RACE-SEX DISCRIMINATION INDICES FOR NON-METROPOLITAN STATE ECONOMIC AREAS

G. K. Kripalani, Western Michigan University

INTRODUCTION

One major group of factors influencing internal migration decisions stems from the "socio-cultural environment" of the areas of origin of potential migrants and their anticipated evaluation of corresponding elements in the areas of prospective in migration. Differences in internal net migration patterns may be analyzed by non-linear iterative least squares estimation procedure to isolate a component which reflects the impact of forces of relative race-sex discrimination in an area vis-a-vis the rest of the nation and which represents the net migration rate that would occur if relative economic opportunity factors in this area were as good as in the rest of the nation. This component of net migration is used as a basis of the definition of the race-sex discrimination index of a state.

The principal premise that underlies this study is that there are at least a few major independent variables affecting net migration and that some of these are non-measurable or non-observable, and that valid data series for such variables do not exist for use in empirical investigations. The method of analyses used is, therefore, designed to recognize and take into account this problem of nonobservability of some of the major explanatory variables. It is further recognized that net migration behavior patterns vary between the races, between the sexes and between age groups within each race-sex category. Consequently, there is need for stratification of an area's population into reasonably small homogeneous age, sex and race groups.

It is hypothesized that factors influencing internal net migration decisions of an age-sex-race group are of three categories:

1. Sub-area-related relative opportunity factors. These factors are the same for all age groups within a race-sex category. These relative opportunity factors are represented by an omnibus variable Z_t which is an index representing all relevant sub-area related relative opportunity factors. It is assumed that Z_t, which is the independent variable, is non-observable.

2. Age-related relative opportunity factors. These factors do not vary over a sub-area t in cross-section analyses, but vary between age groups within a race-sex category. Such age-related relative opportunity factors are denoted by a nonobservable index m_i where i refers to age group.

3. Race-sex-related relative opportunity factors. These factors do not vary over a sub-area in cross-section analyses or between age groups within a race-sex category. But these factors vary between race-sex categories and they reflect the impact of relative race-sex discrimination elements of the socio-cultural environment of the state in question.

In practical language the model separates net migration into three components: race-sex discrimination effect 'autonomous' component α' which is the same for all age groups within a race-sex category. This component would reflect the amount of net migration that would occur if $Z_t = 1$ and $m_i = 1$, that is, if net migration induced by relative opportunity and age-related factors were zero. It is this component which is defined to reflect "race-sex discrimination effect." There are two induced effects, one representing response to area-related omnibus independent variable Z_t representing relative opportunity factors and the other to age-related factors m_i.

For the purpose of this study, we may <u>define</u> the "race-sex discrimination" index of an area as the race-sex related component of internal net migration of that area (component α'). It is, however, recognized that the subset S_1 of elements of a socio-cultural environment S giving rise to what is called "race-sex discrimination" may consist of two types of elements--subset S_{11} consisting of elements which are the same for all age groups within a race-sex category (component α') and a subset S_{12} consisting of elements which vary between age groups within a race-sex category (component m_i). The latter component may reasonably be thought of, in given situations of being the result of "race-sex discrimination" and should appropriately be attributed to it.[1]

A real difficulty comes in the interpretation of the significance of the forces represented by m_i. Some of the forces underlying m_i may stem from those elements of the "socio-cultural environmental" complex as may be said to represent "race-sex discrimination," while it may legitimately be argued that some of these age-related factors stem from the fact that the assumption of a common index of relative opportunity facing all age-groups is unrealistic and that the index of relative opportunity is a function of both t and i. In such a situation,

Z_t would represent an average index of relative opportunity and a part of m_i would represent departures of the omnibus variable for the age group from the average Z_t for the category. Under these conditions, it would be necessary to identify the two subsets of the elements underlying m_i; those that relate to race-sex discrimination and those that reflect the situation that the index of relative opportunity is both age-related and time-related.

The "race-sex discrimination" index of an area may be viewed as a measure of the net migration effects of factors other than age and area related factors. Viewed thus, a comparative analyses of α''s may enable us to answer questions such as, (1) Are females "potentially" more migratory than males when the influences of age-time related factors are eliminated or equalized out or are Southern nonwhite males potentially more migratory than the Southern nonwhite females? and (2) Does the socio-cultural environment of a state discriminate against females or against nonwhites?

