WP#66

THE AGGREGATION PROBLEM IN DEMOGRAPHY

By

Andrei Rogers
Department of City and Regional Planning
Working Paper No. 66

October 1967

THE AGGREGATION PROBLEM IN DEMOGRAPHY[1]

1. Introduction

Most demographers ultimately are confronted with the need to aggregate data over the three fundamental data units of demographic systems: space, people, and time. The need for aggregation may arise because of data limitations or budget constraints. If the latter is the binding consideration, the population analyst must identify a scheme for consolidation that is in some sense optimal.

In this paper we seek to establish more rational decision rules for aggregating demographic data. We draw on the related literature on the aggregation problem in input-output economics to answer the formal question: Under what conditions can we consolidate space, people, and time in interregional population models without deviating from the population estimates that would have been derived in the absence of aggregation?[2] We begin by considering the effects of spatial aggregation, i.e., consolidation of regions. Next, we focus on the consequences of cohort aggregation, e.g., consolidation of color or sex differentiated cohorts. We then analyze the influence of temporal consolidation, i.e., consolidation of time intervals, and follow this with another look at spatial aggregation.

Finally, we investigate the joint impact of all three forms of aggregation and attempt to establish guides for rational consolidation. An illustrative numerical example concludes the paper.

2. Spatial Aggregation

Imagine an interregional demographic system consisting of three regions: i, j, and k. To simplify the analysis further, let us consider only the total population of these regions. Over time, the growth of such an interregional population system may be expressed by the following equations:

$$
\begin{aligned}
w_i(t+1) &= w_i(t) + b_i(t) - d_i(t) - m_{ij}(t) - m_{ik}(t) + m_{ji}(t) + m_{ki}(t) \\
w_j(t+1) &= w_j(t) + b_j(t) - d_j(t) - m_{ji}(t) - m_{jk}(t) + m_{ij}(t) + m_{kj}(t) \qquad (1) \\
w_k(t+1) &= w_k(t) + b_k(t) - d_k(t) - m_{ki}(t) - m_{kj}(t) + m_{ik}(t) + m_{jk}(t)
\end{aligned}
$$

where

$w_i(t)$ = total population of the ith region at time t;

$b_i(t)$ = total number of births in the ith region during the time interval $(t, t + 1)$;

$d_i(t)$ = total number of deaths in the ith region during the time interval $(t, t + 1)$; and

$m_{ij}(t)$ = total number of migrants from the ith region to the jth region during the time interval $(t, t + 1)$.

Expressing (1) in matrix form, we have:

$$\begin{bmatrix} w_i(t+1) \\ \\ w_j(t+1) \\ \\ w_k(t+1) \end{bmatrix} = \begin{bmatrix} \frac{m_{ii}(t)}{w_i(t)} & \frac{m_{ji}(t)}{w_j(t)} & \frac{m_{ki}(t)}{w_k(t)} \\ \\ \frac{m_{ij}(t)}{w_i(t)} & \frac{m_{jj}(t)}{w_j(t)} & \frac{m_{kj}(t)}{w_k(t)} \\ \\ \frac{m_{ik}(t)}{w_i(t)} & \frac{m_{jk}(t)}{w_j(t)} & \frac{m_{kk}(t)}{w_k(t)} \end{bmatrix} \begin{bmatrix} w_i(t) \\ \\ w_j(t) \\ \\ w_k(t) \end{bmatrix} \tag{2}$$

where $m_{ii}(t) = w_i(t) + b_i(t) - d_i(t) - m_{ij}(t) - m_{ik}(t)$,

or more compactly:

$$\underline{w}(t + 1) = G\underline{w}(t) \tag{3}$$

where $g_{ij} = \frac{m_{ij}(t)}{w_i(t)}$.

Consider, now, the consequences of consolidating regions i and j. The consolidated form of (2) is:

$$\begin{bmatrix} w_{(i+j)}(t+1) \\ \\ w_k(t+1) \end{bmatrix} = \begin{bmatrix} \frac{m_{(i+j)(i+j)}(t)}{w_{(i+j)}(t)} & \frac{m_{k(i+j)}(t)}{w_k(t)} \\ \\ \frac{m_{(i+j)k}(t)}{w_{(i+j)}(t)} & \frac{m_{kk}(t)}{w_k(t)} \end{bmatrix} \begin{bmatrix} w_{(i+j)}(t) \\ \\ w_k(t) \end{bmatrix}$$

$$= \begin{bmatrix} g_{(i+j)(i+j)} & g_{k(i+j)} \\ \\ g_{(i+j)k} & g_{kk} \end{bmatrix} \begin{bmatrix} w_{(i+j)}(t) \\ \\ w_k(t) \end{bmatrix} \tag{4}$$

where

$$w_{(i+j)}(t) = w_i(t) + w_j(t)$$
$$m_{(i+j)(i+j)}(t) = m_{ii}(t) + m_{ij}(t) + m_{jj}(t) + m_{ji}(t)$$
$$m_{(i+j)k}(t) = m_{ik}(t) + m_{jk}(t)$$
$$m_{k(i+j)}(t) = m_{ki}(t) + m_{kj}(t) \quad .$$

Thus the consolidated form of (3) may be expressed as:

$$\underline{\hat{w}}(t+1) = G\underline{\hat{w}}(t) \tag{5}$$

We now may analyze the conditions which will lead to perfect aggregation by answering the following question: Under what conditions can regions i and j be consolidated without introducing an error into the population forecast?[3] The elements of G for the consolidated region (i + j) may be defined as follows:

$$g_{(i+j)k} = \frac{m_{(i+j)k}(t)}{w_{(i+j)}(t)} = \frac{m_{ik}(t) + m_{jk}(t)}{w_i(t) + w_j(t)} = \frac{g_{ik}w_i(t) + g_{jk}\quad w_j(t)}{w_i(t) + w_j(t)}$$

$$= \left(\frac{w_i(t)}{w_i(t) + w_j(t)}\right) g_{ik} + \left(\frac{w_j(t)}{w_i(t) + w_j(t)}\right) g_{jk}$$

