Title

Estimating Interregional Population and Migration Operators From Interregional Population Distributions

Author

Rogers, Andrei

Publication Date

1967-04-01

Peer reviewed

ESTIMATING INTERREGIONAL POPULATION AND MIGRATION OPERATORS FROM INTERREGIONAL POPULATION DISTRIBUTIONS

Andrei Rogers
Department of City and Regional Planning
University of California, Berkeley

Prepared for the Mathematical Demography Session of the Annual Meeting of the Population Association of America, Cincinnati, Ohio, April 28-29, 1967

ESTIMATING INTERREGIONAL POPULATION AND MIGRATION OPERATORS FROM INTERREGIONAL POPULATION DISTRIBUTIONS[1]

1. Introduction

Efforts to analyze and forecast interregional population growth and distribution frequently are severely constrained by the general paucity of reliable data describing the behavior of the fundamental components of population change. Migration data, for example, are simply not available for most regional subdivisions of the nation. As a result demographers have had to rely on rather crudely constructed "residual" methods of estimating migration. Recent efforts to express the population growth process in matrix form, however, suggest a means whereby the growth regime of an interregional system may be estimated solely on the basis of a time series of interregional population distributions. In this paper, we present such a method and illustrate it using data on a two-region system consisting o. California and the rest of the United States.

For ease of exposition we shall focus on the simplest possible population model. Such a model is the cohort-survival model with only a single cohort: the total population. This model may be defined by the matrix equation:[2]

$$
\begin{bmatrix} w_1^{(t+1)} \\ w_2^{(t+1)} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ w_m^{(t+1)} \end{bmatrix} = \begin{bmatrix} g_{11} & g_{21} & \cdots & g_{m1} \\ g_{12} & g_{22} & & \cdot \\ \cdot & & \ddots & \cdot \\ \cdot & & & \cdot \\ g_{1m} & \cdots & \cdots & g_{mm} \end{bmatrix} \begin{bmatrix} w_1^{(t)} \\ w_2^{(t)} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ w_m^{(t)} \end{bmatrix} \tag{1.1}
$$

where $w_i^{(t)}$ = total population of region i at time t;

g_{ij} = proportion of people who were in region i at time t and in region j at time t + 1;

g_{ii} = proportion of people who were in region i at time t and remained there until time t + 1 plus the rate of natural increase in region i during the unit time interval (t, t + 1).

2. Population Operator Estimators

Consider the population growth process, in a two-region system, that occurs over two sequential periods. Assume that an unknown schedule of growth, defined by the matrix G in (1.1), remains unchanged during this period of time. We have then:

$$\begin{bmatrix} w_1^{(t+1)} \\ w_2^{(t+1)} \end{bmatrix} = \begin{bmatrix} g_{11} & g_{21} \\ g_{12} & g_{22} \end{bmatrix} \begin{bmatrix} w_1^{(t)} \\ w_2^{(t)} \end{bmatrix}$$

and

$$\begin{bmatrix} w_1^{(t+2)} \\ w_2^{(t+2)} \end{bmatrix} = \begin{bmatrix} g_{11} & g_{21} \\ g_{12} & g_{22} \end{bmatrix} \begin{bmatrix} w_1^{(t+1)} \\ w_2^{(t+1)} \end{bmatrix} .$$

If the w_i's have been observed, and therefore are known, the matrix population growth operator, G, may be derived by solving the following system of 4 equations in 4 unknowns:[3]

$$w_1^{(t+1)} = g_{11}w_1^{(t)} + g_{21}w_2^{(t)}$$

$$w_2^{(t+1)} = g_{12}w_1^{(t)} + g_{22}w_2^{(t)}$$

$$w_1^{(t+2)} = g_{11}w_1^{(t+1)} + g_{21}w_2^{(t+1)}$$

$$w_2^{(t+2)} = g_{12}w_1^{(t+1)} + g_{22}w_2^{(t+1)}$$

Thus, for example, given 1950-1952 population data for the two-region system of California and the rest of the United States (Table 1), we have:

$$11130 = 10643g_{11} + 141225g_{21}$$

$$142852 = 10643g_{12} + 141225g_{22}$$

$$11638 = 11130g_{11} + 142852g_{21}$$

$$144755 = 11130g_{12} + 142852g_{22}$$

The more general problem of estimating such a G on the basis of a time series of distributions, however, requires a criterion of fit, since more equations than unknowns are available. Three estimators have been suggested in the literature: (1) the unrestricted least squares estimator;[4] (2) the minimum absolute deviations estimator;[5] and (3) the restricted least squares estimator.[6] The first adopts the well-known estimation technique traditionally used in regression analysis; the latter two require a mathematical programming formulation. In this paper only the first two estimation techniques are considered. The restricted least squares solution will be described in a forthcoming paper.

<u>The Unrestricted Least Squares (ULS) Estimator</u>

Assume that the observed vectors, $\underline{w}$, denoting the population in each of m regions over n years, are generated by the linear statistical model

TABLE 1.--ESTIMATED TOTAL POPULATION OF CALIFORNIA AND THE REST OF THE UNITED STATES, 1950 - 1960*

Year (July 1)	California (000's)	Rest of the United States (000's)	National Total (000's)
1950	10643	141225	151868
1951	11130	142852	153982
1952	11638	144755	156393
1953	12101	146855	158956
1954	12517	149367	161884
1955	13004	152065	165069
1956	13581	154507	168088
1957	14177	157010	171187
1958	14741	159408	174149
1959	15288	161637**	176925**
1960	15863	163284**	179147**

*Source: Economic Development Agency, California Statistical Abstract (Sacramento, Calif.: State of California, Documents Section, 1963), p. 44.

**Continental U.S. only (i.e., excluding Alaska and Hawaii).

(2.1) $$\underline{w}^{(t+1)} = G\underline{w}^{(t)} + \underline{\epsilon}^{(t)}$$

where $\underline{\epsilon}^{(t)}$ is a $(mn \times 1)$ vector of error terms. Expressing Equation (2.1) in the more familiar linear regression form, we have:

(2.2) $$\underline{y} = X\underline{g} + \underline{\epsilon}$$

where, for our two-region example,

$$\underline{y} = \begin{bmatrix} \underline{y}_1 \\ \\ \underline{y}_2 \end{bmatrix} \quad \underline{g} = \begin{bmatrix} \underline{g}_1 \\ \\ \underline{g}_2 \end{bmatrix} \quad \underline{\epsilon} = \begin{bmatrix} \underline{\epsilon}_1 \\ \\ \underline{\epsilon}_2 \end{bmatrix} \quad X = \begin{bmatrix} X_1 & 0 \\ \\ 0 & X_2 \end{bmatrix} X_1 = X_2$$

or, more specifically,

$$\underset{2n \times 1}{\underline{y}} = \begin{bmatrix} w_1^{(t+1)} \\ w_1^{(t+2)} \\ \cdot \\ \cdot \\ w_1^{(t+n)} \\ \hline w_2^{(t+1)} \\ w_2^{(t+2)} \\ \cdot \\ \cdot \\ w_2^{(t+n)} \end{bmatrix} \quad \underset{2^2 \times 1}{\underline{g}} = \begin{bmatrix} g_{11} \\ g_{21} \\ \hline g_{12} \\ g_{22} \end{bmatrix} \quad \underset{2n \times 1}{\underline{\epsilon}} = \begin{bmatrix} \epsilon_{11} \\ \epsilon_{12} \\ \cdot \\ \cdot \\ \epsilon_{1n} \\ \hline \epsilon_{21} \\ \epsilon_{22} \\ \cdot \\ \cdot \\ \epsilon_{2n} \end{bmatrix}$$

and

$$\underset{2n \times 2^2}{X} = \left[\begin{array}{cc|cc} w_1^{(t)} & w_2^{(t)} & 0 & 0 \\ w_1^{(t+1)} & w_2^{(t+1)} & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ w_1^{(t+n-1)} & w_2^{(t+n-1)} & 0 & 0 \\ \hline 0 & 0 & w_1^{(t)} & w_2^{(t)} \\ 0 & 0 & w_1^{(t+1)} & w_2^{(t+1)} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & w_1^{(t+n-1)} & w_2^{(t+n-1)} \end{array}\right] .$$

