STANDARDIZATION AS A TECHNIQUE FOR TEMPORAL COMPONENT ANALYSIS

By

Dr. ABDEL-FATTAH NASSEF

Institute of National Planning

1.—INTRODUCTION

Since the early 1880's standardization techniques have been in use in demographic research. Their simplest application, which was practiced for most of the period since 1980's, is to secure a reasonably valid comparison between two or more different populations regarding certain demographic phenomenon, such as the crude death rate, independent of one of the major compositional factors, such as age, through holding that factor constant in the populations to be compared.

In recent decades, in addition to mortality, the use of standardization has been extended to other fields of demography and other social sciences. Moreover, multiple standardization has been developed in order to control more than just one factor.

Still more recently, standardization techniques, simple and/or multiple, have been used not only for purposes of «controlled» comparisons but also for factorial analysis. In principle, the latter use is based on the subtraction of the standardized rate, or level, of the phenomenon under investigation from its original, or «crude», counterpart. The result is an estimate of the effect of the factor, or factors, for which standardization was carried out.

Beside the estimates for the «controlled» factors, some experts in the field have applied methods by which they made allowance for the effects of simultaneous changes in more than one factor, (i.e., interaction effect) [1].

---

(1) See J. D. Durand, *The Labor Force in the United States, 1890-1960* (New York : Social Science Research Council, 1948), Appendix B. ; J. D. Durand and A. R. Miller, *Methods of Analysing Census Data on Economic Activities of the Population* (U. N., Department of Economic and Social Affairs, ST/SOA/Ser. A/43, 1968).

The purpose of this paper is not to review the literature on standardization techniques and their various applications [2]. Rather, it is to present, in some detail, an analytical framework for using standardization as a technique for temporal compenent analysis that has been developed during the course of another study, and its application to Egyptian labour force data [3]. The scheme and its applications are described in the following sections.

2. COMPONENT ANALYSIS OF INTERCENSAL CHANGES IN CRUDE ACTIVITY RATE.

The crude activity rate (W) is defined as the percentage of the total population (P) in the labor force (LF). That is,

$$W = \frac{LF}{P} 100$$

This rate is a weighted average of age specific activity rates (W_x), where the weights are the relative shares of various age groups in the total population (P_x /P). Thus,

$$W = \frac{\Sigma P_x W_x}{P}$$

$$= \Sigma C_x W_x \cdot$$

were $c_x = p_x / p$

Therefore, if a country had a changing age structure coupled with a constant pattern of age-specific activity rates, the age structure would be the sole source of change of its crude activity rate. If, on the other hand, the age structure remained stable, the variation in the age pattern of participation in economic activities would be the only source of change in the crude activity rate. In most cases, however, both age structure and age-specific activity rates vary simultaneously,

---

(2) See references in footnote (1) and, A. J. Jaffe, *Handbook of Statistical Methods for Demographers* (Washington : U. S. Government Printing Office, 1960) ; E. M. Kitagawa, «Standardized Comparisons in Population Research», *Demography,* Vol. 1, 1964 ; S. L. Wolfbein and A. J. Jaffe, «Demographic Factors in Labor Force Growth», *American Sociological Review,* Vol. XI, No. 4, August 1946.

(3) A. Nassef, *The Egyptian Labor Force : Its Dimensions and Changing Structure 1907-1970,* Unpublished Ph. D. Dissertation, University of Pennsylvania, 1969, Sections 3.1.2 and 3.5.1 and Appendix C. This paper includes further extensions in the applications of the scheme. Data used here are somewhat different from those in the given reference.

and both contribute in varying degrees to the total change of crude activity rate. Consequently, the question arises as to how much of the total change is attributable to each source. The purpose here is to approach this question by means of standardization.