$$= dg_{ik} + (1-d)g_{jk} \quad . \tag{6}$$

$$g_{(i+j)(i+j)} = \frac{m_{(i+j)(i+j)}(t)}{w_{(i+j)}(t)} = \frac{m_{(ii)}(t) + m_{(ij)}(t) + m_{(jj)}(t) + m_{(ji)}(t)}{w_{(i)}(t) + w_{(j)}(t)}$$

$$= \frac{g_{ii}w_i(t) + g_{ij}w_i(t) + g_{jj}w_j(t) + g_{ji}w_j(t)}{w_i(t) + w_j(t)}$$

$$= \left(\frac{w_i(t)}{w_i(t) + w_j(t)}\right) (g_{ii} + g_{ij}) + \left(\frac{w_j(t)}{w_i(t) + w_j(t)}\right) (g_{jj} + g_{ji})$$

$$= d(g_{ii} + g_{ij}) + (1 - d)(g_{jj} + g_{ji}) \quad . \tag{7}$$

$$g_{k(i+j)} = \frac{m_{k(i+j)}(t)}{w_k(t)} = \frac{m_{ki}(t) + m_{kj}(t)}{w_k(t)} = \frac{g_{ki}w_k(t) + g_{kj}w_k(t)}{w_k(t)}$$

$$= g_{ki} + g_{kj} \quad . \tag{8}$$

The element g_{kk} remains unchanged.

If all the elements of $\hat{G}$ relating to the consolidated region are unaffected by changes in the regional populations over time, no errors will result from the aggregation of regions i and j. Equations (6), (7), and (8) show that there are two conditions for which this will be true:

Consolidation Rule I:

If $g_{ik} = g_{jk}$ and $g_{ii} + g_{ij} = g_{jj} + g_{ji}$, changes in the weight d, resulting from changes in the proportional relationship between w_i and w_j over time, will not affect the consolidated coefficients $g_{(i+j)k}$ and $g_{(i+j)(i+j)}$.

Consolidation Rule II:

If w_i and w_j retain their proportional relationship over time, i.e., if d remains constant, changes in g_{ik}, g_{jk}, g_{ii}, g_{ij}, g_{jj}, g_{ji}, will not affect the consolidated coefficients $g_{(i+j)k}$ and $g_{(i+j)(i+j)}$.

In conclusion, we have identified two rules for consolidation which, if strictly met, will produce perfect aggregation. The first states that two regions whose rates of natural increase in the absence of intermigration and whose rates of outmigration to the unconsolidated region (or regions) are equal may be consolidated. The second asserts that two regions whose populations retain a constant proportional relationship to each other may be consolidated. In practice it is very unlikely that either condition will hold. Thus, the population analyst typically is faced with a choice of combinations of regions which meet these consolidation rules to varying degrees. In such instances, the indirect influence of the consolidation, i.e., the impact of the aggregation on the population forecasts of unconsolidated regions, must also be considered, since the criterion with which each aggregation scheme is evaluated is the average error for all regional population totals.[4]

Before considering the consequences of cohort aggregation, let us summarize our examination of the effects of spatial aggregation in matrix form. This will afford us a compact notation for our subsequent analyses and thereby allow us to proceed more rapidly.

Aggregation and disaggregation operations may be conveniently expressed by matrix multiplication. To consolidate a three dimensional vector, say $\underline{v}$, into a two dimensional vector, say $\hat{\underline{v}}$, whose first element is the sum of the first two elements of $\underline{v}$, we define the <u>aggregation operator</u>:

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{9}$$

such that

$$\hat{\underline{v}} = A\underline{v} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} v_1 + v_2 \\ v_3 \end{bmatrix} .$$

To reverse this consolidation, we may define the <u>disaggregation operator</u>:

$$D = \begin{bmatrix} d & 0 \\ 1 - d & 0 \\ 0 & 1 \end{bmatrix} \tag{10}$$

where $d = \dfrac{v_1}{v_1 + v_2}$.

Thus

$$\underline{v} = D\hat{\underline{v}} = \begin{bmatrix} \dfrac{v_1}{v_1 + v_2} & 0 \\ \dfrac{v_2}{v_1 + v_2} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 + v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} .$$

Note that $AD = I$.

Returning to our three-region population system, we recall Equations (3) and (5):

$$\underline{w}(t+1) = G\underline{w}(t) \tag{3'}$$

$$\hat{\underline{w}}(t+1) = \hat{G}\hat{\underline{w}}(t) \tag{5'}$$

and note that

$$\hat{\underline{w}}(t) = A\underline{w}(t) \tag{11}$$

$$\underline{w}(t) = D\hat{\underline{w}}(t) \quad . \tag{12}$$

Hence, we have:

$$\underline{w}(t+1) = A\underline{w}(t+1) = AG\underline{w}(t) = AGD\underline{w} = G\ddot{\underline{w}}(t)$$

$$\hat{G} = AGD \quad . \tag{13}$$

We now may rederive Consolidation Rules I and II by means of matrix algebra. First we note that the "errorless" population forecast for (t + 1) is $\underline{w}(t+1) = G\underline{w}(t)$. Thus the associated errorless forecast for the consolidated case is:

$$\hat{\underline{w}}(t+1) = A\underline{w}(t+1) = AG\underline{w}(t) \quad . \tag{14}$$

The "consolidated" forecast, however, is:

$$\hat{\underline{w}}(t+1) = \hat{G}\hat{\underline{w}}(t) = \hat{G}A\underline{w}(t) \quad . \tag{15}$$

Thus the two forecasts will be equal if:[5]

<u>Consolidation Rule I</u>: $$AG = \hat{G}A \tag{16}$$

To derive the second consolidation rule, we note that the errorless population forecast for the unconsolidated case is: [6]

$$\underline{w}(t+1) = G\underline{w}(t) = GD_o\hat{\underline{w}}(t) \quad . \tag{17}$$

The corresponding "consolidated" forecast is:

$$\underline{w}(t+1) = D_1\hat{\underline{w}}(t+1) = D_1\hat{G}\hat{\underline{w}}(t) \quad . \tag{18}$$

Therefore these two forecasts will be equal if;

<u>Consolidation Rule II</u>: $$GD_o = D_1\hat{G} \quad . \tag{19}$$

Simple substitution and matrix multiplication may be used to show that Consolidation Rules I and II as expressed by (16) and (19), respectively, are equivalent to those stated earlier in simple algebraic form. The matrix expressions, however, are simpler to apply to situations involving interregional systems with several multiregional consolidations. Moreover, as we shall see below, with appropriate redefinition of the aggregation and disaggregation operators, the consolidation rules of (16) and (19) also apply to cohort and temporal aggregation.