Recalling that the least squares estimator of $\underline{g}$ is one which minimizes the sum of squared deviations of observed from predicted values, we minimize

$$(2.3) \qquad S = \Sigma(y_i - \hat{y}_i)^2 = [\underline{y} - X\underline{g}]'[\underline{y} - X\underline{g}]$$

with respect to $\underline{g}$ and find

$$(2.4) \qquad \hat{\underline{g}} = (X'X)^{-1}X'\underline{y} \quad .$$

Applying Equation (2.4) to our two-region example of Table 1, we have the estimation process illustrated in Figure 1 and the results presented in Figure 2. Table 2 provides data for comparing the discrepancies between observed and predicted values.

Results. The results, presented in Figure 2, are very unrealistic. First of all, the presence of a negative migration rate in two of the estimated growth operators is, by definition, unacceptable since it implies a negative migration flow. Second, in the one case where this particular migration

FIG. 1.--UNRESTRICTED LEAST SQUARES ESTIMATE OF THE ANNUAL POPULATION GROWTH OPERATOR FOR CALIFORNIA AND THE REST OF THE UNITED STATES: 1950-1960 TOTAL POPULATION TIME SERIES

<u>Estimator</u>:

$$\underline{g} = (X'X)^{-1}X'\underline{y}$$

<u>Data</u>:

$$X = \left[\begin{array}{cc|cc}
10643 & 141225 & 0 & 0 \\
11130 & 142852 & 0 & 0 \\
11638 & 144755 & 0 & 0 \\
12101 & 146855 & 0 & 0 \\
12517 & 149367 & 0 & 0 \\
13004 & 152065 & 0 & 0 \\
13581 & 154507 & 0 & 0 \\
14177 & 157010 & 0 & 0 \\
14741 & 159408 & 0 & 0 \\
15288 & 161637 & 0 & 0 \\
\hline
0 & 0 & 10643 & 141225 \\
0 & 0 & 11130 & 142852 \\
0 & 0 & 11638 & 144755 \\
0 & 0 & 12101 & 146855 \\
0 & 0 & 12517 & 149367 \\
0 & 0 & 13004 & 152065 \\
0 & 0 & 13581 & 154507 \\
0 & 0 & 14177 & 157010 \\
0 & 0 & 14741 & 159408 \\
0 & 0 & 15288 & 161637
\end{array}\right] \qquad \underline{y} = \left[\begin{array}{c}
11130 \\ 11638 \\ 12101 \\ 12517 \\ 13004 \\ 13581 \\ 14177 \\ 14741 \\ 15288 \\ 15863 \\
\hline
142852 \\ 144755 \\ 146855 \\ 149367 \\ 152065 \\ 154507 \\ 157010 \\ 159408 \\ 161637 \\ 163284
\end{array}\right]$$

<u>Estimate</u>:

$$\hat{\underline{g}} = \left[\begin{array}{r}
1.0149 \\ .0022 \\ \hline -.0557 \\ 1.0194
\end{array}\right]$$

FIG. 2.--UNRESTRICTED LEAST SQUARES ESTIMATES: TOTAL POPULATION

A. 1950-1960 Time Series:

$$\begin{bmatrix} 1.0149 & .0022 \\ -.0557 & 1.0194 \end{bmatrix}$$

$R^2 = 1.0000$

B. 1950-1955 Time Series:

$$\begin{bmatrix} .9456 & .0076 \\ .7839 & .9522 \end{bmatrix}$$

$R^2 = 1.0000$

C. 1955-1960 Time Series:

$$\begin{bmatrix} .9603 & .0072 \\ -.6169 & 1.0700 \end{bmatrix}$$

$R^2 = 1.0000$

TABLE 2.--TOTAL POPULATION OF CALIFORNIA AND THE REST OF THE UNITED STATES, 1950-1960: OBSERVED AND PREDICTED

A. California (000's):

Year (July 1)	Observed*	Predicted** 1950-1960	Predicted** 1950-1955	Predicted** 1955-1960
1951	11130	11110	11139	
1952	11638	11608	11612	
1953	12101	12128	12106	
1954	12517	12602	12560	
1955	13004	13030	12973	
1956	13581	13530		13587
1957	14177	14121		14158
1958	14741	14732		14749
1959	15288	15309		15308
1960	15863	15809		15849

B. Rest of the United States (000's):

Year (July 1)	Observed*	Predicted** 1950-1960	Predicted** 1950-1955	Predicted** 1955-1960
1951	142852	143367	142819	
1952	144755	144998	144750	
1953	146855	146910	146961	
1954	149367	149025	149323	
1955	152065	151562	152041	
1956	154507	154285		154682
1957	157010	156743		156939
1958	159408	159261		159249
1959	161637	161674		161467
1960	163284	163916		163515

* Source: Table 1.

** Derived by estimated growth operators in Figs. 2A, 2B, and 2C, respectively.

rate is nonnegative it is unrealistically large. Finally, the diagonal element denoting the growth of California in the absence of in-migration is less than unity on two occasions, a most unlikely event. Thus it appears that the ULS estimator is not very satisfactory for our purposes. We, therefore, turn next to a weighted estimator in order to explore whether the peculiar results are in any way related to the great difference in the size of the populations of our two regions.

<u>The Weighted Unrestricted Least Squares (WULS) Estimator</u>

The unweighted ULS estimator assumes that

$$E(\underline{\epsilon\epsilon}') = \delta^2 I \quad , \tag{2.5}$$

i.e., that the error terms are independently distributed with a constant variance. Both assumptions are likely to be violated in our estimation problem. Because of the lagged variable relationship in our model, a dependence exists between the error terms and the explanatory variables. And since the populations of the regions differ considerably in size, it is unlikely that the condition of homoscedasticity is present. For example, one might reasonably expect the variance of the error term to be related to the size of the population being estimated.

The problem of lagged variables is a thorny one, and its resolution in our problem context is beyond the scope of this paper. We must bear in mind, therefore, that our least squares estimates are likely to be biased. The presence of heteroscedasticity, however, may be taken into account by a suitable transformation of the data and, therefore, will be briefly explored here.