The significance of the positive or negative sign of α' may be clearly understood. Since total internal net migration of a color-sex category for the nation as a whole must be zero, it is easy to see that for each race-sex category:

$$\sum_s \alpha'_{cs} \cdot W_{cs} = 0 \text{ (c = WM, WF, NM, NF)}$$

where α'_{cs} equals race-sex discrimination index of category c in state s, and W_{cs} equals proportion of category c population in state s (as proportion of the total category population in the nation).

Consequently, index α' is an index of relative "discrimination," in relation to the average for the nation which is zero. A positive α' does not signify that "discrimination" however defined, is absent in that state; it only signifies that "discrimination," if any in this state, is less than the average for the nation as a whole.

The results reported in this paper are with reference to the third component of internal net migration which stems from forces which are constant over age groups within a race-sex category, but which vary between the four race-sex categories, namely, white males (WM), white females (WF), nonwhite males (NM), and nonwhite females (NF). These results are based on the analyses of nonmetropolitan state economic areas (NSEA) data for 1950-60 decade (Table 1). Results based on the analyses of 1950-60 net migration data for metropolitan state economic areas (MSEA) as units of study were reported in a paper read at the 1970 Detroit meetings.

The nonlinear iterative least squares estimation procedure developed by Johnston and Tolley [1] in their study "Supply of Farm Operators" was used to estimate values of model parameters α', and m_i and the nonobservable variable Z_t. The basic properties of this model were, however, crucially different in some respects from the properties of Johnston-Tolley model and consequently necessary modifications were introduced in evaluation procedures. Estimation method is dealt with in Section II.

Empirical Results

Some interesting results were: (a) Inter-sex comparisons showed that NSEA's which are predominantly rural areas are relatively more favorable to white males than white females. For nonwhites the evidence was not clear. (b) Inter-racial comparisons between white and nonwhite females provided no clear evidence; for males, however, there was some evidence that NSEA's are relatively favorable to white males than to nonwhite males. (c) The indices of relative race-sex discrimination were generally negative except for some interesting cases, viz., (i) white females for North Carolina, (ii) both nonwhite males and nonwhite females for Virginia and Mississippi; and (iii) nonwhite females for Florida. These exceptional cases indicate that there would be positive net inmigration of these categories into the NSEA's of these states, if the opportunity factors in these areas were as good as in the rest of the nation.

Section II

Model and the Method of Estimation

Consider the model:

$$Y'_{it} = a'_i + \beta_i Z'_t + e'_{it} \quad (1)$$

where primed variables represent logarithms of the original variables. The subscript i refers to the age group and subscript t refers to the sub-area in cross-section analyses. In this model the value of the dependent variable Y'_{it}, which depends upon i and t, is known while the independent variable on the right-hand side Z'_t which is nonmeasurable and hence unknown, is independent of i and is a function of t alone. Thus, the analysis of net migration data by the use of the above model would imply a critical assumption viz., that the net migration rates for different age groups i = 1, 2, ... i in a given nonmetropolitan state economic area (NSEA) are all functions of the same variable Z_t. This means that all the age groups in a given NSEA face the same index of relative opportunity. While the response coefficients for different age groups will be different, the critical assumption is that the independent non-observable omnibus variable to which these age groups are responding is the same.

Johnston and Tolley (1968) in their analysis of the supply of farm operators investigated a model of the above form using the NILES iterative procedure. An essential property of this iterative procedure is that the sequence of parameter estimates obtained at various iteration stages converges to the underlying parameter value not in the absolute sense, but in a relative sense. Hence, the character of parameter estimates obtained at the final stage is not cardinal but ordinal.

Consider the hypothesis: The number of persons of age i who will be staying in a nonmetropolitan state economic area t at the end of a decade (when the only cause of decrement or increment operating on the group is net migration) is a function of the index of relative opportunity Z_{it}, the exposed to risk of net migration E_{it} being the proportionality factor. This hypothesis gives rise to the multiplicative model of the form:

$$P_{it} = a_i\, E_{it}\, Z_{it}^{\beta_i} e_{it} \qquad (2)$$

where $$P_{it} = E_{it} + M_{it} \qquad (3)$$

The notation is:

E_{it} = Population of age group i exposed to risk of net migration in sub-area t during the decade, i.e., the population of age group i which would be in area t in the absence of any net migration.