3. Cohort Aggregation

Imagine a cohort differentiated population that is closed to migration. Assume that there are only two cohorts and that these are differentiated by color, say white and nonwhite (the same results may be derived for sex differentiated cohorts). Let each cohort be denoted by a four dimensional column vector $\underline{w}$, with each element denoting the population in a fifteen-year age group. Thus we have only the following four age groups: 0-14, 15-29, 30-44, and 45+. The growth of such a population system may be expressed by the matrix expression: [7]

$$\underline{w}(t+1) = G\underline{w}(t) \tag{20}$$

or

$$\left[\begin{array}{c} w_{m1}(t+1) \\ w_{m2}(t+1) \\ w_{m3}(t+1) \\ w_{m4}(t+1) \\ \hline w_{n1}(t+1) \\ w_{n2}(t+1) \\ w_{n3}(t+1) \\ w_{n4}(t+1) \end{array}\right] = \left[\begin{array}{cccc|cccc} 0 & b_{m2} & b_{m3} & 0 & 0 & 0 & 0 & 0 \\ s_{m1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & s_{m2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & s_{m3} & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & b_{n2} & b_{n3} & 0 \\ 0 & 0 & 0 & 0 & s_{n1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & s_{n2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & s_{n3} & 0 \end{array}\right] \left[\begin{array}{c} w_{m1}(t) \\ w_{m2}(t) \\ w_{m3}(t) \\ w_{m4}(t) \\ \hline w_{n1}(t) \\ w_{n2}(t) \\ w_{n3}(t) \\ w_{n4}(t) \end{array}\right] \tag{21}$$

where

$w_{mi}(t)$, $w_{ni}(t)$ = respectively, the white and nonwhite populations in the ith age group at time t;

b_{mi}, b_{ni} = respectively, the number of white and nonwhite surviving births, per person, attributable to the ith age group, during the time period (t, t + 1); and

s_{mi}, s_{ni} = respectively, the white and nonwhite survivorship proportions, of the ith age group, during the time period (t, t + 1).

Let us now aggregate the white and nonwhite age distributions. We define the <u>aggregation operator</u>:

$$A = \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \tag{22}$$

such that:

$$\underline{\hat{w}}(t) = A\underline{w}(t) = \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \left[\begin{array}{c} w_{m1}(t) \\ w_{m2}(t) \\ w_{m3}(t) \\ w_{m4}(t) \\ \hdashline w_{n1}(t) \\ w_{n2}(t) \\ w_{n3}(t) \\ w_{n4}(t) \end{array}\right] = \left[\begin{array}{c} w_{(m+n)1}(t) \\ w_{(m+n)2}(t) \\ w_{(m+n)3}(t) \\ w_{(m+n)4}(t) \end{array}\right] \tag{23}$$

where $w_{(m+n)i}(t) = w_{mi}(t) + w_{ni}(t)$.

Next, we define the <u>disaggregation</u> <u>operator</u>:

$$D = \left[\begin{array}{cccc} d_1 & 0 & 0 & 0 \\ 0 & d_2 & 0 & 0 \\ 0 & 0 & d_3 & 0 \\ 0 & 0 & 0 & d_4 \\ \hdashline 1-d_1 & 0 & 0 & 0 \\ 0 & 1-d_2 & 0 & 0 \\ 0 & 0 & 1-d_3 & 0 \\ 0 & 0 & 0 & 1-d_4 \end{array}\right] \tag{24}$$

where $d_i = \dfrac{w_{mi}(t)}{w_{mi}(t) + w_{ni}(t)}$.

Thus

$$\underline{w}(t) = D\hat{\underline{w}}(t) = \left[\begin{array}{cccc} d_1 & 0 & 0 & 0 \\ 0 & d_2 & 0 & 0 \\ 0 & 0 & d_3 & 0 \\ 0 & 0 & 0 & d_4 \\ \hdashline 1-d_1 & 0 & 0 & 0 \\ 0 & 1-d_2 & 0 & 0 \\ 0 & 0 & 1-d_3 & 0 \\ 0 & 0 & 0 & 1-d_4 \end{array}\right] \left[\begin{array}{c} w_{(m+n)1}(t) \\ w_{(m+n)2}(t) \\ w_{(m+n)3}(t) \\ w_{(m+n)4}(t) \end{array}\right] = \left[\begin{array}{c} w_{m1}(t) \\ w_{m2}(t) \\ w_{m3}(t) \\ w_{m4}(t) \\ \hdashline w_{n1}(t) \\ w_{n2}(t) \\ w_{n3}(t) \\ w_{n4}(t) \end{array}\right] \tag{25}$$

Having defined the appropriate <u>aggregation</u> and <u>disaggregation</u> operators, we now may draw on Equation (13) to express the consolidated growth matrix $\hat{G}$ as a function of G:

$$\hat{G} = AGD = \begin{bmatrix} 0 & \hat{b}_2 & \hat{b}_3 & 0 \\ \hat{s}_1 & 0 & 0 & 0 \\ 0 & \hat{s}_2 & 0 & 0 \\ 0 & 0 & \hat{s}_3 & 0 \end{bmatrix} \tag{26}$$

where

$$\hat{b}_i = d_i b_{mi} + (1-d_i) b_{ni} \tag{27}$$

and

$$\hat{s}_i = d_i s_{mi} + (1-d_i) s_{ni} \quad . \tag{28}$$

Equations (27) and (28) exhibit a striking similarity to Equation (6). And the argument used to derive Consolidation Rules I and II in matrix form may be adopted, without change, to establish the following consolidation rules for cohort aggregation:

<u>Consolidation Rule III</u>: $$AG = \hat{G}A \quad . \tag{29}$$

If $b_{mi} = b_{ni}$ and $s_{mi} = s_{ni}$ for all $i (i = 1, 2, 3)$, changes in the weights d_i, resulting from changes in the proportional relationship between w_{mi} and w_{ni} over time, will not affect the consolidated coefficients $\hat{b}_i$ and $\hat{s}_i (i = 1, 2, 3)$.

