of the existing housing stock (current emphasis is here). But grants have been awarded communities for upgrading the water system, a health/recreational center, street improvements and for fire stations, among other projects. No attempt is made by HUD to tie these grants to a direct improvement in economic conditions, but awards are based upon need, and the number and percentage of families below the poverty level is one of the factors considered. The results of this study showed no correlation between the housing-related amenity variables and our output variables. We did not investigate such public safety factors as adequacy of police and fire protection.

The major thrust of rural economic development comes from the Economic Development Administration (EDA) and the Four Corners Regional Commission (FCRC). A wide range of development-related projects can be funded by these two groups, which often cooperate to supply quite substantial sums to smaller communities that cannot, or will not, supply much local match. All types of

community facilities have been built, water and service lines installed and roads completed as a result of these agencies. Most of the money goes for public works, planning and assisting small business. The results of this study indicate that such projects do not appear to improve the vitality of the smaller communities. The central role played by PCTHS, and the lack of importance of the infrastructure variables, would indicate that a re-examination of EDA and FCRC policies could be in order. It is probably true that EDA and FCRC grants and loans have created some new jobs, and saved existing jobs, in many communities. However, the evidence in this study indicates that the upgrading of the existing work force is the single most beneficial policy that could be followed.

CHAPTER 6

CONCLUSIONS AND RECOMMENDATIONS

Except for the importance of the education-related variables, the results of this study are at variance with much of the traditional economic development literature. A companion study (Ayer and Weidman 1976) used a similar, not not identical, data base and similar variables and reached a number of different conclusions. It is suggested that some improvements need to be made to both the data base and the model and this general line of analysis continued in an effort to obtain a deeper insight into the nature and operation of the, perhaps, more basic determinants of town vitality. The following recommendations for further analysis are suggested:

1. The sample size should be expanded to include both larger and smaller communities. Property class data may have to be synthesized by judicious use and adjustment of the best available taxing jurisdiction information.
2. As an alternate, or as a separate study, ignore property class data and expand the sample size by using data from states "similar" to Arizona. Utah, New Mexico, Nevada, Western Colorado and Western Texas are possibilities.

3. The production function form should be generalized to include the variable returns to scale model suggested by Ulveling and Fletcher (1970). In view of the high values of R^2 which were found with the present model, the goal should be to retain the high R^2, but to use fewer terms in the production function.
4. Consider time-lagging some of the variables. Generally, causes precede effects and the real effect of variables such as those connected with education and capital formation might not show up for a number of years. In our case, we must use census data from 1970, so that some of the other variables should be taken from the 1960-1970 period.

It can be concluded from this study that use of a production function approach to investigating rural development policy is an interesting and useful technique. Use of the production function enables one to directly estimate the economic efficiency of suggested policies in any particular community. It also provides a framework for collecting data and classifying variables so that some additional insights into the question of suitable economic development policy can be gained. A considerable amount of model refinement, extension and further work needs to be done.

APPENDIX A

DATA BASE DETAILS

This material discusses the data used in the analysis and supplements the material in Chapter 3. The sources and limitations of the data are discussed more fully here. The philosophy behind "proxy" variables and variables that were used in early stages of the analyses and later discarded is discussed in Chapter 3.

Data Sources and Limitations

1970 Census of Population and Housing

Much of the needed data are available from this source. Data on personal income, poverty levels, population, work force, education and housing is generally available in printed form, on computer tape or on microfilm. Unfortunately, complete data are only available for cities and places of over 2,500 population. This proved to be a severe limitation when so many Arizona cities, especially the small rural cities of primary interest in this study, are below this size.

One possibility, which was tried and did not prove very successful, was to use CCD-level data. A CCD is a Census County Division and is a precisely defined sub-area of a county having

"more-or-less" the same characteristics. It may contain none, one, or more cities with populations over 2,500. If the CCD does contain such cities, the data for these places can be subtracted out of the total, leaving a remainder which, hopefully, represents more closely the smaller embedded cities.