Assume, for example, that the standard deviation of the error term in the estimating equation for $w_i^{(t+1)}$ is directly proportional to the size of $w_i^{(t)}$. Then in our model

$$(2.6) \qquad E(\underline{\epsilon\epsilon}') = \delta^2 W^2$$

where

$$(2.7) \qquad W = \left[\begin{array}{ccccc|ccccc}
w_1^{(t)} & \cdots & & & 0 & 0 & \cdots & & & 0 \\
\vdots & w_1^{(t+1)} & & & \vdots & \vdots & \ddots & & & \vdots \\
\vdots & & \ddots & & \vdots & \vdots & & \ddots & & \vdots \\
0 & \cdots & & & w_1^{(t+n-1)} & 0 & \cdots & & & 0 \\
\hline
0 & \cdots & & & 0 & w_2^{(t)} & \cdots & & & 0 \\
\vdots & \ddots & & & \vdots & \vdots & w_2^{(t+1)} & & & \vdots \\
\vdots & & \ddots & & \vdots & \vdots & & \ddots & & \vdots \\
0 & \cdots & & & 0 & 0 & \cdots & & & w_2^{(t+n-1)}
\end{array}\right].$$

If now we premultiply our original model by W^{-1}, we have

$$(2.8) \qquad W^{-1}\underline{y} = W^{-1}X\underline{g} + W^{-1}\underline{\epsilon}$$

or

$$(2.9) \qquad \underline{y}^* = X^*\underline{g} + \underline{\epsilon}^*$$

where

$$\underline{y}^* = W^{-1}\underline{y}\ , \quad X^* = W^{-1}X\ , \quad \text{and} \quad \underline{\epsilon}^* = W^{-1}\underline{\epsilon}\ .$$

The variance-covariance matrix for the transformed vector of error terms, $\underline{\epsilon}^*$, is

$$E(\underline{\epsilon}^{*}\underline{\epsilon}^{*\prime}) = E(W^{-1}\underline{\epsilon\epsilon}'W^{-1}) \tag{2.10}$$

$$= W^{-1}E(\underline{\epsilon\epsilon}')W^{-1}$$

$$= W^{-1}\delta^{2}W^{2}W^{-1}$$

$$= \delta^{2}I \quad .$$

Hence we may apply the ULS estimator directly to the transformed model in Equation (2.9). Table 3 outlines the computational sequence for transforming the data. Figure 3 illustrates the weighted estimation procedure. Figure 4 presents the three estimated population growth operators.

Results. The weighted ULS estimates in Figure 4 do not appreciably differ from the unweighted ULS estimates in Figure 2. Thus it appears that weighting does not significantly alter the estimation results. We, therefore, turn to a different criterion of fit: one which gives less weight to observations with large deviations from the mean.

The Minimum Absolute Deviations (MAD) Estimator

The minimum absolute deviations estimator finds the vector of parameter estimates which minimizes the sum of the absolute values of the deviations between observed and predicted values. That is, it derives the vector $\underline{g}$ which minimizes

$$S = \Sigma|y_i - \hat{y}_i| = |\underline{y} - X\underline{g}|'\underline{1} \quad , \tag{2.11}$$

subject to the constraints

$$\underline{y} = X\underline{g} + \underline{\epsilon} \tag{2.12}$$

$$\underline{g} \geqq \underline{0} \quad . \tag{2.13}$$

TABLE 3.--WEIGHTED TOTAL POPULATION OF CALIFORNIA AND THE REST OF THE UNITED STATES, 1950-1960

Year	Unweighted Populations (000's)				Weights		Weighted Populations (000's)					
	Cal.	R.U.S.	Cal.	R.U.S.	Cal.	R.U.S.	Cal.	R.U.S.	Cal.	R.U.S.	Cal.	R.U.S.
	$\underline{w}_1$	$\underline{w}_2$	$\underline{y}_1$	$\underline{y}_2$			$\underline{w}_{11}$	$\underline{w}_{21}$	$\underline{w}_{12}$	$\underline{w}_{22}$	$\underline{y}_1$	$\underline{y}_2$
t	$w_1^{(t)}$	$w_2^{(t)}$	$w_1^{(t+1)}$	$w_2^{(t+1)}$	$w_1^{(t)}$	$w_2^{(t)}$	$w_1^{(t)}/w_1^{(t)}$	$w_2^{(t)}/w_1^{(t)}$	$w_1^{(t)}/w_2^{(t)}$	$w_2^{(t)}/w_2^{(t)}$	$w_1^{(t+1)}/w_1^{(t)}$	$w_2^{(t+1)}/$
1950	10643	141225	11130	142852	10643	141225	1	13.2693	.0754	1	1.0458	1.011
1951	11130	142852	11638	144755	11130	142852	1	12.8349	.0779	1	1.0456	1.013
1952	11638	144755	12101	146855	11638	144755	1	12.4381	.0804	1	1.0398	1.014
1953	12101	146855	12517	149367	12101	146855	1	12.1358	.0824	1	1.0344	1.017
1954	12517	149367	13004	152065	12517	149367	1	11.9331	.0838	1	1.0389	1.018
1955	13004	152065	13581	154507	13004	152065	1	11.6937	.0855	1	1.0444	1.016
1956	13581	154507	14177	157010	13581	154507	1	11.3767	.0879	1	1.0439	1.016
1957	14177	157010	14741	159408	14177	157010	1	11.0750	.0903	1	1.0398	1.015
1958	14741	159408	15288	161637	14741	159408	1	10.8139	.0925	1	1.0371	1.014
1959	15288	161637	15863	163284	15288	161637	1	10.5728	.0946	1	1.0376	1.010

FIG. 3.--UNRESTRICTED LEAST SQUARES ESTIMATE OF THE ANNUAL POPULATION GROWTH OPERATOR FOR CALIFORNIA AND THE REST OF THE UNITED STATES: 1950-1960 WEIGHTED TOTAL POPULATION TIME SERIES

Estimator: $\underline{g} = (X'X)^{-1}X'\underline{y}$

Data:

$$X = \left[\begin{array}{cc|cc} 1 & 13.2693 & 0 & 0 \\ 1 & 12.8349 & 0 & 0 \\ 1 & 12.4381 & 0 & 0 \\ 1 & 12.1358 & 0 & 0 \\ 1 & 11.9331 & 0 & 0 \\ 1 & 11.6937 & 0 & 0 \\ 1 & 11.3767 & 0 & 0 \\ 1 & 11.0750 & 0 & 0 \\ 1 & 10.8139 & 0 & 0 \\ 1 & 10.5728 & 0 & 0 \\ \hline 0 & 0 & .0754 & 1 \\ 0 & 0 & .0779 & 1 \\ 0 & 0 & .0804 & 1 \\ 0 & 0 & .0824 & 1 \\ 0 & 0 & .0838 & 1 \\ 0 & 0 & .0855 & 1 \\ 0 & 0 & .0879 & 1 \\ 0 & 0 & .0903 & 1 \\ 0 & 0 & .0925 & 1 \\ 0 & 0 & .0946 & 1 \end{array}\right] \qquad \underline{y} = \left[\begin{array}{c} 1.0458 \\ 1.0456 \\ 1.0398 \\ 1.0344 \\ 1.0389 \\ 1.0444 \\ 1.0439 \\ 1.0398 \\ 1.0371 \\ 1.0376 \\ \hline 1.0115 \\ 1.0133 \\ 1.0145 \\ 1.0171 \\ 1.0181 \\ 1.0161 \\ 1.0162 \\ 1.0153 \\ 1.0140 \\ 1.0102 \end{array}\right]$$

Estimate:

$$\hat{\underline{g}} = \left[\begin{array}{c} 1.0138 \\ .0023 \\ \hline -.0249 \\ 1.0167 \end{array}\right]$$