<u>Consolidation Rule IV</u>: $$GD = D\hat{G} \quad . \tag{30}$$

If w_{mi} and $w_{ni} (i = 1, 2, 3, 4)$ retain their proportional relationship over time, i.e., if all d_i remain constant, changes in b_{mi}, b_{ni}, s_{mi}, and s_{ni}, will not affect the consolidated coefficients $\hat{b}_i$ and $\hat{s}_i$.

Consolidation Rules III and IV specify the conditions under which we may consolidate two cohorts without introducing error into our population forecasts. The first rule states that cohorts exposed to the same age-specific schedule of fertility and mortality may be consolidated. The second asserts that cohorts whose age groups maintain a constant proportional relationship to one another may be consolidated. As in the case of spatial aggregation, however, it is very unlikely that either condition will be met in practice. Nevertheless the rules do suggest how errors introduced by consolidation may be minimized.

4. Temporal Aggregation

Temporal aggregation differs slightly from spatial and cohort aggregation in that two or more time intervals are involved instead of one. This occurs because in temporal aggregation we are in fact consolidating age groups and with this action are simultaneously expanding the time interval. Thus, for example, in consolidating five-year age groups into ten-year age groups we are at the same time expanding the time interval from five to ten years.[8] A simple example will clarify this point. Imagine a regional population that is undisturbed by migration. Define this population by a four dimensional vector $\underline{w}$, with each element denoting the population in a fifteen-year age group. Thus we have, once again, only the following four age groups: 0-14, 15-29, 30-44, and 45+. We have seen, in the previous section, that the growth of such a population may be expressed by the matrix equation:

$$\underline{w}(t+1) = G\underline{w}(t) \tag{31}$$

or

$$\begin{bmatrix} w_1(t+1) \\ w_2(t+1) \\ w_3(t+1) \\ w_4(t+1) \end{bmatrix} = \begin{bmatrix} 0 & b_2 & b_3 & 0 \\ s_1 & 0 & 0 & 0 \\ 0 & s_2 & 0 & 0 \\ 0 & 0 & s_3 & 0 \end{bmatrix} \begin{bmatrix} w_1(t) \\ w_2(t) \\ w_3(t) \\ w_4(t) \end{bmatrix} .$$

Assume that the matrix G in (31) remains constant for another time interval. We have then:

$$\underline{w}(t+2) = G\underline{w}(t+1) = G^2\underline{w}(t) \tag{32}$$

or

$$\begin{bmatrix} w_1(t+2) \\ w_2(t+2) \\ w_3(t+2) \\ w_4(t+2) \end{bmatrix} = \begin{bmatrix} b_2s_1 & b_3s_2 & 0 & 0 \\ 0 & b_2s_1 & b_3s_1 & 0 \\ s_1s_2 & 0 & 0 & 0 \\ 0 & s_2s_3 & 0 & 0 \end{bmatrix} \begin{bmatrix} w_1(t) \\ w_2(t) \\ w_3(t) \\ w_4(t) \end{bmatrix} .$$

We now may consider the consequences of consolidating the first two and the last two age groups. First we define the <u>aggregation operator</u> A:

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \tag{33}$$

such that

$$\hat{\underline{w}}(t) = A\underline{w}(t) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} w_1(t) \\ w_2(t) \\ w_3(t) \\ w_4(t) \end{bmatrix} = \begin{bmatrix} w_{(1+2)}(t) \\ w_{(3+4)}(t) \end{bmatrix} \tag{34}$$

Next, we define the corresponding <u>disaggregation operator</u> D:

$$D = \begin{bmatrix} d_1 & 0 \\ 1-d_1 & 0 \\ 0 & d_3 \\ 0 & 1-d_3 \end{bmatrix} \tag{35}$$

where $d_i = \dfrac{w_i(t)}{w_i(t) + w_{(i+1)}(t)}$.

Thus

$$\underline{w}(t) = D\hat{\underline{w}}(t) = \begin{bmatrix} d_1 & 0 \\ 1-d_1 & 0 \\ 0 & d_3 \\ 0 & 1-d_3 \end{bmatrix} \begin{bmatrix} w_{(1+2)}(t) \\ w_{(3+4)}(t) \end{bmatrix} = \begin{bmatrix} w_1(t) \\ w_2(t) \\ w_3(t) \\ w_4(t) \end{bmatrix} . \tag{36}$$

We now recall Equation (13), replace G by G^2, and derive the consolidated growth matrix $\hat{G}$:[9]

$$\hat{G} = AG^2D = \begin{bmatrix} 1 & 1 & 0 & 0 \\ & & & \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} b_2s_1 & b_3s_2 & 0 & 0 \\ 0 & b_2s_1 & b_3s_1 & 0 \\ s_1s_2 & 0 & 0 & 0 \\ 0 & s_2s_3 & 0 & 0 \end{bmatrix} \begin{bmatrix} d_1 & 0 \\ 1-d_1 & 0 \\ 0 & d_3 \\ 0 & 1-d_3 \end{bmatrix}$$

$$= \begin{bmatrix} b_2s_1d_1 + (b_3s_2+b_2s_1)(1-d_1) & b_3s_1d_3 \\ & \\ s_1s_2d_1 + s_2s_3(1-d_1) & 0 \end{bmatrix} = \begin{bmatrix} \hat{b}_1 & \hat{b}_2 \\ & \\ \hat{s}_1 & 0 \end{bmatrix} . \tag{37}$$

Thus, in addition to the unconsolidated projection defined in (32), we have the following consolidated forecast:

$$\underline{\hat{w}}(t+2) = \hat{G}\underline{\hat{w}}(t) = \begin{bmatrix} \hat{b}_1 & \hat{b}_2 \\ \hat{s}_1 & 0 \end{bmatrix} \begin{bmatrix} w_{(1+2)}(t) \\ w_{(3+4)}(t) \end{bmatrix} = \begin{bmatrix} w_{(1+2)}(t+2) \\ w_{(3+4)}(t+2) \end{bmatrix} . \quad (38)$$