M_{it} = Net migration of age group i into or out of sub-area t during the decade. M_{it} is positive when there is net inmigration and negative when there is net outmigration.

P_{it} = $E_{it} + M_{it}$ = Population of age group i in sub-area t at the end of the decade (if the only cause of decrement or increment operating on the group was net migration).

Y_{it} = $P_{it}/E_{it} = 1 + M_{it}/E_{it}$ = 'Survival' rate against net migration where P_{it} is the quantity of supply of population of age group i in sub-area t and E_{it} is the supply shifter.

Z_{it} = Nonobservable independent variable representing the index of relative opportunity.

β_i = Age group i's net migration response coefficient (elasticity) to relative opportunity index Z_{it} facing it.

a_i = Constant term for age group i.

e_{it} = Disturbance term.

(2) may be written as

$$Y'_{it} = a'_i + \beta_i\, Z'_{it} + e'_{it} \qquad (4)$$

where primed variables represent logarithms of the original variables. Note that (4) will become (1) if it is assumed that $Z_{it} = Z_t$ for all i.

Consider a given sex-color group in a NSEA in a state, say white males in a particular NSEA t in state s. This group is subdivided into nine age groups, 0-9, 10-14, 15-19, 20-24, 25-34, 35-44, 45-54, 55-64 and 65+ at the start of the decade 1950-60. The data regarding the number of net migrants (M_{it}) for each age group for each of the several NSEA's in each state for 1950-60 decade with the "appropriate" exposed to risk of net migration (E_{it}) were taken from the statistics published by the Economic Research Service of the U.S. Department of Agriculture (1965) in their Population Migration Report giving net migration numbers and rates by age, sex and color separately for metropolitan and nonmetropolitan state economic areas. The values of $Y_{it} = (M_{it} + E_{it})$ are given for $i = 1, 2, \ldots, 9$ and $t = 1, 2, 3, \ldots,$ depending upon the number of NSEA's in a state.

The procedure consists of taking logarithm of the functional relationship $Y_{it} = a_i\, Z_t^{\beta_i}\, e_{it}$ and minimizing the sum of squares of the random term e'_{it} ($e'_{it} = \log\, e_{it}$), the summation being a double summation over i and t. The process starts with an arbitrarily selected set of values for Z_t, $t = 1, 2, \ldots, t$, $(N_t = t)$.

$$(5)\quad SSE = \sum_i \sum_t e'^2_{it} = \sum_i \sum_t (Y'_{it} - a'_i - \beta_i\, Z'_t)^2$$

To obtain least squares estimates for parameters a'_i and β_i, we have the usual normal equations by taking partial derivatives of relation (5) with respect to a'_i and β_i, and solving these we have:

$$(6)\quad \hat{a}'_i = \frac{\sum_t Y'_{it} - \beta_i \sum_t Z'_t}{N_t}$$

$$(7) \quad \hat{\beta}_i = \frac{N_t \sum_t Y'_{it} Z'_t - \sum_t Y'_{it} \sum_t Z'_t}{N_t \sum_t Z'^2_t - \left(\sum_t Z'_t\right)^2}$$

Setting the partial derivative of SSE with respect to Z'_t equal to zero, we obtain an additional normal equation which when solved for Z'_t gives the least squares estimate of Z'_t in terms of a'_i, β_i and Y'_{it}.

$$(8) \quad \hat{Z}'_t = \left(\sum_i \hat{\beta}_i Y'_{it} - \sum_i \hat{a}_i \hat{\beta}_i\right) \Big/ \sum_i \hat{\beta}^2_i$$

Thus, the estimate of the value of Z'_t is obtained in terms of the estimated values of β_i and a'_i. To summarize, the process consists in starting with an assumed arbitrary set of values for Z'_t. Using this set, we arrive in the usual way, via normal equations, at the least squares estimates of a'_i and β_i. Then using these estimates of a'_i and β_i and the given values of Y'_{it}, we obtain the new set of estimates for Z'_t from equation (8). This completes the first iteration and we are at the second "stage" of the iteration process having given values of Y'_{it} as before but a new set of values of Z'_t. The process is repeated and the new set of Z'_t used to obtain a new set of a''s and β's in the first step of the second iteration and the process is repeated until estimates are approximately equal from the K^{th} and $(K+1)^{th}$ iteration. We will assume, for our purpose, certain convergence properties of this iterative procedure and the following results demonstrated by Johnston and Tolley (1968). The final estimates of a'_i, β_i and Z'_t depend upon the initial set of Z'_t's arbitrarily chosen as the starting point of the iterative procedure. It can, however, be shown that the ratios of β's, the ratio of differences for the Z'_t's and a certain linear function of a''s and β's have the property of convergence to the underlying value. Thus,