FIG. 4.--UNRESTRICTED LEAST SQUARES ESTIMATES: WEIGHTED TOTAL POPULATION

A. 1950-1960 Time Series:

$$\begin{bmatrix} 1.0138 & .0023 \\ -.0249 & 1.0167 \end{bmatrix}$$

$R^2 = 1.0000$

B. 1950-1955 Time Series:

$$\begin{bmatrix} .9446 & .0077 \\ .7899 & .9517 \end{bmatrix}$$

$R^2 = 1.0000$

C. 1955-1960 Time Series:

$$\begin{bmatrix} .9593 & .0073 \\ -.6042 & 1.0688 \end{bmatrix}$$

$R^2 = 1.0000$

To transform (2.11) into a form that is solvable by linear programming methods we express each ϵ_i as the difference of two nonnegative variables u_i and v_i. Thus we now seek the vector $\underline{g}$ which minimizes

(2.14) $$S = [\underline{u} + \underline{v}]'\underline{1}$$

subject to the constraints

(2.15) $$\underline{y} = X\underline{g} + \underline{\epsilon} = X\underline{g} + [I, -I]\begin{bmatrix} \underline{u} \\ \underline{v} \end{bmatrix}$$

(2.16) $$\underline{g}, \underline{u}, \underline{v} \geq \underline{0} .$$

The problem can be solved by using the simplex algorithm on the tableau presented in Table 4. Since, in (2.15), the vectors $\underline{u}$ and $\underline{v}$ are linearly dependent, elements of one or the other, but not both, will enter the optimal solution.[7]

A simplex tableau made up of the population data in Table 1 appears in Figure 5. Figure 6 presents the three population growth operators that result from the application of the unweighted and weighted estimators, respectively.

<u>Results</u>. The MAD estimators appear to resolve one of the problems presented by the ULS and WULS estimators: namely, that of the possible entry of negative elements into the growth operator. However, the MAD estimator does not seem to provide improved results in two important respects:

(1) Although the MAD estimator is sure to produce a nonnegative population growth operator, experiments conducted in the course of preparing this paper suggest that what appears as a negative element in the ULS and WULS estimates often merely assumes a zero value in the MAD estimate. Zero migration, though a more reasonable result than negative migration, is, nevertheless, still a rather improbable event.

TABLE 4.--SIMPLEX TABLEAU FOR FINDING THE MINIMUM ABSOLUTE DEVIATIONS ESTIMATE

	$\underline{0}'$	$\underline{0}'$	...	$\underline{0}'$	$\underline{1}'$	$\underline{1}'$	...	$\underline{1}'$	$\underline{1}'$	$\underline{1}'$	...	$\underline{1}'$
B	$\underline{g}_1'$	$\underline{g}_2'$	...	$\underline{g}_r'$	$\underline{u}_1'$	$\underline{u}_2'$	...	$\underline{u}_r'$	$\underline{v}_1'$	$\underline{v}_2'$	...	$\underline{v}_r'$
$\underline{y}_1$	X_1				I_n				$-I_n$			
$\underline{y}_2$		X_2				I_n				$-I_n$		
...			...				...				...	
$\underline{y}_r$				X_r				I_n				$-I_n$

FIG. 5.--SIMPLEX TABLEAU FOR MINIMUM DEVIATIONS ESTIMATE OF THE ANNUAL POPULATION GROWTH OPERATOR FOR CALIFORNIA AND THE REST OF THE UNITED STATES: 1950-1960 TOTAL POPULATION TIME SERIES

	0	0	0	0	1	1	. . . .	1	1	. . . .
B	g_{11}	g_{21}	g_{12}	g_{22}	u_{11}	u_{12}	. . . .	v_{11}	v_{12}	. . . .
11130	10643	141225			1			-1		
11638	11130	142852				1			-1	
12101	11638	144755								
12517	12101	146855								
13004	12517	149367								
13581	13004	152065								
14177	13581	154507								
14741	14177	157010								
15288	14741	159408								
15863	15288	161637								
142852			10643	141225						
144755			11130	142852						
146855			11638	144755						
149367			12101	146855						
152065			12517	149367						
154507			13004	152065						
157010			13581	154507						
159408			14177	157010						
161637			14741	159408			1			-1
163284			15288	161637			1			-1

FIG. 6.--MINIMUM ABSOLUTE DEVIATIONS ESTIMATES: TOTAL POPULATION AND WEIGHTED TOTAL POPULATION

A. 1950-1960 Time Series:

Unweighted

$$\begin{bmatrix} 1.0057 & .0030 \\ .0774 & 1.0083 \end{bmatrix}$$

Z = .3123

Weighted

$$\begin{bmatrix} 1.0054 & .0030 \\ .0808 & 1.0080 \end{bmatrix}$$

Z = .0431

B. 1950-1955 Time Series:

Unweighted

$$\begin{bmatrix} .9504 & .0072 \\ .7753 & .9531 \end{bmatrix}$$

Z = .0284

Weighted

$$\begin{bmatrix} .9500 & .0072 \\ .7857 & .9523 \end{bmatrix}$$

Z = .0101

C. 1955-1960 Time Series:

Unweighted

$$\begin{bmatrix} .9577 & .0074 \\ .0 & 1.0153 \end{bmatrix}$$

Z = .1350

Weighted

$$\begin{bmatrix} .9575 & .0074 \\ .0 & 1.0153 \end{bmatrix}$$

Z = .0122

(2) The elements of the population growth matrix occasionally take on very unlikely dimensions. In particular, the diagonal elements frequently are less than unity and the off-diagonal entries, in order to compensate for this, become excessively large.

In the following section of the paper we wish to analyze the causes of such behavior and to find a means for "smoothing" the data, prior to the estimation process, in order to obtain more reasonable estimates. To do this we draw on the frequently used result in regression analysis which establishes that any multiple linear regression problem can be decomposed and solved as a series of simple linear regressions.[8] Such a technique forms the basis of most stepwise regression procedures.

3. Estimation With Smoothed Data

The Population Model as a Series of Simple Linear Regressions

Consider the fitted regression equation

$$Y_i = b_0 + b_1 X_{1i} + b_2 X_{2i} \quad . \tag{3.1}$$

We can obtain the vector of coefficients, $\underline{b}$, simultaneously, by using the matrix least squares estimator

$$\underline{b} = (X'X)^{-1} X' \underline{y} \quad . \tag{3.2}$$

Alternatively, we may adopt the following stepwise procedure to obtain the same result:

(1) Regress Y on X_1 and compute the residuals $Y_i - \hat{Y}_i$ where $\hat{Y}_i = a_0 + a_1 X_{1i}$.

(2) Regress X_2 on X_1 and compute the residuals $X_{2i} - \hat{X}_{2i}$ where $\hat{X}_{2i} = c_0 + c_1 X_{1i}$.

(3) Regress the residuals $Y_i - \hat{Y}_i$ on $X_{2i} - \hat{X}_{2i}$ by fitting the model

$$(3.3) \qquad (Y_i - \hat{Y}_i) = \delta(X_{2i} - \hat{X}_{2i}) + \epsilon_i \quad .$$

No intercept term is required, since we are dealing with two sets of residuals whose sums are zero.