Several CCD's have more than one small (i.e., under 2,500) city; one has three. With this method, CCD-level data are imputed to the small city and we have several samples with partially identical input data. This is not very realistic especially when some of the identical input data are very significant in predicting income levels, as was found to be the case.

Another problem is that, for the large rural-like counties, the area in a CCD is very large and the population of the CCD, even though sparse, is considerably larger than that of the small embedded cities.

1972 Abstract of Assessment Rolls--City Level

The Arizona Department of Property Valuation (DPV) annually publishes the abstract of the assessment roll for all property in the state. The usual published information is aggregated at the county and state levels. However, Dr. Arlan Larson, while serving as a special Legislative Assistant and consultant on school tax reform to the Arizona Legislature, obtained a special computer run from the DPV computer which aggregated the data at a city, school district and special tax district level. He kindly

furnished a copy of these data for this study. Unfortunately, the data are not complete at the city level for Maricopa, Pima and Greenlee counties. Efforts to obtain the missing data failed and hence this potentially valuable data source lost some of its effectiveness and served primarily to further reduce the sample size with full data since all cities in the three omitted counties were eliminated from the data set.

The data that are contained in the Abstract lists the full cash value, the assessment ratio and the net assessed value for all property in each taxing jurisdiction, in each of about 90 property classes. The assessment ratio can be used to separate the commercial and non-commercial uses of residential-type property such as apartment houses, motels, trailer parks, etc. Since residential property was assessed at 18% of full cash value (FCV) when these data were furnished (1972), and commercial property was assessed at 25% of FCV, any assessment ratio between these two figures is a measure of the relative value of the non-commercial and commercial property values. That is, assume an assessment ratio of R, where $.18 \leq R \leq .25$. For a property with FCV of P' dollars, the value of the commercial portion (P') is given by:

$$P' = P \times \frac{(R - .18)}{(.25 - .18)}$$

or $$P' = P \times \frac{R - .18}{(.07)}$$

It was reasoned that this adjustment should be made, especially for small cities, to take proper account of the numerous "mom and pop" motels, small stores and trailer courts. In larger cities where there are numerous big motels and stores, this adjustment is probably not worth the trouble.

It should also be noted that the only available data on property valuation were for the year 1972 and thus some error is introduced when the remainder of the data was obtained in 1970. No attempt was made to deflate 1972 data to the 1970 base year.

Community Profiles

The Arizona Office of Economic Planning and Development publishes a "Community Profile" for most of the cities in Arizona and this proved a fruitful source of data on amenity factors such as communication and health facilities, transportation routes, community institutions and climate. Some of these factors were originally used in the analysis but were later discarded as the sample size shrunk due to incomplete amenity data on some cities.

Annual Report of the State Superintendent of Public Instruction

This report contains information about the community's school system. In particular the average daily membership, the total expenditures and the number of teachers are listed for each school and some estimate of the quality of local education can then be obtained from these quantities.

Miscellaneous Data Sources

The current Arizona Highway Department map was used to obtain distances between cities. The Arizona Blue Book provided data on the altitude and area of the cities being studied.

Data Used in the Analysis

Selection of Sample

Ideally, the sample size used in a cross-sectional analysis should be as large as possible, limited only by the availability of data and the need to have the samples to be selected from the same universe that is being analyzed. Because of the critical importance of census data, it was decided to use data from all Arizona communities for which data were given in the 1970 census. Cities in the metropolitan area of Phoenix (i.e., Scottsdale, Tempe, Mesa, Paradise Valley and Glendale) and Tucson (Tucson and South Tucson) were eliminated since these areas are really outside the scope of interest of this study. In addition, the very specialized areas of Fort Huachuca, Luke and Williams Air Force Base and Yuma Station were eliminated because of their almost total dependence on nearby military bases. Sun City was eliminated because of its specialized function also. The remaining places, for which data were available, are listed below.