$$(9) \quad \lim_{s \to \infty} \frac{\beta^s_i}{\beta^s_j} = \frac{\beta_i}{\beta_j}$$

$$(10) \quad \lim_{s \to \infty} \frac{Z'^s_1 - Z'^s_j}{Z'^s_1 - Z'^s_2} = \frac{Z'_1 - Z'_j}{Z'_1 - Z'_2}$$

$$(11) \quad \lim_{s \to \infty} \frac{a'^s_\ell \beta^s_k - a'^s_k \beta^s_\ell}{Y'_{\ell t} \beta^s_k - Y'_{kt} \beta^s_\ell} = \frac{a'_\ell \beta_k - a'_k \beta_\ell}{Y'_{\ell t} \beta_k - Y'_{kt} \beta_\ell}$$

where $Y'_{\ell t}$ and Y'_{kt} are the known values of the dependent variable and s refers to the parameter estimates at the s^{th} iterative stage.

An important assumption underlying Johnston's model ($Y'_{it} = a'_i + \beta_i Z'_t + e'_{it}$) was that the variable Z is independent of i and depends on t only; that is, all age groups within a color-sex category in a NSEA faced the same index of relative opportunity. The index is thus assumed to change over NSEA's in a cross-section analysis, but it does not vary from age group to age group within a color-sex category in a given NSEA. It is proposed to relax this assumption and to regard the variable Z as a function of both i and t, and to replace it by a less exacting assumption that the index of relative opportunity varies from age group to age group but the ratio of any two Z's is fixed over NSEA's, i.e., it does not change over NSEA's in a cross-section analysis (or over time in a time series analysis). Mathematically, this is equivalent to assuming

$$(4.1) \quad Z_{it} = k_i Z_t \quad (i = 1, 2, \ldots i, t = 1, 2, \ldots t)$$

Hence,

$$(4.2) \quad \frac{Z_{it}}{Z_{jt}} = \frac{k_i}{k_j} = f(i, j)$$

It is important to note that k_i's are relative and it will be valid to regard $k_i = 1$ for $i = i_o$ and express all other k_i's ($i = 0, 1, \ldots i, i \neq i_o$) in relation to $k_i = 1$ for $i = i_o$. The condition that k_i's are relative enables us to put a constraint on k_i's, e.g. $\prod_{i=1}^{i} k_i = 1$. Z_t is a sort of average of individual Z_{it}'s depending upon the constraint imposed on k_i's. If $\prod_{i=1}^{i} k_i = 1$, $\prod_{i=1}^{i} Z_{it} = (Z_t)^i$ or Z_t is the unweighted geometric average of Z_{it}'s.

The model may be written as:

$$(4.3) \quad Y_{it} = \alpha_i Z_{it}^{\beta_i} e_{it}$$

Substituting $Z_{it} = k_i Z_t$ and taking logarithm, we have

$$(4.4) \quad Y'_{it} = \alpha'_i + \beta_i (k'_i + Z'_t) + e'_{it}$$

$$(4.5) \quad = (\alpha'_i + \beta_i k'_i) + \beta_i Z'_t + e'_{it}$$

The problem now is to minimize $S = \sum\sum_{it} e'_{it}{}^2$.

$$(4.6) \quad S = \sum\sum_{it} (Y'_{it} - \alpha'_i - \beta_i k'_i - \beta_i Z'_t)$$

Taking partial derivatives with respect to the unknown parameters α'_i, β_i and k'_i, (assuming the initial arbitrary set of values for Z'_t) and equating the expressions to zero to obtain the normal equations, we have:

$$(4.7) \quad \frac{\partial S}{\partial \alpha'_i} = \sum_t (Y'_{it} - \alpha'_i - \beta_i k'_i - \beta_i Z'_t) = 0$$

$$(4.8) \quad \frac{\partial S}{\partial \beta_i} = \sum_t (Y'_{it} - \alpha'_i - \beta_i k'_i - \beta_i Z'_t)(k'_i + Z'_t) = 0$$

$$(4.9) \quad \frac{\partial S}{\partial k'_i} = \sum_t (Y'_{it} - \alpha'_i - \beta_i k'_i - \beta_i Z'_t)\beta_i = 0$$