If we now consider the fitted equation

$$(3.4) \qquad (Y_i - \hat{Y}_i) = d(X_{2i} - \hat{X}_{2i}) \quad ,$$

and denote $\hat{Y}$ and $\hat{X}_2$ as functions of X_1, we have

$$(3.4) \qquad [Y_i - (a_0 + a_1X_1)] = d[X_{2i} - (c_0 + c_1X_{1i})] \ ,$$

or, collecting terms and attaching a caret to Y to identify it as a fitted value,

$$(3.5) \qquad \hat{Y}_i = (a_0 - dc_0) + (a_1 - dc_1)X_{1i} + dX_{2i}$$

$$= b_0 + b_1X_{1i} + b_2X_{2i} \quad .$$

Applying the above decomposition procedure to our population model, we find that

$$(3.6) \qquad w_1^{(t+1)} = (a_1 - g_{21}c_1)w_1^{(t)} + g_{21}w_2^{(t)} \quad ,$$

and

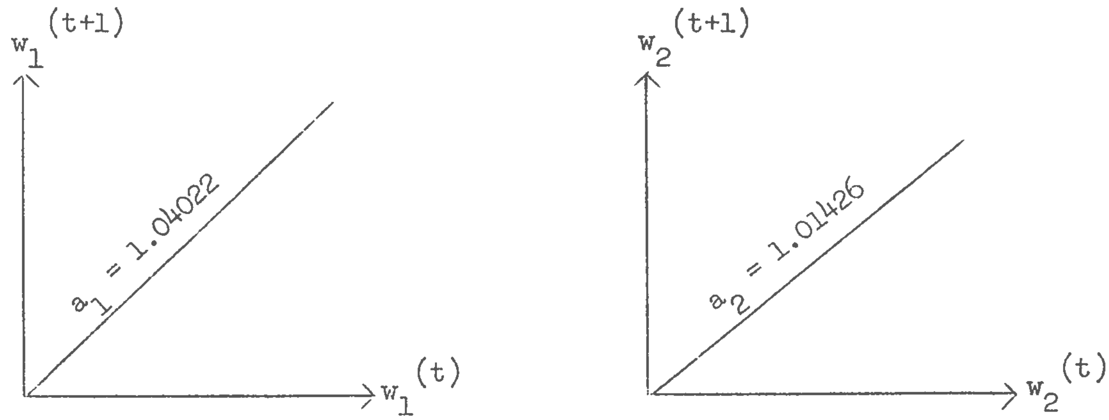
$$(3.7) \qquad w_2^{(t+1)} = (a_2 - g_{12}c_2)w_2^{(t)} + g_{12}w_1^{(t)} \quad .$$

Since, in our population model, the linear equation is forced through the origin, all intercept terms are zero.

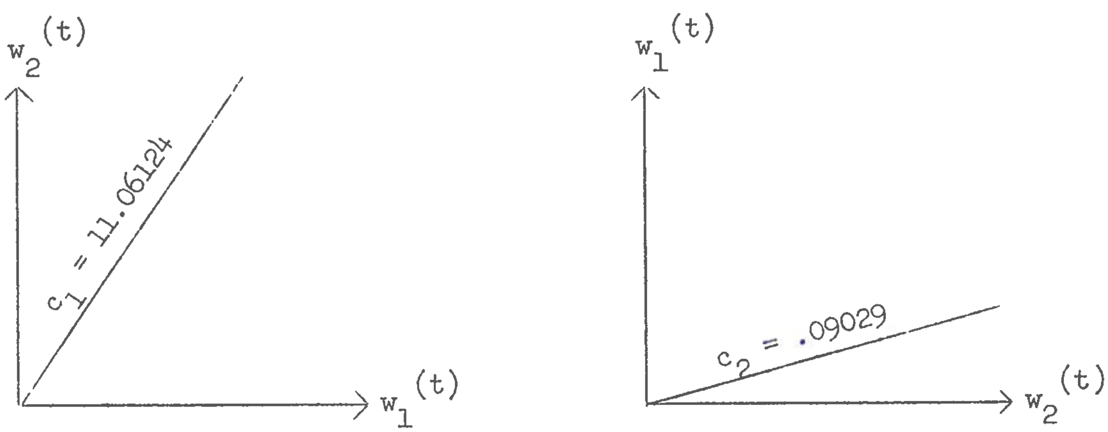
Figure 7 illustrates the decomposition for the 1955-1960 population growth operator that appears in Figure 2C. Using the 1955-1960 time series,

Fig. 7.--THE POPULATION MODEL AS A SERIES OF

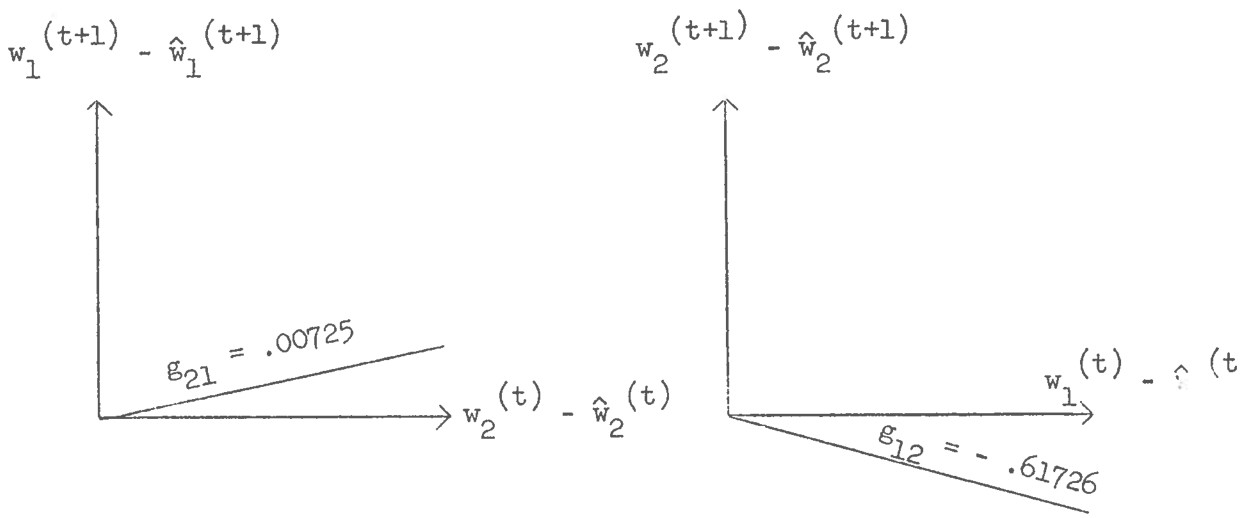
SIMPLE LINEAR REGRESSIONS



A. Interperiod Growth Path



B. Intraperiod Allocations



C. Migration Rate

we obtain[9]

$$w_1^{(t+1)} = [1.04022 - .00725(11.06124)]w_1^{(t)} + .00725w_2^{(t)}$$

$$= .9603w_1^{(t)} + .0073w_2^{(t)} \quad ,$$

and

$$w_2^{(t+1)} = [1.01426 + .61726(.09029)]w_2^{(t)} - .6173w_1^{(t)}$$

$$= 1.0700w_2^{(t)} - .6173w_1^{(t)} \quad .$$

<u>Analysis of the negative migration rate</u>

Consider the negative migration rate $g_{12} = - .6169$. The least squares estimator for the slope of a line going through the origin of an (X, Y) co-ordinate system is

$$b = \frac{\Sigma X_i Y_i}{\Sigma X_i^2} \quad . \tag{3.8}$$

Hence, the least squares estimator for g_{12} is

$$g_{12} = \frac{\Sigma[(w_1^{(t)} - c_2w_2^{(t)})(w_2^{(t+1)} - a_2w_2^{(t)})]}{\Sigma[w_1^{(t)} - c_2w_2^{(t)}]^2} \quad . \tag{3.9}$$