In general, and according to standardization techniques, when estimating the effect of a factor on a phenomenon one allows this factor to change while holding other factors at constant levels. In studying intercensal changes the problem arises, in what direction the changing factors should be allowed to changs, and what would be the implied levels of the constant factor (or factors). For convenience, let it be called the «Forward» method which takes the direction of the changing factor as it actually occurred with the paseage of time. The intercensal change in that factor is then equal to its terminal minus its initial levels. Thus, in order to conform with the above mentioned basic principles of standardization and the estimation of components of change, the implied levels of the factors to be held constant are those of the terminal date. The «Reverse» method, on the other hand, deals with the changing factor in the opposite direction and takes the initial levels for the constant factors. Most likely, the results obtained by the two methods will be different due to the interaction effect. One way of eliminating the interaction term in the results is to use the «Average» method, implying the use of the mean value of the constant factor. The procedures for estimating intercensal components of change in crude activity rate by the three methods may be described as follows :

Let $\Delta W(c_x)$ represent the change in crude activity rate due to change in the age structure of the population (i. e., age component) ; $\Delta W(w_x)$, the change in crude activity rate due to changes in age-specific activity rate (i. e., activity component) ; $\Delta W(c_x, w_x)$, the change in activity rate due to the interaction of changes in the age structure and the pattern of activity rates by age (i. e., interaction effect) ; and $\Delta W(T)$, the total change in crude activity rate. Then :

(i) *Forward Method :*

$$\Delta W(c_x) = \sum_{x=i}^{\infty} (c_{x,2} - c_{x,1}) , W_{x,2} \qquad (F.1)$$

$$\Delta W(w_x) = \sum_{x=i}^{\infty} (w_{x,2} - w_{x,1}) \cdot c_{x,2} \qquad (F.2)$$

$$\Delta W(c_x, w_x) = \sum_{x=i}^{\infty} (c_{x2} - c_{x,1}) (w_{x,2} - w_{x,1}) \qquad (F.3)$$

(ii) *Reverse Method :*

$$\Delta W(c_x) = \sum_{x=i}^{\infty} (c_{x,1} - c_{x,2}) \cdot w_{x,1} \qquad (R.1)$$

$$\Delta W(w_x) = \sum_{x=i}^{\infty} (w_{x,1} - w_{x,2}) \cdot c_{x,1} \qquad (R.2)$$

$$\Delta W(c_x, w_x) = \sum_{x=i}^{\infty} (c_{x,1} - c_{x,2})(w_{x,1} - w_{x,2}) \qquad (R.3)$$

(iii) *Average Method :*

$$\Delta W(c_x) = \sum_{x=i}^{\infty} (c_{x,2} - c_{x,1})(w_{x,1} + w_{x,2})/2 \qquad (A.1\text{—}a)$$

$$\text{or} \quad = \sum_{x=i}^{\infty} (c_{x,1} - c_{x,2})(w_{x,1} + w_{x,2})/2 \qquad (A.1\text{—}b)$$

$$\Delta W(w_x) = \sum_{x=i}^{\infty} (w_{x,2} - w_{x,1})(c_{x,1} + c_{x,2})/2 \qquad (A.2\text{—}a)$$

$$\text{or} \quad = \sum_{x=i}^{\infty} (w_{x,1} - w_{x,2})(c_{x,1} + c_{x,2})/2 \qquad (A.2\text{—}b)$$

where suffixes 1 and 2 denote the initial and terminal magnitudes.

It can be shown that the equation for the total change in crude activity rate is,

$$\Delta W(T) = \Delta W(c_x) + \Delta W(w_x) - \Delta W(c_x, w_x)$$

whether we use the forward or the reverse method. The equation still holds for the average method, given that the interaction effect is equal to zero.

It may be noted that from the equations given above, the estimates for each by forward and reverse methods have opposite signs : a fact that should be expected if two-way standardization is carried out with regard to any specific variable.

The interaction effect, as defined by the two methods, has a definite and identical value and with the same sign. Aside from the sign implied in equations F. 3 and R. 3 for the interaction effect, its value is to be subtracted algebraically in the equation of the total change in crude activity rate.