Now a difficulty arises in the solubility of this system. Equations (4.9) are the same as equations (4.7); and equations (4.8) simplify to:

$$\sum_t (Y'_{it} - \alpha'_i - \beta_i k'_i - \beta_i Z'_t)\ Z'_t = 0$$

because the terms $k'_i \Sigma_t (Y'_{it} - \alpha'_i - \beta_i k'_i - \beta_i Z'_t)$ vanish by virtue of (4.7). We are thus left with the following normal equations:

$$(4.10)\quad \alpha'_i \sum_t 1 + \beta_i k'_i \sum_t 1 + \beta_i \sum_t Z'_t = \sum_t Y'_{it}$$

$$(4.11)\quad \alpha'_i \sum_t Z'_t + \beta_i k'_i \sum_t Z'_t + \beta_i \Sigma Z'^2_t = \sum_t Y'_{it} Z'_t$$

that is,

$$(4.12)\quad (\alpha'_i + \beta_i k'_i) \sum_t 1 + \beta_i \Sigma Z'_t = \sum_t Y'_{it}$$

$$(4.13)\quad (\alpha'_i + \beta_i k'_i) \sum_t Z'_t + \beta_i \Sigma Z'^2_t = \sum_t Y'_{it} Z'_t$$

The solutions are:

$$(4.14)\quad \alpha'_i + \beta_i k'_i = \frac{1}{\Delta}\left[\sum_t Z'^2_t \cdot \sum_t Y'_{it} - \sum_t Z'_t \sum_t Y'_{it} Z'_t\right]$$

$$(4.15)\quad \beta_i = \frac{1}{\Delta}\left\{-\sum_t Z'_t \sum_t Y'_{it} + \sum_t 1 \cdot \sum_t Y'_{it} Z'_t\right\}$$

where

$$(4.16)\quad \Delta = \sum_t 1 \cdot \sum_t Z'^2_t - \left(\sum_t Z'_t\right)^2$$

Denote

$$(4.17)\qquad a'_i = \alpha'_i + \beta_i k'_i$$

Then

$$(4.18)\qquad k'_i = (a'_i - \alpha'_i)/\beta_i$$

After the iterative procedure has run to termination, the estimates for a'_i and β_i are obtained. To solve this model, we have to solve system (4.18) containing i equations and $2i$ unknowns k'_i and α'_i.

To make the system solvable, two alternatives may be considered by imposing constraints on the model:

$$(4.19)\quad \text{Case I: } \alpha'_1 = \alpha'_2 = \ldots = \alpha'_i = 0$$

$$(4.20)\quad \text{Case II: } \alpha'_1 = \alpha'_2 = \ldots = \alpha'_i = \alpha' \text{ and } k'_1 + k'_2 + \ldots + k'_i = 0$$

The rationale for a constraint on α'_i is dictated by an important consideration arising from the arbitrary nature of the estimates of these parameters. As Johnston has shown, the estimated values of α'_i's depend upon the arbitrary set of values chosen as the starting Z_t's for the iterative process. The regression line gives the intercept α'_i on the Y'_{it} axis when $Z'_{it} = 0$. On theoretical grounds, however, it may be argued that when $Z'_{it} = 0$, the component of net migration induced by relative opportunity factors will be zero; but an autonomous component of net migration may, however, still occur on account of factors other than relative opportunity factors which are assumed to have the same non-zero impact on all age groups within a color-sex category in a state. In the terminology of this study, this autonomous component represents the index of race-sex discrimination. Consequently, it is Case II which conforms to the hypothesis developed in this study.