Since the denominator is never negative, g_{12} is negative only when the numerator in (3.9) is negative. This will occur when there is a preponderance of observations for which either

$$w_2^{(t+1)} < a_2w_2^{(t)} \ \underline{\text{and}} \ w_1^{(t)} > c_2w_2^{(t)} \quad , \tag{3.10}$$

or

$$w_2^{(t+1)} > a_2w_2^{(t)} \ \underline{\text{and}} \ w_1^{(t)} < c_2w_2^{(t)} \quad . \tag{3.11}$$

The conditions which will lead to a positive g_{ij} may be illustrated graphically. For example, consider g_{12} once again. For g_{12} to be positive we must have either the situation illustrated by Figure 8A or Figure 8B. That is, the data points in the $(w_2^{(t+1)}, w_2^{(t)})$ coordinate system must follow the same pattern as those in the $(w_1^{(t)}, w_2^{(t)})$ coordinate system. Either both must exhibit a declining rate of growth or both must follow an increasing rate of growth (for example, the dotted lines in Figure 8). This will ensure that the conditions specified by (3.10) and (3.11) are not satisfied and, hence, that g_{12} is positive.

The necessary conditions for diagonal elements to be greater than unity may be established in an analogous manner. However, the problem is a bit more difficult and is beyond the scope of this paper.

Smoothing the data

The following smoothing operation is suggested by the conditions illustrated in Figure 8. We begin by fitting quadratic functions (forced through the origin) to the data. We have, then, the following four equations:

$$w_1^{(t+1)} = \alpha_1 w_1^{(t)} + \beta_1 [w_1^{(t)}]^2 \quad , \tag{3.12}$$

$$w_2^{(t)} = \alpha_2 w_1^{(t)} + \beta_2 [w_1^{(t)}]^2 \quad , \tag{3.13}$$

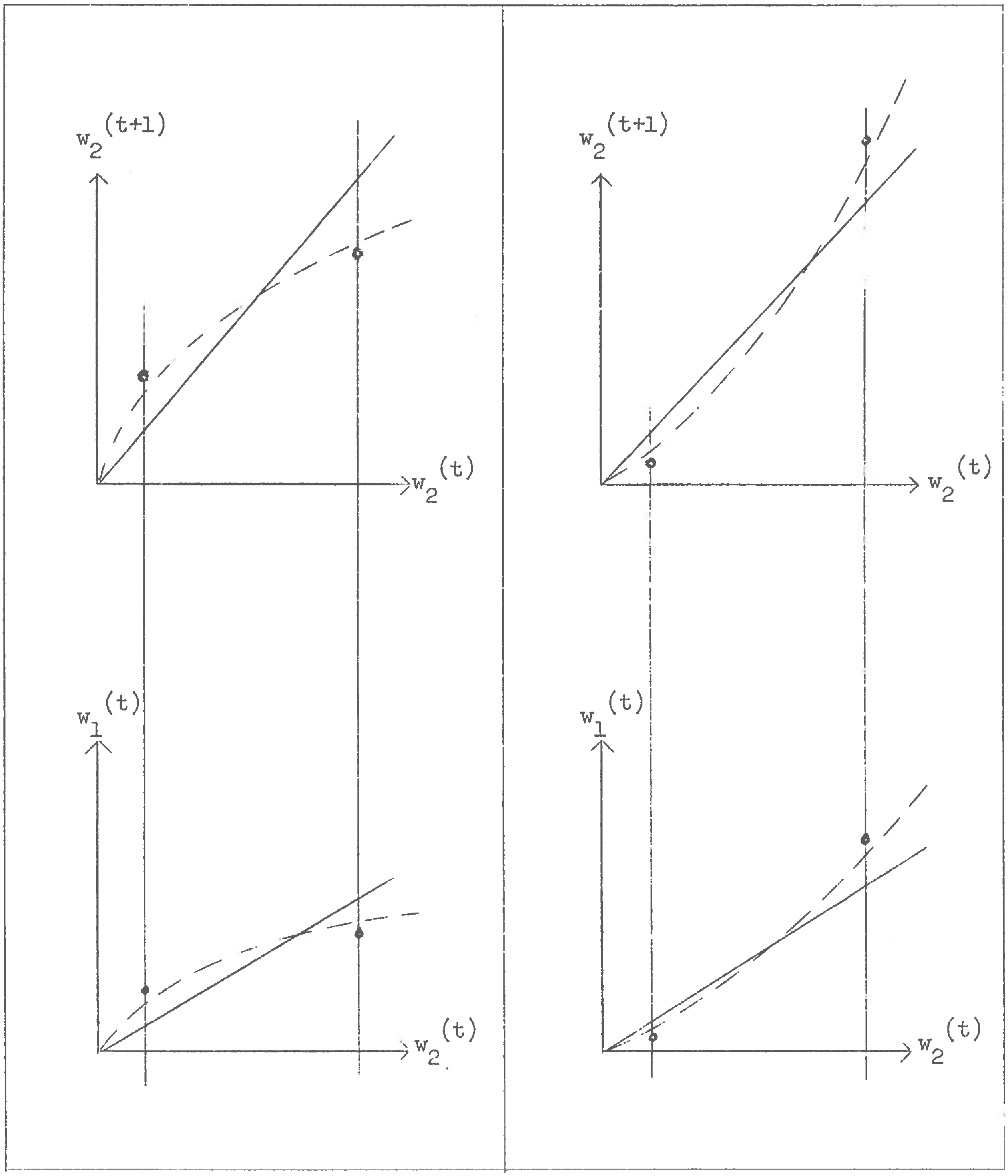
$$w_2^{(t+1)} = \alpha_3 w_2^{(t)} + \beta_3 [w_2^{(t)}]^2 \quad , \tag{3.14}$$

$$w_1^{(t)} = \alpha_4 w_2^{(t)} + \beta_4 [w_2^{(t)}]^2 \quad . \tag{3.15}$$

Next, we examine the fitted equations to determine whether (3.12) and (3.13) have the same form and then repeat the examination for (3.14) and

FIG. 8.--CONDITIONS FOR A POSITIVE MIGRATION RATE FROM

REGION 1 TO REGION 2



A. Decreasing Rate of Growth B. Increasing Rate of Growth

(3.15). For example, returning to our analysis of the 1955-1960 time series we find,

$$(3.16) \qquad w_1^{(t+1)} = 1.09029w_1^{(t)} - .0000035[w_1^{(t)}]^2$$

$$(3.17) \qquad w_2^{(t)} = 18.01309w_1^{(t)} - .0004878[w_1^{(t)}]^2$$

$$(3.18) \qquad w_2^{(t+1)} = 1.10608w_2^{(t)} - .0000006[w_2^{(t)}]^2$$

$$(3.19) \qquad w_1^{(t)} = -\ .05800w_2^{(t)} + .0000009[w_2^{(t)}]^2$$

Clearly (3.16) and (3.17) are of the same form. Equations (3.18) and (3.19), however, are not. This is to be expected since g_{12} was negative. We select (3.18) for smoothing and transform it from a decreasing rate-of-growth function to an increasing rate-of-growth function by ensuring that both coefficients in (3.14) remain strictly positive. Our procedure from here on is quite arbitrary and is only meant to be illustrative.