We now may derive Consolidation Rules V and VI. First, we note that the errorless population forecast for (t + 2) is given by Equation (32). Hence the associated errorless consolidated forecast is:

$$\underline{\hat{w}}(t+2) = A\underline{w}(t+2) = AG^2\underline{w}(t) \quad . \quad (39)$$

According to (38), the consolidated forecast is:

$$\underline{\hat{w}}(t+2) = \hat{G}\underline{\hat{w}}(t) = \hat{G}A\underline{w}(t) \quad . \quad (40)$$

Thus, the two forecasts will be identical, if:

<u>Consolidation</u> <u>Rule</u> <u>V</u>: $\quad AG^2 = \hat{G}A \quad . \quad (41)$

To derive the second consolidation rule, we observe that the errorless unconsolidated population forecast is: [10]

$$\underline{w}(t+2) = G^2\underline{w}(t) = G^2D_0\underline{\hat{w}}(t) \quad . \quad (42)$$

The corresponding consolidated forecast is:

$$\underline{w}(t+2) = D_2\underline{\hat{w}}(t+2) = D_2\hat{G}\underline{\hat{w}}(t) \quad . \quad (43)$$

Hence the two projections will be equal, if:

Consolidation Rule VI: $G^2 D_0 = D_2 \hat{G}$. (44)

The consolidation rules may be expressed in nonmatrix terms as before. However, the necessary and sufficient equalities increase in number and complexity throughout the rest of this paper. Since little is gained by a mere listing of such equalities, and because they are not an integral part of our subsequent arguments regarding rational consolidation, we no longer shall identify them individually.

5. Spatial Aggregation Revisited

Before considering the consequences of multidimensional consolidation, i.e., the simultaneous consolidation of regions, cohorts, and time intervals, we need to reconsider the problem of spatial aggregation and extend our previous results to populations disaggregated by age. This may be accomplished by dealing with matrices where earlier we dealt with numbers.

Let us return to the three region demographic system discussed in Section 2, and assume that each region's population may be denoted by a four-dimensional vector. Thus we have the vectors: $\underline{w}_i$, $\underline{w}_j$, and $\underline{w}_k$ and may denote their change over time by a matrix of matrices which is the direct analogue of Equation (2): [11]

$$\underline{w}(t+1) = G\underline{w}(t) \tag{45}$$

where

$$G = \left[\begin{array}{cccc|cccc|cccc}
0 & b_{i2} & b_{i3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
s_{i1} & 0 & 0 & 0 & {}_1g_{ji} & 0 & 0 & 0 & {}_1g_{ki} & 0 & 0 & 0 \\
0 & s_{i2} & 0 & 0 & 0 & {}_2g_{ji} & 0 & 0 & 0 & {}_2g_{ki} & 0 & 0 \\
0 & 0 & s_{i3} & 0 & 0 & 0 & {}_3g_{ji} & 0 & 0 & 0 & {}_3g_{ki} & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & b_{j2} & b_{j3} & 0 & 0 & 0 & 0 & 0 \\
{}_1g_{ij} & 0 & 0 & 0 & s_{j1} & 0 & 0 & 0 & {}_1g_{kj} & 0 & 0 & 0 \\
0 & {}_2g_{ij} & 0 & 0 & 0 & s_{j2} & 0 & 0 & 0 & {}_2g_{kj} & 0 & 0 \\
0 & 0 & {}_3g_{ij} & 0 & 0 & 0 & s_{j3} & 0 & 0 & 0 & {}_3g_{kj} & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & b_{k2} & b_{k3} & 0 \\
{}_1g_{ik} & 0 & 0 & 0 & {}_1g_{jk} & 0 & 0 & 0 & s_{k1} & 0 & 0 & 0 \\
0 & {}_2g_{ik} & 0 & 0 & 0 & {}_2g_{jk} & 0 & 0 & 0 & s_{k2} & 0 & 0 \\
0 & 0 & {}_3g_{ik} & 0 & 0 & 0 & {}_3g_{jk} & 0 & 0 & 0 & s_{k3} & 0
\end{array}\right]$$

$$= \left[\begin{array}{c|c|c}
G_{ii} & G_{ji} & G_{ki} \\
\hline
G_{ij} & G_{jj} & G_{kj} \\
\hline
G_{ik} & G_{jk} & G_{kk}
\end{array}\right]$$

and

$$\underline{w}(t) = \left[\begin{array}{c}
\underline{w}_i(t) \\
\hline
\underline{w}_j(t) \\
\hline
\underline{w}_k(t)
\end{array}\right] .$$

Notice that the migration proportions now have an additional subscript in order to identify the relevant age group at the region of origin. It is assumed that upon arriving at the region of destination the migrants enter the next age group.

Now let us consider once again the consequences of consolidating regions i and j. First we define the <u>aggregation operator</u> A:

$$A = \left[\begin{array}{cccc|cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \tag{46}$$

$$= \left[\begin{array}{c|c|c} I & I & 0 \\ \hline 0 & 0 & I \end{array}\right]$$

such that

$$\hat{\underline{w}}(t) = A\underline{w}(t) = \left[\begin{array}{c|c|c} I & I & 0 \\ \hline 0 & 0 & I \end{array}\right] \left[\begin{array}{c} \underline{w}_i(t) \\ \hline \underline{w}_j(t) \\ \hline \underline{w}_k(t) \end{array}\right] = \left[\begin{array}{c} \underline{w}_{(i+j)}(t) \\ \hline \underline{w}_k(t) \end{array}\right] . \tag{47}$$

Next, we define the <u>disaggregation operator</u> D:

$$D = \left[\begin{array}{cccc|cccc} d_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & d_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & d_3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & d_4 & 0 & 0 & 0 & 0 \\ \hline 1-d_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1-d_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1-d_3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1-d_4 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \tag{48}$$

$$= \left[\begin{array}{c|c} D & 0 \\ \hline I-D & 0 \\ \hline 0 & I \end{array}\right]$$

where $d_h = \dfrac{w_{ih}(t)}{w_{ih}(t) + w_{jh}(t)}$.