Ajo (U)	Casa Grande
Avondale	Cashion
Benson	Chandler
Bisbee	Clifton
Buckeye	Cottonwood

Douglas	Prescott
El Mirage	Safford
Eloy	San Carlos (U)
Flagstaff	San Manuel (U)
Globe	Sierra Vista
Holbrook	Superior (U)
Kearny	Tolleson
Kingman	West Yuma (U)
Miami	Wickenburg
Nogales	Willcox
Peoria	Winslow
Phoenix	Yuma

Total N = 34

Dependent (Output) Variables

Per capita personal income (YPCAP) measured in 1970 dollars and the percentage of residents living above the poverty level (NONPOOR) were chosen as the dependent variables. YPCAP is tabulated in Census Report PC(1)-C4, Arizona, General Social and Economic Characteristics, Tables 107-118. NONPOOR must be computed by subtracting the percentage of persons living below the poverty level from 100. The necessary data are also in Tables 107 and 118.

Independent Variables Concerned with Available Capital

Ideally, we would use the actual cash values of all real and personal private and public property available in each of the places in our sample. This capital list would also include all public investment in social overhead capital such as streets, sewerage collection and treatment, schools, etc., however this information is generally not available. In addition, the

available Abstract of the Assessment Roll did not contain city-level data for cities in Pima, Maricopa and Greenlee counties as mentioned above. Data were only available for a total of 20 cities and none of the unincorporated places. Some attempt was made to synthesize the data for the missing samples by using the elementary school or fire district data for approximately the same areas, but this method is of doubtful validity and may be subject to large errors. The list of cities for which real property values are available is given below.

Benson	Kearny
Bisbee	Kingman
Casa Grande	Miami
Coolidge	Nogales
Cottonwood	Prescott
Douglas	Safford
Eloy	Sierra Vista
Flagstaff	Willcox
Globe	Winslow
Holbrook	Yuma

Total N = 20

For each of these cities, the value of real property in each of about 90 property classes was made available through the assistance of Dr. Arlan Larson, as mentioned previously. This mass of data was aggregated into six general property classes.

RESTOT--total value of "residential" property used for commercial purposes. Includes property classes 04, 05 and 16. (See Appendix B for definitions of classes.)

COMTOT--total value of "commercial" property used for general commercial purposes. Includes property classes 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27 and 73.

INDTOT--total value of "industrial" property. Includes property classes 30, 31, 32, 33, 35 and 74.

UTLTOT--total value of property owned by utility companies. Includes classes 50, 52, 53, 54, 55, 56 and 57.

TRATOT--total value of property used for transportation purposes. Includes property classes 37, 51, 58 and 59.

NRTOT-- total assessed value of all natural resources and related property. Includes property classes 36, 60, 61, 63 and 75. For this property class the area over which the property values were summed included the entire elementary school district which contained the city in question. This was done to include the effect of large mines or mineral deposits located near the city, but outside the city limits.

Earlier in the analysis, additional property class groupings were used such as AUTTOT, the commercial property used for automotive-related purposes; AMUTOT, the value of property used for amusement or entertainment purposes and MEDTOT, the value of property used for medically-related purposes. These additional classes did not seem to improve the power of the analysis and hence these classes were assumed into the COMTOT class.

Although no data on the value of publically-owned property, which comprises a large portion of the social-overhead capital, are available, it was decided to use the number of federal, state and local government employees as a proxy for this class of capital. Accordingly, three additional capital-related variables were defined.

FGTOT--total number of federal government employees in city

SGTOT--total number of state government employees in city

LGTOT--total number of local government employees in city

In using these variables, it is assumed that on the average, each government employee is supported by some level of capital investment. This level probably varies considerably from one city to another, however including these terms was considered to be better than ignoring this area entirely.

It is necessary to use the 4th Count, File C, Census of Population computer tapes available at the University of Arizona Computer Center in order to obtain these data. (Table 67 lists the needed information.) The University of Arizona Library, Government Documents Center, has a similar computer printout with this information. The data published in PC(1)-C4 is not sufficiently detailed.

All of these capital variables were normalized by dividing by city population (PEOPLE) so that cities of different sizes could be fairly compared. The corresponding population-normalized variables were: RESPER, COMPER, INDPER, UTLPER, TRAPER, NRPER, FGPER, SGPER and LGPER.