We select for the coefficient α_3, in (3.14), the lowest $w_2^{(t+1)}/w_2^{(t)}$ ratio in our time series. Thus we have

$$\alpha_3 = \frac{163284}{161637} = 1.01019 \quad .$$

To derive β_3, we take the first two data points and solve for the β_3 that will reproduce the second data point when the first is entered into Equation (3.14), i.e.,

$$154507 = 1.01019[152065] + \beta_3[152065]^2 \ ,$$

and

$$\beta_3 = .00000004 \quad .$$

Thus we replace (3.18) with the "smoothed" fitted equation

$$(3.20) \qquad w_2^{(t+1)} = 1.01019 w_2^{(t)} + .00000004[w_2^{(t)}]^2 \quad ,$$

and derive the smoothed time series for the rest of the United States that appears in Table 5A. The series presented in 5B and 5C result from setting $\beta_3 = .00000002$ and $\beta_3 = .00000003$, respectively.

Results. Figure 9 presents the ULS and MAD estimates, of the 1955-1960 population growth operator, that correspond to the three alternate smoothed data series in Table 5. A quick glance reveals the sensitivity of g_{12} to a unit change in the seventh decimal place in the value of β_3. This underscores the importance of establishing a more rational method for smoothing the data points than is presented here.

4. Migration Operator Estimators

Up to this point we have assumed that, for our interregional system, data on births and deaths are totally unavailable. This is a level of data availability that exists in many less developed parts of the world. However, it hardly typifies conditions in the United States. Data on births and deaths are regularly collected for most subdivisions of the nation down to the county level. Where such information exists, it should be incorporated into the estimation process and used to reduce the problem to one of estimating migration alone. This removes variation that may be attributable to fluctuation in birth and death rates. It also allows us to introduce a side constraint in the estimation process: namely, that the rows of the migration operator sum to unity. As a result, only $m^2 - m$ parameters need to be estimated. The diagonal elements of the operator are strictly determined by the side constraint.

TABLE 5.--ESTIMATED TOTAL POPULATION OF CALIFORNIA AND "SMOOTHED" TOTAL POPULATION OF THE REST OF THE UNITED STATES, 1955-1960

Year (July 1)	California (000's)	Rest of the United States* (000's)			
		Observed	A	B	C
1955	13004	152065	152065	152065	152065
1956	13581	154507	154516	154077	154308
1957	14177	157010	157022	156122	156595
1958	14741	159408	159584	158200	158926
1959	15288	161637	162203	160313	161303
1960	15863	163284	164882	162461	163727

* Series A: $\beta_3 = .00000004$

Series B: $\beta_3 = .00000002$

Series C: $\beta_3 = .00000003$

FIG. 9.--UNRESTRICTED LEAST SQUARES AND MINIMUM ABSOLUTE DEVIATIONS ESTIMATES: SMOOTHED DATA, 1955-1960 TIME SERIES

A. Series A:

ULS Estimate

$$\begin{bmatrix} .9573 & .0075 \\ .0450 & 1.0123 \end{bmatrix}$$

$R^2 = 1.0000$

MAD Estimate

$$\begin{bmatrix} .9575 & .0074 \\ .0431 & 1.0124 \end{bmatrix}$$

$Z = .0062$

B. Series B:

ULS Estimate

$$\begin{bmatrix} .9663 & .0067 \\ .0170 & 1.0118 \end{bmatrix}$$

$R^2 = 1.0000$

MAD Estimate

$$\begin{bmatrix} .9656 & .0067 \\ .0167 & 1.0118 \end{bmatrix}$$

$Z = .0060$

C. Series C:

ULS Estimate

$$\begin{bmatrix} .9618 & .0071 \\ .0298 & 1.0122 \end{bmatrix}$$

$R^2 = 1.0000$

MAD Estimate

$$\begin{bmatrix} .9616 & .0071 \\ .0286 & 1.0123 \end{bmatrix}$$

$Z = .0061$

To transform the population distributions of our two-region system into a form suitable for estimating the migration operator, we need to remove the effects of natural increase from the data. This can be done by subtracting the time series data on natural increase presented in Table 6 from the population data appearing in Table 1 and then normalizing the results so that the national population after the application of the migration operator (i.e., the elements of the vector $\underline{y}$) remains equal to the national population as represented by the operand (i.e., the row elements of X). We thereby ensure that during the migration operation the population remains constant. The results of such a data transformation for our two-region system appear in Table 7.

<u>The ULS and MAD Estimators</u>

Figure 10 illustrates the ULS estimation process for the migration operator. Figure 11 presents the three migration operators as estimated by the ULS and MAD estimators, respectively.

<u>Results</u>. The results are considerably improved. The problem of negative elements does not seem to arise and the off-diagonal elements, though still much too large in two instances, now take on more plausible values. Finally, the problem of diagonal elements being less than unity, of course, no longer is present.

TABLE 6.--NATURAL INCREASE IN CALIFORNIA AND THE REST OF THE UNITED STATES, 1950-1960*

Interval (July 1)	California (000's)	Rest of the United States (000's)	National Total (000's)
1950-1951	151	2136	2287
1951-1952	165	2184	2349
1952-1953	180	2241	2421
1953-1954	192	2366	2558
1954-1955	198	2416	2614
1955-1956	207	2390	2597
1956-1957	220	2510	2730
1957-1958	225	2405	2630
1958-1959	227	2424	2651
1959-1960	234	2349	2583

Source: Financial and Population Research Section, California Population 1964 (Sacramento, Calif.: Department of Finance, 1964), p. 9 and U.S. Department of Commerce, Bureau of the Census, Current Population Reports, Series P-25, July 1966. January 1 data were linearly interpolated to provide July 1 data.

TABLE 7.--TOTAL POPULATION WITHOUT NATURAL INCREASE OF CALIFORNIA AND THE REST OF THE UNITED STATES, 1950-1960*

Year (July 1)	California (000's)	Rest of the United States (000's)
1951	10991	140876
1952	11468	142514
1953	11910	144483
1954	12296	146660
1955	12761	149123
1956	13323	151540
1957	13926	154162
1958	14488	156699
1959	15050	159099
1960	15571	161353

* Data were derived by subtracting the natural increases in Table 6 from the populations in Table 1 and then normalizing the results so that the national population at time t + 1 added up to the national population at time t, e.g., for 1951 we have 10979 + 140716 = 151695 and, since the national population in 1950 was 151868, we normalize to find 10991 + 140876 = 151868.