Thus

$$\underline{w}(t) = D\underline{\hat{w}}(t) = \left[\begin{array}{c|c} D & 0 \\ \hline I-D & 0 \\ \hline 0 & I \end{array}\right] \begin{bmatrix} \underline{w}_{(i+j)}(t) \\ \underline{w}_k(t) \end{bmatrix} = \begin{bmatrix} \underline{w}_i(t) \\ \underline{w}_j(t) \\ \underline{w}_k(t) \end{bmatrix} . \tag{49}$$

Given the growth matrix G and the operators A and D, we may once again use Equation (13) to derive the consolidated growth matrix $\hat{G}$:

$$\hat{G} = AGD = \left[\begin{array}{c|c|c} I & I & 0 \\ \hline 0 & 0 & I \end{array}\right] \left[\begin{array}{c|c|c} G_{ii} & G_{ji} & G_{ki} \\ \hline G_{ij} & G_{jj} & G_{kj} \\ \hline G_{ik} & G_{jk} & G_{kk} \end{array}\right] \left[\begin{array}{c|c} D & 0 \\ \hline I-D & 0 \\ \hline 0 & I \end{array}\right]$$

$$= \left[\begin{array}{c|c} (G_{ii}+G_{ij})D + (G_{jj}+G_{ji})(I-D) & G_{ki}+G_{kj} \\ \hline G_{ik}D + (I-D)G_{jk} & G_{kk} \end{array}\right] . \tag{50}$$

With these redefinitions of G, A, D, and $\hat{G}$, we may once again invoke Consolidation Rules I and II, as defined by Equations (16) and (19), to express the sufficient conditions for perfect aggregation to be possible.

6. The Joint Impact of Aggregations and A Guide for Consolidation

Having considered the consequences of spatial, cohort, and temporal aggregation, when individually applied, we now extend our results to include the effects of multidimensional aggregation by applying the aggregation and disaggregation operators for each form of aggregation in sequence. An example will clarify the method.

Imagine a three region, two cohort, and four age-group demographic system. Such a population may be denoted by a column vector containing 24 elements, and its growth regime may be defined by a 24 by 24 matrix. The fundamental matrix growth equation for this population system, as before, is:

$$\underline{w}(t+1) = G\underline{w}(t) \tag{51}$$

where the matrix G now consists of submatrices arranged as follows:

$$
G = \left[\begin{array}{ccc|ccc}
G^m_{ii} & G^m_{ji} & G^m_{ki} & 0 & 0 & 0 \\
G^m_{ij} & G^m_{jj} & G^m_{kj} & 0 & 0 & 0 \\
G^m_{ik} & G^m_{jk} & G^m_{kk} & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & G^n_{ii} & G^n_{ji} & G^n_{ki} \\
0 & 0 & 0 & G^n_{ij} & G^n_{jj} & G^n_{kj} \\
0 & 0 & 0 & G^n_{ik} & G^n_{jj} & G^n_{kk}
\end{array}\right]
$$

The superscripts m and n differentiate white and nonwhite operators, respectively, and the subscripts i, j, and k differentiate regional operators.

Now let us consolidate our demographic system into one consisting of two regions, (i + j) and k; a single cohort, total population; and a doubled time interval. We begin by performing the temporal consolidation first. The order of each successive consolidated growth matrix is listed below it, and each aggregation and disaggregation operator is distinguished by a subscript. Thus

$$\underset{12\times12}{\hat{G}} = A_t G^2 D_t \quad .$$

Next we aggregate the two cohorts:

$$\underset{6\times6}{\hat{\hat{G}}} = A_c \hat{G} D_c = A_c A_t G^2 D_t D_c \quad .$$

Finally, we consolidate regions i and j:

$$\underset{4\times4}{\hat{\hat{\hat{G}}}} = A_r \hat{\hat{G}} D_r = A_r A_c A_t G^2 D_t D_c D_r$$

$$= AG^2 D \tag{52}$$

where $A = A_r A_c A_t$ and $D = D_t D_c D_r$.

By now it should be clear that the consolidation rules for multi-dimensional aggregation are:

<u>Consolidation Rule VII</u>: $AG^2 = \hat{\hat{G}}A$. (53)

<u>Consolidation Rule VIII</u>: $GD_0 = D_2\hat{\hat{G}}$. (54)

As before, the subscripts on D distinguish between the beginning and end of the time interval (t, t + 2).

The consolidation rules expressed by (53) and (54) probably are never met in practice. Thus a demographer considering aggregation is immediately confronted with a decision problem, namely: Which one of two or more alternative consolidation schemes most closely approximates the necessary conditions for perfect aggregation? Rough indices are offered by the coefficients of correlation, say R_1 and R_2, between the elements of AG^2 and $\hat{\hat{G}}A$ and between the elements of GD_0 and $D_2\hat{\hat{G}}$, respectively. Thus, for example, if Alternative I offers $R_1(I)$ and $R_2(I)$ each of which is greater than the corresponding correlation coefficient of Alternative II, $R_1(II)$ and $R_2(II)$, say, then it would seem that Alternative I is the superior consolidation scheme. However, if $R_1(I) > R_1(II)$ but $R_2(I) < R_2(II)$, the decision problem becomes more complicated. Moreover, in the event that the population projection extends over several time intervals, both indices will suffer to the extent that G and D change over time. Whereas in some instances we may be willing to assume that G remains constant over the projection period, it is quite another matter to make the same assumption with regard to D. Thus it appears that the first consolidation rule offers the most useful guide for rational consolidation. We now illustrate its use with a numerical example.

7. Example

The data for this numerical example are drawn from tables published in George W. Barclay's Techniques of Population Analysis (New York: John Wiley, 1958), p. 213 and p. 236. The demographic system consists of a single region, Chile; two cohorts, male and female; and eighteen five-year age groups for each cohort. The base period is 1940, and the unit time interval is five years long. Figure 1 presents the data in matrix form and includes the projection to 1970 that is generated by the model under the assumption of an unchanging schedule of growth. (Notice the slight modification in the matrix operator that has been introduced in order to apply birth rates only to the female cohort.)