These and other facts, to be discussed below may be illuctrated by a numerical example. For convenience, the definitional equations given above may be rewritten along with the results of the computations for the Egyptian males during 1947/1960 intercensal period as follows :

(i) *Forward Method :*

$$\Delta\ W\ (c_x) = W_2 - S_1 = -4.02 \quad (F.1)$$
$$\Delta\ W\ (w_x) = W_2 - S_2 = -3.52 \quad (F.2)$$
$$\Delta\ W\ (c_x, w_x) = W_1 + W_2 - S_1 - S_2 = .14 \quad (E.3)$$

(ii) *Reverse Method :*

$$\Delta\ W\ (c_x) = W_1 - S_2 = 4.16 \quad (R.1)$$
$$\Delta\ W\ (w_x) = W_1 - S_1 = 3.66 \quad (R.2)$$
$$\Delta\ W\ (c_x'\ w_x) = W_1 + W_2 - S_1 - S_2 = .14 \quad (R.3)$$

(iii) *Average Method :*

$$\Delta\ W\ (c_x) = \tfrac{1}{2}(W_2 + S_2 - W_1 - S_2) = -4.09 \quad (A.1-a)$$
$$\text{or} = \tfrac{1}{2}(W_1 + S_1 - W_2 - S_2) = 4.09 \quad (A.1-b)$$
$$\Delta\ W\ (w_x) = \tfrac{1}{2}(W_2 + S_1 - W_1 - S_2) = -3.59 \quad (A.2-a)$$
$$\text{or} = \tfrac{1}{2}(W_1 + S_2 - W_2 - S_1) = 3.59 \quad (A.2-b)$$

where W_1 (62.83) is the crude activity rate for the initial date of the period ; W_2 (55.15), the crude activity for the terminal date ; S_1 (59.17), the standardized activity rate as of the terminal date with the age structure of the population of the initial date as weights, i. e., $\sum_{x=i}^{\infty} (w_{x,2}\ x_{x,1})$; and S_2 (58.67), the standardized activity for the initial date with the age structure of the terminal date as weights, i. e. $\sum_{x=i}^{\infty} (w_{x,1} \cdot c_{x,2})$.

Many of the statements above, especially those related to the magnitudes of the effect of each factor as estimated by the three methods as well as their signs, are clearly given by the numerical example and need no further comments. However, the imposed nature of the interaction effect ; its value in relation to the estimates of other components by both forward and reverse methods ; and the determination of the sign of different estimates of the components of change calculated by the average method as well as the way in which the latter method eliminates the interaction effects, require further examination. The results of such examination may be summarized as follows :

(i) For estmating the effects of changes in the age structure of the population by the forward and reverse methods (equation F. 1 and R. 1 respectively), the age structure is

allowed to change from $c_{x,1}$ to $c_{x,2}$ or $c_{x,2}$ to $c_{x,1}$ In other words, the value of the first term in the two equations is the same with, of course, different signs. If the other factor, which is to be held constant, were the same for the two dates, the absolute value of the estimates of the age component would be the same regardless of the method used for estimation. However, when the value of the other factor also changes (in the equations), the interaction effect must be taken into account. This effect is implied, only once, in one or the other of the two estimates depending on how one interprets the data. The same holds with respect to estimates of the activity component.

(ii) Since the interaction effect is to be imputed only once in each of the two estimates of a component, its value should be equal to the difference between the two estimates, regardless of sign. Or, equivalently, since the two estimates have opposite signs, the interaction effect should be equal to the algebraic sum of the two estimates. For example, estimates of the age component for males by the forward and reverse methods are —4.02 and +4.16 respectively, and both the difference between their absolute values and their algebraic sum are equal to the interaction effect +.14 estimated independently by equation F. 3 or R. 3. The difference between the absolute values of the two estimates, however, does not provide the sign of the interaction component. This relationship can also be proved mathematically by means of equations F. 1, R. 1 and either F. 3 or R. 3. Thus, it will be seen that equation F. 3 or R. 3 is the algebraic sum of equations F. 1 and R. 1. This finding is equally true with regard to the activity component in relation to the interaction component.