Independent Variables Concerned with Available Labor Supply

We are in considerably better shape here than in the area of capital-related variables. Most of the labor-related variables are published in PC(1)-C4, but some must be found in census data tapes. The variables of interest concern the composition of the work force in terms of skill levels. It is assumed that the presently employed workers in each city are working at their

maximum skill levels. Underemployment is ignored here. Hence, the occupationa breakdown of the total work force will indicate absolute or relative percentages of workers trained as managers, for example, as compared to service workers. The following work force composition classes were used as labor-related independent variables.

Group 1--number of workers in professional, technical management and administrative positions

Group 2--number of workers in craftsmen, foremen and related positions

Group 3--number of workers who are operatives, including transport equipment operatives

Group 4--number of workers in clerical, sales, service and related occupations

Group 5--all other workers--includes laborers, farmers, farm managers and farm foremen and private household workers. (Group 5 was really computed by subtracting the sum of Group 1 through Group 4 from the total number of employed persons in the community.)

Another important characteristic of the available labor force is its educational level. There are a number of ways to express this variable and, since earlier analyses showed educational level was highly correlated with YPCAP, it was decided that these data should not be highly aggregated but should express the various educational levels that existed in each community. Accordingly, the following education-related variables describing the work force (population over 25 years of age) were used in the analysis:

EDUCAT--the mean education for all persons over 25 years of age. (This variable was later discarded.)

COLGRA--the number of college graduates in the over-25 year old population

SOMCOL--the number of over-25 years old population with over one year of college

HIGHSC--the number of over-25 year population who have graduated from high school

All of the needed data are available in PC(1)-C4, Tables 103 and 117.

As was the case with the capital-related variables, all the above variables (except EDUCAT) were normalized by dividing by the appropriate population. The corresponding normalized variables were these: PERGP1, PERGP2, PERGP3, PERGP4, PERGP5, LABORATE (total workers per capita), PCTCOL, PCTSOM and PCTHS.

Geographic Variables

This group of variables is included so that the effect of an unfavorable geographic position on personal income can be estimated. Three new independent variables were created to examine this possibility:

ALTITUDE--in the author's opinion, an altitude of about 5,000 feet is ideal for Arizona. Cities below this level are too hot in the summer--cooling costs will add to manufacturing costs--while cities above this altitude may have transportation problems in winter. ALTITUDE was processed by the computer to form the equation:

CLIMATE = (city's altitude in feet--5,000)^2, so that a positive quantity was generated which became increasingly large as the departure from 5,000 feet increased.

ISOLATE--this is a measure of the city's separation from its major market center--presumably the nearest SMSA unless a much more careful analysis is done. ISOLATE is the measured distance in road miles from each city to the nearest SMSA. For most Arizona cities, this is Tucson or Phoenix. For a few cities, Las Vegas is closer.

SEPARATE--the separation, in miles, between the city in question and the nearest larger city.

All of the above information can be obtained from a good map and The Arizona Blue Book (for altitudes).

Infrastructure Variables

This group of variables relates to the mechanism by which the city functions as a producing unit. A city, or place, with a municipal water and sewer system, a city manager, an adopted comprehensive plan, adequate housing, including a supply of rental units, and adequate transportation facilities would seem to have the prerequisites for economic expansion.

NOSEWERS--this factor is the percentage of residents who are not on public sewers. The remainder use septic tanks or other disposal means. Data are available from the 1970 Census of Housing, HC(1)-B4, Detailed Housing Characteristics, Arizona.

INCORP--unincorporated places suffer a great disadvantage in attracting most kinds of new business because they do not have established municipal services or a city manager, or similar full-time official, to keep things running smoothly. This independent variable is entered as a "dummy" variable with a value of "1" if the place is incorporated and a "0" if it is not.