Assume that budget constraints are such that a population projection using the model in Figure 1 is too costly and that a consolidation of time (and therefore age groups) or a consolidation of cohorts, or both, must be considered. What are the costs in terms of "information" loss ? The simplest way to measure this information loss would be to run all three alternative projections to the end of the forecasting period, say 1970, and then compare the projected vector of each with that generated in the absence of any aggregation. Of course this is completely unrealistic in practice since the costs of such a procedure would far exceed those of the unconsolidated projection. Thus a method that does not require any projections at all would be much preferred. Earlier we suggested the use of the coefficient of correlation between AG and $\hat{G}A$ as an index with which to rate alternative consolidation schemes. Let us now see which of the three consolidations this measure would recommend for the Chile example.

We begin by applying Equation (13) to derive the consolidated growth operators. This produces the models presented in Figures 2, 3, and 4.

Table 1 lists the "errorless" and consolidated 1970 projections that are generated by each consolidation scheme under the assumption of a constant regime of growth. Also included are the coefficients of correlation between "errorless" and consolidated projections.

It appears that the consolidation of the five-year time intervals into fifteen-year intervals is to be preferred to a consolidation of male and female cohorts. The latter consolidation scheme introduces an upward bias of about 70,000 into the population projection, whereas the former only increases the projected population by about 46,000. The distributional properties of the two projections, as measured by the correlation coefficient, seem to be rather similar, but again the temporal consolidation appears to be slightly more accurate ($r = .999947$ vs. $r = .999490$). The consolidation of both time and cohorts overprojects the "true" 1970 population by more than 93,000 and clearly is the least desirable consolidation alternative. However, with respect to projection costs, it is by far the most economical one to use since it collapses the original 36 by 36 matrix growth operator to one which is only 6 by 6. Thus its selection would depend on the severity of the budget constraint.

Let us now see how these three consolidation alternatives would be ranked by our index of performance. The coefficient of correlation between AG^3 and $\hat{G}A$ for the temporal consolidation alternative is .94092, for the cohort consolidation it is .88384, and for both consolidations we find $r = 62796$. Thus it appears that the index of performance we have suggested ranks the alternatives in the correct order and, therefore, does lead to a reasonable decision rule for consolidation. This index may be used to rank alternatives according to their expected "information loss," and the size of the growth operator may be used as a very crude indicator of projection costs. The compromise between accuracy and cost then can be decided upon in accordance with the budget constraint.

8. Conclusion

In this paper we set out to identify the conditions under which perfect aggregation, i.e., aggregation without the introduction of error, is possible in demographic analysis. This question first led us to a consideration of how one might express consolidation operations in matrix terms. Equation (13) provides us with a means for consolidating space, people, and time in interregional population models without requiring a return to the basic data once the disaggregated matrix growth operator has been estimated. In terms of economies in computer processing this is a most useful result. With the matrix expression for consolidation operations we next proceed to derive the sufficient, but not necessary, conditions for perfect aggregation to be possible. In general, these conditions may be summarized by the two matrix expressions: $AG = \hat{G}A$ and $GD_0 = D_1\hat{G}$. Finally, we have suggested an index which appears to be a useful guide for making rational consolidation decisions. Although much more needs to be said with regard to the aggregation problem in demography, it is hoped that this essay will provide a useful point of departure for such efforts.

FIGURE 1--PROJECTION OF 1940 POPULATION OF CHILE TO 1970: "ERRORLESS" PROJECTION

																																				w(t)		w(t+6)
0	0	0	.13165	.35108	.41993	.40415	.31028	.15783	.05822	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	308336		486184
.93128	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	312905		431541
0	.98632	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	302490		405840
0	0	.97524	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	256482		377771
0	0	0	.96096	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	238059		335894
0	0	0	0	.95495	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	222506		295840
0	0	0	0	0	.95364	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	174974		241717
0	0	0	0	0	0	.95127	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	165587		250559
0	0	0	0	0	0	0	.94697	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	131692		232557
0	0	0	0	0	0	0	0	.94150	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	106840		190364
0	0	0	0	0	0	0	0	0	.93096	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	87940		171176
0	0	0	0	0	0	0	0	0	0	.90904	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	66408		152299
0	0	0	0	0	0	0	0	0	0	0	.87886	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	62780		110375
0	0	0	0	0	0	0	0	0	0	0	0	.82850	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	33784		90973
0	0	0	0	0	0	0	0	0	0	0	0	0	.75887	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	28537		57979
0	0	0	0	0	0	0	0	0	0	0	0	0	0	.66301	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14166		33125
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.56667	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11109		16596
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.45554	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9018		6280
0	0	0	.13530	.36083	.43162	.41539	.31899	.16405	.05977	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	313667	=···=	500353
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.93073	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	321952		443701
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.98651	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	306807		420343
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.97807	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	256321		389488
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.96277	0	0	0	0	0	0	0	0	0	0	0	0	0	0	218871		346987
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.95452	0	0	0	0	0	0	0	0	0	0	0	0	0	204395		305455
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.95279	0	0	0	0	0	0	0	0	0	0	0	0	177431		246643
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.94834	0	0	0	0	0	0	0	0	0	0	0	155405		257948
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.93959	0	0	0	0	0	0	0	0	0	0	136983		234124
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.92642	0	0	0	0	0	0	0	0	0	109145		185266
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.90615	0	0	0	0	0	0	0	0	86081		148898
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.87681	0	0	0	0	0	0	0	68181		127728
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.84095	0	0	0	0	0	0	55443		97862
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.78806	0	0	0	0	0	32586		71228
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.71279	0	0	0	0	22562		47629
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.60001	0	0	0	11799		24578
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.59391	0	0	6958		12706
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.42424	0	5339		4869
																																				5023539		7756276

Source: Constructed from data published in George W. Barclay, Techniques of Population Analysis (New York: John Wiley, 1958), p. 213 and p. 236