(iii) Finally, it may be noted that though each pair of equations of the average method provides the same absoluts value of the estimates of the component of change under investigation, the sign is different, depending upon what equations are used for estimation (A. 1-a and A. 2-a, or A. 1-b and A. 2-b above). The obvious reason is that the process of averaging the values of the other factor to be held constant, in effect neutralizes that factor and, therefore, eliminates the effects of simultaneous changes of the two factors.

Nevertheless, when using the average method for estimating components of change, one still needs to know the direction of these estimates. Equations A. 1-a and A. 2-a vs. A. 1-b and A. 2-b provide the answer. The first two give the sign of the effects when the changing factor moves from its initial to its terminal levels, while the two other equations give the sign when the movement is in the other direction. In fact, equations A. 1-a and A. 2-a correspond to equations F. 1 and F. 2 of the forward method ; while equations A. 1-b and A. 2-b correspond to equations R. 1 and R. 2 of the reverse method. The only difference is the averaging process of the «controlled» factor in the average method. To put it differently, equations A. 1-a and A. 1-a give the components of W_2-W_1 ; while A. 1-b and A. 2-b indicate the components of W_1 — W_2.

It may be noted that every equation in the analytical scheme described above may answer a specific question under certain restric tions. Thus, the analyst has the choice of selecting any particular equation or combination of equations depending on the question under consideration and the availability of data. But he must always be aware of the relations between the different estimates obtained by different methods and the role of the interaction effect for a proper interpretation of his results.

Wolfbein and Jaffe in a leading article on demographic vs. socic-economic factors of the growth of the U. S. labor force used what may be described as equations F. 2 and R. 1 given above (provided that the demographic variables they dealt with are represented by c_x values).[4] In fact, the latter equation (R. 1) was used by them with a different sign. This implied that in estimating the two components of changes in the crude activity rate they maintained one direction for the changing variables in the two equations while exchanging the levels of the factors to be held constant. It also implied the estimation of the socio-economic component by subtracting the standardized rate from its crude counterpart, while the demographic component was estimated by subtracting the crude activity rate from its corresponding rate.[5] Aside from these inconsistencies of procedure, their

(4) Wolfbein and Jaffe, *Ibid.*, pp. 392-396.

(5) Similar procedure has been used in Durand and Miller, *op. cit.*, pp. 43-46.

analysis ignored the interaction effect, which might have been of significant value.[6]

Table (1) provides estimates of the effects of changes in the age structure of the total population on the crude activity rate as against the effects of changes in age specific activity rates due to all other factors, for each of the two intercensal periods between 1937 and 1960.[7] The two components of change are estimated by the average method (equations A. 1-a and A. 2-a).

The results show that the effects of changes in the age composition were to increase the crude activity rate during the 1937/1947 intercensal period. The reverse was true between 1947 and 1960. Furthermore, the positive effects during the earlier intercensal period were much less significant than the major negative effects in the latter period. This trend is explained by the relative stability of the Age structure during 1937/1947 intercensal period, followed by a significant increase in the proportion of children between 1947 and 1960 as a result of the substantial decline in infant mortality coupled with a relatively stable level of fertility.