PLANNING--communities with an adopted and current comprehensive plan are possibly better able to attract new business because they can show some evidence of knowing where they are heading and how the surrounding land may develop. It was not possible to poll each city to determine the existence and age of its plan, so a proxy quantity was determined by assuming that every city with a Planning and/or Zoning Commission had an ongoing planning process or was at least trying to control its future. The data source was the League of Arizona Cities and Towns' (1972) publication, Directory of Arizona City and Town Officials. This variable was also entered as a dummy variable with "1" denoting the existence of some planning function and "0" denoting no planning function existed.

The general condition of the housing stock was measured by means of two variables:

NEWHOUSE--the percentage of all housing built since 1950

VACANCY-- the percentage of year-around housing units that are vacant

These data are available in the 1970 Census of Housing, HC(1)-A4, Tables 18 and 23 and in HC(1)-B4, Tables 56 and 58.

Transportation facilities are measured by means of four dummy variables:

INTERSTA--places near an interstate highway have superior access to the nation's highway network, and hence enjoy an advantage over less fortunately located places. This variable is scored as "1" if a community is within five miles of an interstate interchange, otherwise the score is "0."

AIR1-- basic airport facilities only, paved runway

AIR2-- second level airport services; all of above plus hanger facilities plus lighted runway(s), gasoline service and beacon

AIR3-- third level airport services; all of above plus hanger facilities, FAA radio and mechanics available

All three of these airport variables were coded as dummy variables, scored as "0" if the service was not available and as "1" if it was.

Amenity Variables

Klaassen (1968) suggested in Chapter 2 that factors such as housing supply, level of education services, medical care facilities, recreation opportunities, shopping facilities and physical planning, when supported by sound local financing, are important determinants in new plant location decisions, and hence have a direct effect on the level of local income through an increase in the demand for local labor. Some of these factors, namely housing and physical planning, have already been included in the model.

Level of educational facilities. The Annual Report of the State Superintendent of Public Instruction (Shofstall 1969)

contains the desired information. Three numbers were extracted for the high school district in which each of our sample points was located: average daily membership, number of teachers and grant total expenditures. It was reasoned that the two ratios: tachers per pupil (denoted as TEACHPER) and dollars (of grand total expenditure) per pupil (denoted as BUCKSPER), would provide "input-level" measures on the quality of local education.

In addition to these quantitative measures of the educational system of the town, it was reasoned that the inclusion of a dummy variable, called COLLEGE, was important. This variable was scored as "1" if the town contained a community college or a branch of a state college.

Medical care facilities. Two measures of medical care facilities were used in the analysis: hospital beds per person (BEDSPER) and doctors per person (DOCSPER). It is recognized that these medical facilities may be designed to serve a larger population than that of the city in which they are located, but the availability of a large hospital and a large number of doctors would seem to be a positive factor in stimulating economic growth, following Klaassen's argument. The necessary data were found in the Arizona State Plan for Hospital and Medical Facilities (Arizona Health Planning Authority, Arizona State Department of Health 1971).

Recreation facilities. It was reasoned that almost all cities now have some form of municipal park and swimming pool. However, not all have golf courses available, private or public;

therefore cities with golf courses should show some benefits in terms of attracting industries. This variable was coded as a dummy variable since it is not clear that the benefit of an 18-hole course is twice that of a 9-hole course. If a community had a golf course, the variable GOLF was scored as a "1," otherwise the score was "0." The data were obtained from Amazing Arizona, a publication of the Travel Information Section of the Arizona Office of Economic Planning and Development (1971).

Communications facilities. In addition to Klaassen's (1968) factors, presented above, the existence of a weekly and/or a daily newspaper, and a local radio station in the community were considered as important factors.

Three dummy variables, defined as indicated below, were used to qualitatively measure the communication infrastructure in the town. The variables are:

DAILYPAP--existence of a daily newspaper published in the town

WEEKPAP-- existence of a weekly, or more frequent paper (but not a daily paper) published in the town

RADIO-- existence of one or more radio stations in the town

It was first decided to include TV stations, but they are so infrequent in all except the very largest cities, that little contribution was expected to be made to a study of differences in smaller towns and cities.

All dummy variables were coded as a "1" if the facility existed and "0" otherwise.