Figure 2--PROJECTION OF 1940 POPULATION OF CHILE TO 1970: TEMPORAL CONSOLIDATION

$$
\overset{G}{\left[\begin{array}{cccccc:cccccc}
.18979 & .96513 & .51367 & .02188 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
.90519 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & .86721 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & .82745 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & .66759 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & .32522 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hdashline
.19504 & .99221 & .52933 & .02245 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & .90886 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & .86337 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .78569 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .58902 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .27465 & 0 \\
 & & & & & & & & & & &
\end{array}\right]}
\overset{\underline{w}(t)}{\left[\begin{array}{c}
923731 \\ 717047 \\ 472253 \\ 261188 \\ 125101 \\ 34293 \\ \hdashline 942426 \\ 679587 \\ 469819 \\ 263407 \\ 110591 \\ 24096 \\ 5023539
\end{array}\right]}
=\cdots=
\overset{\underline{w}(t+6)}{\left[\begin{array}{c}
1346701 \\ 1009880 \\ 725119 \\ 514533 \\ 260871 \\ 56707 \\ \hdashline 1385162 \\ 1042887 \\ 739505 \\ 460992 \\ 217426 \\ 42613 \\ 7802396
\end{array}\right]}
$$

Source: Figure 1

Figure 3--PROJECTION OF 1940 POPULATION OF CHILE TO 1970: COHORT CONSOLIDATION

$$
\overset{G}{\begin{bmatrix}
0 & 0 & 0 & .13352 & .37091 & .44384 & .40691 & .32456 & .15777 & .05836 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
.93100 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & .98642 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & .97667 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & .96186 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & .95474 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & .95323 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & .94979 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & .94340 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .93381 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .91842 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .89310 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .84966 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .80953 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .73625 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .63519 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .57905 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .44349 & 0
\end{bmatrix}}
\overset{\underline{w}(t)}{\begin{bmatrix}
622003 \\ 634857 \\ 609297 \\ 512803 \\ 456930 \\ 426901 \\ 352405 \\ 320992 \\ 268675 \\ 215985 \\ 174021 \\ 134589 \\ 118223 \\ 66370 \\ 51099 \\ 25965 \\ 13067 \\ 14357
\end{bmatrix}}
= \cdots =
\overset{\underline{w}(t+6)}{\begin{bmatrix}
1018122 \\ 898189 \\ 845653 \\ 774911 \\ 683853 \\ 601282 \\ 488367 \\ 508521 \\ 466762 \\ 375602 \\ 319562 \\ 279286 \\ 205499 \\ 159540 \\ 104215 \\ 56987 \\ 28948 \\ 11118
\end{bmatrix}}
$$

$$5023539 \qquad 7826417$$

Source: Figure 1

FIGURE 4--PROJECTION OF 1940 POPULATION OF CHILE TO 1970: COHORT AND TEMPORAL CONSOLIDATION

$$
\begin{array}{ccccc}
G & & t & & t+6 \\
\begin{bmatrix}
.19049 & 1.00492 & .52285 & .02207 & 0 & 0 \\
.90704 & 0 & 0 & 0 & 0 & 0 \\
0 & .86534 & 0 & 0 & 0 & 0 \\
0 & 0 & .80662 & 0 & 0 & 0 \\
0 & 0 & 0 & .62814 & 0 & 0 \\
0 & 0 & 0 & 0 & .30149 & 0
\end{bmatrix}
& &
\begin{bmatrix}
1866157 \\ 1396634 \\ 942072 \\ 524595 \\ 235692 \\ 58389
\end{bmatrix}
& \doteq &
\begin{bmatrix}
2780779 \\ 2052750 \\ 1464743 \\ 974851 \\ 477320 \\ 99347
\end{bmatrix} \\
& & 5023539 & & 7849790
\end{array}
$$

Source: Figure 2

Table 1--CORRELATION BETWEEN "ERRORLESS" AND CONSOLIDATED 1970 PROJECTIONS OF CHILE'S POPULATION

Temporal Consolidation	
Consolidated Projection	Errorless Projection
1346701	1326965
1009880	1009505
725119	724833
514533	513839
260871	259327
56707	56001
1385162	1364397
1042887	1041930
739505	738715
460992	461892
217426	216719
42613	42153
7802396	7756276
r = .999947	

Cohort Consolidation	
Consolidated Projection	Errorless Projection
1018122	986837
898189	875342
845653	829183
774911	767259
683853	682881
601282	601295
488367	488360
508521	508507
466762	466681
375602	375630
319562	320074
279286	280027
205499	208237
159540	162201
104215	105608
56987	57703
28948	29302
11118	11149
7826417	7756276
r = .999490	

Temporal and Cohort Consolidation	
Consolidated Projection	Errorless Projection
2780779	2691362
2052750	2051435
1464743	1463548
974851	975731
477320	476046
99347	98154
7849790	7756276
r = .999643	

References

1. Keyfitz, Nathan, "The Population Projection as a Matrix Operator," Demography, I (1964), pp. 56-73

2. Rogers, Andrei, Matrix Analysis of Interregional Population Growth and Distribution (Berkeley, California: University of California Press, 1968), in press.

3. Stone, Richard, Input-Output and National Accounts (Paris: O.E.C.D., 1961)

FOOTNOTES

1. The author gratefully acknowledges the support of the American-Yugoslav Project in Regional and Urban Planning Studies in Ljublana, Yugoslavia, and the computational efforts of his able research assistant, Gregory Zore.
2. See, for example: Stone (3; pp. 101-104).
3. In answering this question we assume, in the absence of other information, that G remains constant over the projection period.
4. Of course, one has to prespecify the relevant forecasting period. In some applications this period might be infinitely long. The evaluation then may be made against the stable population growth rate and distribution.
5. Notice that this is a sufficient, but not necessary, condition.
6. The D in (17) is not necessarily the same D that appears in (18). Thus we distinguish them by the subscripts 0 and 1, respectively, to differentiate the D of time t from that of time (t+1).
7. See, for example: Keyfitz (1).
8. Although it is possible to consolidate age groups without expanding the time interval accordingly (see footnote 9), such cohort aggregation is rarely used in practice except to consolidate the populations of all age groups into a total population figure.
9. Note that if, instead, we apply Equation (13) without modification (i.e., with G and not G^2), we obtain an aggregation of age groups without temporal consolidation.
10. Once again we introduce subscripts to distinguish between the D's at the beginning and end of our time interval.
11. See: Rogers (2; Chapter 2).