FIG. 10.--UNRESTRICTED LEAST SQUARES ESTIMATE OF THE ANNUAL MIGRATION FLOW OPERATOR FOR CALIFORNIA AND THE REST OF THE UNITED STATES: 1950-1960 TOTAL POPULATION WITHOUT NATURAL INCREASE TIME SERIES

Estimator: $\underline{g} = (X'X)^{-1}X'\underline{y}$

Data:

$$X = \left[\begin{array}{cc|cc} 10643 & 141225 & 0 & 0 \\ 11130 & 142852 & 0 & 0 \\ 11638 & 144755 & 0 & 0 \\ 12101 & 146855 & 0 & 0 \\ 12517 & 149367 & 0 & 0 \\ 13004 & 152065 & 0 & 0 \\ 13581 & 154507 & 0 & 0 \\ 14177 & 157010 & 0 & 0 \\ 14741 & 159408 & 0 & 0 \\ 15288 & 161637 & 0 & 0 \\ \hline 0 & 0 & 10643 & 141225 \\ 0 & 0 & 11130 & 142852 \\ 0 & 0 & 11638 & 144755 \\ 0 & 0 & 12101 & 146855 \\ 0 & 0 & 12517 & 149367 \\ 0 & 0 & 13004 & 152065 \\ 0 & 0 & 13581 & 154507 \\ 0 & 0 & 14177 & 157010 \\ 0 & 0 & 14741 & 159408 \\ 0 & 0 & 15288 & 161637 \end{array}\right] \qquad \underline{y} = \left[\begin{array}{c} 10991 \\ 11468 \\ 11910 \\ 12296 \\ 12761 \\ 13323 \\ 13926 \\ 14488 \\ 15050 \\ 15571 \\ \hline 140876 \\ 142514 \\ 144483 \\ 146660 \\ 149123 \\ 151540 \\ 154162 \\ 156699 \\ 159099 \\ 161353 \end{array}\right]$$

Estimate:

$$\hat{\underline{g}} = \left[\begin{array}{c} .9832 \\ .0034 \\ \hline .0171 \\ .9964 \end{array}\right]$$

FIG. 11.--UNRESTRICTED LEAST SQUARES AND MINIMUM ABSOLUTE DEVIATIONS ESTIMATES: TOTAL POPULATION WITHOUT NATURAL INCREASE

A. 1950-1960 Time Series:

Unweighted

$$\begin{bmatrix} .9832 & .0034 \\ .0171 & .9964 \end{bmatrix}$$

$R^2 = 1.0000$

Unweighted

$$\begin{bmatrix} .9689 & .0048 \\ .0311 & .9952 \end{bmatrix}$$

$Z = .0852$

B. 1950-1955 Time Series:

Unweighted

$$\begin{bmatrix} .8711 & .0122 \\ .1295 & .9877 \end{bmatrix}$$

$R^2 = 1.0000$

Unweighted

$$\begin{bmatrix} .8824 & .0113 \\ .1176 & .9887 \end{bmatrix}$$

$Z = .0196$

C. 1955-1960 Time Series:

Unweighted

$$\begin{bmatrix} .9562 & .0059 \\ .1614 & .9832 \end{bmatrix}$$

$R^2 = 1.0000$

Unweighted

$$\begin{bmatrix} .9288 & .0085 \\ .0712 & .9915 \end{bmatrix}$$

$Z = .0243$

5. Conclusion

In this paper we have demonstrated that it is possible to estimate the structure of interregional population growth and migration solely on the basis of historical data on interregional population distributions. The application of the method to data generated by a changing growth operator, however, provides unsatisfactory results. Recourse to data "smoothing," therefore, appears to be necessary in instances where the operator undergoes considerable changes over time. Furthermore, data permitting, estimation of migration operators rather than population growth operators seems to lead to improved results.

6. References

1. Draper, N. R. and H. Smith, Applied Regression Analysis (New York: John Wiley, 1966).
2. Fisher, W. D., "A Note on Curve Fitting with Minimum Deviations by Linear Programming," Journal American Statistical Association, vol. 56 (1961), pp. 359-362.
3. Judge, G. G., and T. Takayama, "Inequality Restrictions in Regression Analysis," Journal American Statistical Association, vol. 61 (1966), pp. 166-181.
4. Lee, T. C., G. G. Judge, and T. Takayama, "On Estimating the Transition Probabilities of a Markov Process," Journal of Farm Economics, vol. 43 (1965), pp. 742-762.
5. Madanksy, A., "Least Squares Estimation in Finite Markov Process," Psychometrika, vol. 25 (1959), pp. 137-144.
6. Miller, George A., "Finite Markov Processes in Psychology," Psychometrika, vol. 17 (1952), pp. 149-167.
7. Rogers, Andrei, "The Multiregional Matrix Growth Operator and the Stable Interregional Age Structure," Demography, vol. 3, no. 2 (1966).
8. Telser, Lester G., "Least Squares Estimates of Transition Probabilities," in Measurement in Economics, F. Christ (ed.) (Palo Alto, California: Stanford University Press, 1963), pp. 272-292.
9. Theil, Henry and Guido Rey, "A Quadratic Programming Approach to the Estimation of Transition Probabilities," Management Science, vol. 12 (1966), pp. 714-721.
10. Wagner, Harvey M., "Linear Programming Techniques for Regression Analysis," Journal American Statistical Association, vol. 54 (1959), pp. 206-212.

FOOTNOTES

1 The author wishes to acknowledge the invaluable services rendered by his research assistant, Edward Kampmann, who performed most of the data processing chores generated by this paper. Thanks also are due to the Center for Planning and Development Research and to the Institute of Social Sciences, both of the University of California, for financial support.

2 Andrei Rogers, "The Multiregional Matrix Growth Operator and the Stable Interregional Age Structure," Demography, vol. 3, no. 2 (1966).

3 Assuming, of course, that the following matrix is non-singular:

$$\begin{bmatrix} w_1^{(t)} & w_2^{(t)} \\ w_1^{(t+1)} & w_2^{(t+1)} \end{bmatrix} .$$

4 A. Madanksy, "Least Squares Estimation in Finite Markov Processes," Psychometrika, vol. 25 (1959), pp. 137-144; George A. Miller, "Finite Markov Processes in Psychology," Psychometrika, vol. 17 (1952), pp. 149-167; Lester G. Telser, "Least Squares Estimates of Transition Probabilities," Measurement in Economics, F. Christ, ed. (Palo Alto, California: Stanford University Press, 1963), pp. 272-292.

5 W. D. Fisher, "A Note on Curve Fitting with Minimum Deviations by Linear Programming," Journal American Statistical Association, vol. 56 (1961), pp. 359-362; T. C. Lee, G. G. Judge, and T. Takayama, "On Estimating the Transition Probabilities of a Markov Process," Journal of Farm Economics, vol. 43 (1965), pp. 742-762.

6 G. G. Judge and T. Takayama, "Inequality Restrictions in Regression Analysis," Journal American Statistical Association, vol. 61 (1966), pp. 166-181; Henry Theil and Guido Rey, "A Quadratic Programming

Approach to the Estimation of Transition Probabilities," Management Science, vol. 12 (1966), pp. 714-721.

7 If a particular u_i and v_i were both strictly positive, with $u_i > v_i$ say, the solution could not be optimal since it would be possible to find a smaller value for S in Equation (2.14) by replacing u_i by $u_i - v_i$ and setting the original v_i equal to zero.

8 For a lucid exposition see: N. R. Draper and H. Smith, Applied Regression Analysis (New York: John Wiley, 1966), pp. 107-115.

9 Discrepancies are due to rounding.

10 Notice that g_{12} is zero if $w_2^{(t+1)} = a_2 w_2^{(t)}$ and is undefined if $w_1^{(t)} = c_2 w_2^{(t)}$.