Data for newspapers were obtained from the Arizona Newspapers Association (1970) publication 1970 Advertising Rates and Data of Newspapers and Other Publications in Arizona. Data on radio stations were found in Directory: Arizona Radio and TV Stations, 1970, published by the Arizona Broadcasters Association (1970).

A Scale Factor Variable

The variable PEOPLE was already mentioned as a normalization factor on such variables as those associated with the capital and labor supply. The variable itself was used as an independent variable because "growth pole" theory is still popular (Hoover 1967). Traditionally, the larger cities have been considered better candidates for economic development than smaller communities.

In addition to these variables, various weighting factors were, at one stage in the analysis, applied to labor-skill level variables to account for the widely differing average wages received by these worker groups. This was felt to be an unnecessary complication since the regression analysis itself will determine the best weighting factors to apply to the labor-skill level variables.

APPENDIX B

PROPERTY CLASSES USED BY DEPARTMENT OF PROPERTY VALUATION IN ASSESSING REAL PROPERTY IN ARIZONA

The following material was obtained from the Arizona Department of Revenue (1972), formerly the Department of Property Valuation.

Standard Land Use Classification Codes

00 RESIDENTIAL PROPERTY
01 Single Family
02 Double & Duplex
03 Apartments
04 Hotel
05 Motel
06 Resorts
07 Condominiums
08 Trailer Parks
09 Improved Rural Homesite (Non-Subdiv.)

10 COMMERCIAL PROPERTY
11 Commercial Buildings--One Story
12 Store & Office Comb.
13 Dept. Store Mult. Story
14 Shopping Centers
15 Office Buildings & Mult. Story
16 Bank & Savings & Loan
17 Service Stations--Petroleum Bulk Plants
18 Auto Sales Repair & Storage
19 Nursing Homes
20 Restaurants & Bars
21 Medical Buildings
22 Parking Lots
23 Nurseries & Gree Houses
24 Drive In Theaters, Golf Courses, Race Tracks
25 Theaters, Bowling Alleys, Skating Rinks
27 Clubs & Lodge Halls

Number Not Used: 26

30 INDUSTRIAL PROPERTY
31 Manufacturing--Durable & Non-Durable Products
32 Manufacturing--(Food & Kindred Products)
34 Lumbering, Saw Mills, Planing Mills
35 Cotton Gins & Compress
36 Mining, Quarrying & Processing
37 Warehousing, Storage & Truck Terminals

40 FARM PROPERTY
41 Field Crops
42 Vineyards
43 Other Tree Crops
44 Citrus
45 High Density Production of Livestock, Fish, Poultry, or Produce
46 Immature Citrus, Vineyards, & Tree Crops

47 RANCH PROPERTY
48 Pasture Land
49 Fallow Crop Land

*50 UTILITIES PROPERTY
*51 Railroad Operating Property
*52 Telephone & Telegraph Operating Property
*53 Pipe Line Operating Property
*54 Gas & Electric Utility Operating Property
*55 Water Utility Operating Property
*56 Community Antenna T.V. Systems Operating Property
*57 Salt River Project
*58 Airline Flight Property
*59 Private Car Companies

60 NATURAL RESOURCES
*61 Producing Mines
*63 Producing Oil & Gas Interests
66 Non-Producing Mine Property
68 Mineral Rights Only
69 Standing Timber

70 PERSONAL PROPERTY
71 Household
72 Mobile Homes
73 Commercial
74 Industrial
75 Mining, Quarring & Processing
76 Farm & Ranch Mach. & Equipment
77 Merchants & Manufacturing Inventory

Numbers Not Used: 33, 38, 39, 62, 64, 65, 67
*Appraised by the Department

78 LIVESTOCK & MISCELLANEOUS ANIMALS
79 Beef Cattle
80 Dairy Cattle
81 Sheep
82 Swine
83 Horses
84 Poultry

85 UNDEVELOPED RURAL LANDS
86 Undeveloped Rural Lands--with Misc. Imps.
87 Non-Qualifying Agriculture Land
88 Lands not Included in any other class

90 TAX EXEMPT PROPERTY
91 Schools
92 Religious & Charitable
93 Hospitals
94 Public
95 Cemetery

Numbers Not Used: 89, 96, 97, 98, 99

LIST OF REFERENCES

Alonso, W. 1971. "The Economics of Urban Size." Papers of the Regional Science Association. Volume 26.