TABLE 1

Estimates of Intercensal Components of Change in Crude Activity Rate by sex, Egypt, 1937-1960

Period	Age Component	Activity Component	Total Change	Age Component	Activity Component	Total Change
		Percentage Points			*% of Initial Rate*	
			Males			
1937/1947	+ .26	—2.47	—2.21	0.39	—3.79	— 3.40
1947/1960	—4.09	—3.59	—7.68	—6.50	—5.71	—12.21
			Females			
1937/1947	+ .15	— .25	— .10	+1.90	—3.17	— 1.27
1947/1960	— .23	—2.70	—2.93	—2.96	—34.74	—37.70
			Both Sexes			
1937/1947	+ .39	—1.83	—1.44	+1.06	—5.01	— 3.95
1947/1960	—1.99	—2.91	—4.90	—5.67	—8.30	—13.97

(6) The estimates of the two components by the equations used by them add up to the total change. By the same token, equations F. 1 and R. 2, with similar modification in sign, give estimates that add up also to the total change but with different magnitudes due to the interaction effct.

(7) The small numbers of persons whose age is not given are excluded.

The effects of all factors other than age structure, as reflected by the changes in age-specific activity rates, were significant and of negative nature during the intercensal periods.

3.—OTHER APPLICATIONS

In general, the scheme presented above is applicable in studying various socio-economic phenomena for which the required data are available. It may be adviseable to add few other examples using labor force data.

Thus, the total change in the size of labor force during given period of time Δ LF (T) may be decomposed into two components ; namely, population component, Δ LF (P), i. e., the change in labor force size due to change in population size ; and activity rate component, Δ LF (W) i. e., the change in the size of labor force due to change in crude activity rate. According to the average method, the two components may be comuted as follows : (8)

$$\Delta \text{ LF (P)} = (P_2 - P_1)\ (W_1 + W_2)\ /2$$

$$\Delta \text{ LF (W)} = (W_2 - W_1)\ (P_1 + P_2)\ /2$$

The results of this application is given in Table (2). It is clear that, in general, the contribution of population growth to the changing size of the Egyptian labor force overshadowed that attributable to the changes in the rate of participation in economic activites. In fact, the declining trend of activity rate is, of course, responsible for its negative effect on the size of labor force during the two intercensal periods. This is particularly so in the case of females between 1947 and 1960.

From the foregoing discussion, it is possible to decompose the growth of labor force into three components. Thus, according to the average method, and using the same notation, the change in labor force size may be subdivided into the following three components :

(i) The change in the size of the labor force due to population growth which, as indicated before, is equal to

$$\tfrac{1}{2}\ (P_2 - P_1)\ (W_1 + W_2)$$

---

(8) The corresponding equations in the Forward and Reverse methods can be derived as explained in the preceding section.

TABLE 2

Estimates of Intercensal Components of Labor Force Growth by sex, Egypt, 1937-1960

Components	1937/47	1947/60	1937/47	1947/60
	Absolute Numbers		*% of Initial Labor Force Size*	
		Males		
△ LF (P)	— 906 732	2 183 338	17.53	37.09
△ LF (W)	191 249	— 861 892	— 3.69	—14.64
△ LF (T)	715 483	1 321 446	13.84	22.45
		Females		
△ LF (T)	116 588	— 116 093	18.65	—15.65
△ LF (W)	— 9 122	— 328 854	— 1.45	—44.35
△ LF (P)	125 710	212 761	20.10	28.70
		Both Sexes		
△ LF (P)	1032 442	2 396 099	17.80	36.14
△ LF (W)	— 200 371	—1 190 746	— 3.45	—17.96
△ LF (T)	832 071	1 205 353	14.35	18.18

TABLE 3

Estimates of Intercensal Components of Labor Force Growth due to Changes in age Structure and Age-Specific Activity Rates Egypt, 1937—1960 *

Components	1937/47	1947/60	1937/47	1947/60
	Absolute Numbers		% of Initial Labor Force Size	
		Males		
LF (cx)	+ 22 581	—459 479	+ .44	— 7.81
LF (Wx)	—213 860	—402 675	—4.14	— 6.84
		Females		
LF (c x)	+ 13 048	— 26 429	+2.09	— 3.56
LF (W x)	— 22 053	—302 742	—3.53	—40.83
		Both Sexes		
LF (c x)	+ 35 629	—485 908	+ .61	— 7.33
LF (Wx)	—235 913	—705 417	—4.07	—10.65

(*) Due to rounding errors, the sum of the two components may not be exactly equal to the corresponding values of LF (W) in Table (2).