Arizona Broadcasters Association. 1970. Directory: Arizona Radio and TV Stations. Phoenix, Arizona.

Arizona Department of Aeronautics. 1971. Arizona Aeronautical Chart. Phoenix, Arizona.

Arizona Department of Highways. 1971. Arizona Highways Map. Phoenix, Arizona.

Arizona Department of Revenue. 1972. Abstract of Assessment Rolls. Special computer run prepared for Dr. Arlan Larson. Phoenix, Arizona.

Arizona Health Planning Authority, Arizona State Department of Health. 1971. Arizona State Plan for Hospital and Medical Facilities. Phoenix, Arizona.

Arizona Newspapers Association. 1970. 1970 Advertising Rates and Data of Newspapers and Other Publications in Arizona. Phoenix, Arizona.

Arizona Office of Economic Planning and Development. 1971. Amazing Arizona. Phoenix, Arizona.

_____. 1972. Community Profiles. Phoenix, Arizona.

Arizona State Library Association. 1971. The Arizona Blue Book. R. Yates, Editor. Tucson, Arizona.

Ayer, H. W., and J. S. Weidman. 1976. "The Rural Town as a Producing Unit: An Empirical Analysis and Implications for Rural Development Policy." Southern Journal of Agricultural Economics. December, 1976. Pp. 79-87.

Douglas, P. H. 1948. "Are There Laws of Production?" The American Economic Review. Volume 27, March, 1948. Pp. 1-41.

Haurin, D. R., and G. S. Tolley. 1972. "The Rural Town and the Scale Question." Chicago, Illinois. The University of Chicago, Urban Economics Report No. 82. Presented at the meeting of the Western Agricultural Economics Association. Logan, Utah, July 26, 1972.

Hoover, E. M. 1967. Some Old and New Issues in Regional Development. Pittsburgh, Pennsylvania. University of Pittsburgh, Center for Regional Economic Studies Occasional Paper No. 5.

Klaassen, L. H. 1968. Social Amenities in Area Economic Growth. Paris: Organization for Economic Cooperation and Development. Chapter 2.

League of Arizona Cities and Towns. 1972. Directory of Arizona City and Town Officials. Phoenix, Arizona.

Nie, N. H., C. H. Hall, J. G. Jenkins, K. Steinbrenner, and D. H. Bent. 1975. Statistical Package for the Social Sciences. Second Edition. New York: McGraw-Hill Book Co.

Perloff, H. S., and L. Wingo. 1968. Issues in Urban Economics. Baltimore: Johns Hopkins Press.

Shofstall, W. P. 1969. Annual Report of the Superintendent of Public Instruction for the Fiscal Year 1968-1969. Phoenix, Arizona. State of Arizona, Department of Public Instruction.

Thompson, W. R. 1968. "Internal and External Factors in the Development of Urban Economies," in Issues in Urban Economics, editors H. S. Perloff and L. Wingo. Baltimore: Johns Hopkins Press.

Tolley, G. S. 1969. "The Welfare Economics of City Bigness." Chicago, Illinois. Urban Economics Report #31. University of Chicago.

Ulveling, E. F., and L. B. Fletcher. 1970. "A Cobb-Douglas Production Function with Variable Returns to Scale." American Journal of Agricultural Economics, Volume 52, No. 2, May, 1970, 322-326.

U. S. Bureau of the Census. 1971. General Social and Economic Characteristics, Arizona. PC(1)-C4 Ariz. Washington, D. C.: U. S. Government Printing Office.

_____. 1972. Detailed Housing Characteristics, Arizona. HC(1)-B4 Ariz. Washington, D. C.: U. S. Government Printing Oiffce.