(*ii*) The change in labor force size due to changes in the age structure of the population, Δ LF (c_x). The equation of estimating this component is

$$\Delta \text{ LF } (c_x) = \tfrac{1}{2} (P_1 + P_2) \sum_{x=i}^{\infty} (c_{x,2} - C_{x,1}) (w_{x,1} + w_{x,2})$$

(*iii*) The change in the size of labor force due to changes in age-specific activity rates, Δ LF (w_x). This component may be estimated as follows

$$\Delta \text{ LF } (w_x) = \tfrac{1}{2} (P_1 + P_2) \sum_{x=i}^{\infty} (w_{x,2} - w_{x,1} (c_{x,1} + c_{x,2})$$

Since Table (2) includes the estimated values of the first of these components, estimates of the other two components are shown in Table (3).

Finally, it is to be emphasized that the estimated magnitudes of components of change should not be viewed strictly as measures of causal effects of the corresponding factors on labor force growth. Because of possible interrelations between these factors directly or through other intermediate variables, the exact magnitudes representing such causal effects are indeterminate. Nevertheless, the estimated components may be considered as measures of the relative degrees of influence which the given factors exert on the growth of labor force.

4.—CONCLUDING REMARKS

This paper describes an analyticall scheme presenting the possibilities of using standardization as a tool for temporal component analysis. The application of the scheme to labor force data gives indicative measures of the relative influence of changes in population size, age structure and age-specific activity rates on the growth of Egyptian labor force.

The results of such application clearly indicate that population growth dominated all other factors as a determinant of the labor force size. This fact is of wider pertinence to countries characterized by high rates of population growth. In such countries, the contribution of rapidlly growing population outweights the effects of changes in the level of participation in economic activities when the latter depends on the slow changes in social, economic and political factors.

Secondly, although the contribution of changes in the extent of participation in economic activities was small compared to that of population growth, it is still of economic significance. Furthermore, the relative influence of changes in age-specific activit yrates was somewhat more important than that of changes in the age structure of the population.

It is noteworthy that in addition to the negative effects of its changes on labor force size during the 1947/60 intercensal period, the youthful age structure has been a major factor responsible for an overall low level of crude activity rate, implying a heavy load of dependency. The youghfulness of the age structure throughout the period for which data are available is primarily the result of the high and relatively constant level of fertility.

Therefore, a significant decline in fertility is necessary for moderating the pressure on the labor market through the resulting reduction in the rates of population growth ; and for alleviating the load of dependency through its favorable effect on the age structure of the population.

ACKNOWLEDGEMENT

Appreciative recognition is given to Mrs. Nour Hamed for her help in most of the calculations included in this paper.

REFERENCES

1. DURAND, J. D., *The Labor Force in the United States, 1890-1960* (New York : Social Science Research Council, 1948).

2. DURAND, J. D. and A. R. MILLER, *Methods of Analysing Census Data on Economic Activities of the Population* (U. N., Department of Economic and Social Affairs, ST/SOA/Ser. A/43, 1968).

3. JAFFE, A. J., *Handbook of Statistical Methods for Demographers* (Washington : U. S. Government Printing Office, 1960).

4. KITAGAWA, E. M., «Standardized Comparisions in Population Research», *Demography*, Vol. 1, 1964.

5. NASSEF, A. E., *The Egyptian Labor Force : Its Dimensions and Changing Structure 1907-1960*, Unpublished Ph.D. Dissertation University of Peunsylvania, 1969.

6. WOLFBEIN, S. L. and A. J. JAFFE, «Demographic Factors in Labor Force Growth», *American Sociological Review*, Vol. XI, No. 4, August 1946.