The rationale for Case II may be argued this way. The quantities k'_i's represent relative magnitudes and hence one constraint can be imposed on k'_i's. Let $\sum_i k'_i = 0$ be assumed. Theoretical considerations in a study may warrant provision in the model for a constant term which will be the same for all age groups within a color-sex category in a state and which would reflect the impact of race-sex discrimination. Such an effect can be provided for in the model by the introduction of the parameter α' independent of i and t. This is equivalent to the assumption $\alpha'_1 = \alpha'_2 = \ldots = \alpha'_i = \alpha'$ though the significance of common α' is different from the significance of different a'_i's when they were included in the model. The alteration of the model reduces the number of parameters to be estimated to $(i + 1)$. The number of equations i and the constraint $\sum_i m'_i = 0^2$ give $i + 1$ equations, thus making the system solvable. The new model designated as the Race-Sex Discrimination Effect Autonomous Component Model is:

$$(4.25)\qquad Y_{it} = \alpha Z_{it}^{\beta_i} e_{it}$$

Putting $Z_{it} = m_i Z_t$, taking logarithm and simplifying this assumes the form

$$(4.26)\qquad Y'_{it} = \mu'_i + \beta_i Z'_t + e'_{it}$$

where

$$(4.27)\qquad \mu'_i = \alpha' + \beta_i m'_i .$$

The model thus has the same basic form as that of Johnston's model. The iterative procedure yields estimates of μ'_i, β_i and Z'_t where the value of μ'_i will be the same as that of a'_i in Johnston's model. Using the constraint $\prod_i m_i = 1$ or $\Sigma m'_i = 0$, and simplifying, we have:

$$(4.28)\qquad m'_i = (a'_i - \alpha')/\beta_i$$

and

$$(4.29)\qquad \alpha' = \sum_i (a'_i/\beta_i) / \sum_i (1/\beta_i)$$

$$(4.30)\qquad = \sum_i a'_i c_i / \sum_i c_i$$

where

$$c_i = 1/\beta_i$$

α' is thus the weighted arithmetic mean of a'_i's the system of weights being $c_i = 1/\beta_i$, the reciprocal of β_i. α' may be written as $\alpha' = \bar{a}'_i(c_i)$ to denote the weighted average of a'_i's, the weights being c_i.

Now

(4.31) $m'_i = (a'_i - \alpha')/\beta_i$

(4.32) $= c_i \, (a'_i - \alpha')$

(4.33) $= k'_i - c_i \alpha'$

Relation (4.33) is the relationship between m'_i of Case II and k'_i of Case I.

In terms of the parameters of the initial multiplicative model, viz.

$Y_{it} = \alpha m_i^{\beta_i} Z_t^{\beta_i} e_{it}$, we have $\mu_i = \alpha m_i^{\beta_i}$

that is,

$$(4.34) \quad m_i = \left(\frac{\mu_i}{\alpha}\right)^{1/\beta_i} = \left(\frac{\mu_i}{\alpha}\right)^{c_i}$$

The constraint $\sum_i m'_i = 0$ is equivalent to $\prod_i m_i = 1$. Therefore, $1 = \prod_i m_i =$

$$\prod_i \left(\frac{\mu_i}{\alpha}\right)^{c_i} \quad \text{or} \quad \alpha^{\sum_i c_i} = \prod_i \mu_i^{c_i} \quad \text{or}$$

$$(4.35) \quad \alpha = \left[\prod_i \mu_i^{c_i}\right]^{1/\sum_i c_i}$$

i.e. α is the weighted geometric mean of μ_i's (or a'_i) the system of weights being c_i, reciprocal of β_i. α being known, m_i may be estimated by the relation $m_i = (\mu_i/\alpha)^{c_i}$.

Our object is to know the estimates of α (or α') and m (or m ') in the final iteration which serves as the solution. It is not necessary to estimate α and m_i at intermediate stages in the iterative procedure. The iterative procedure for the Race-Sex Discrimination Effect Autonomous Component Model is the same as that for Johnston-Tolley model. It is only when the estimates of β_i and a'_i (or μ_i in Case II) are obtained in the final iteration of the procedure that values of α' and m'_i may be estimated by the relationships developed.

Like β and Z parameters, m' can be estimated in relative terms only. In the case of m''s, the invariant quantity is given by

$$(4.38) \quad \frac{(Y'_{jt}/n_{ji} - Y'_{it}) - \alpha\,(1/n_{ji} - 1)}{(Y'_{\ell t}/n_{\ell i} - Y'_{it}) - \alpha\,(1/n_{\ell i} - 1)} = \frac{m'_j - m'_i}{m'_\ell - m'_i}$$

Hence under certain assumptions when either n_{ji} and $n_{\ell i}$ are nearly equal or the quantities $\alpha\,(1/n_{ji} - 1)$ and $\alpha\,(1/n_{\ell i} - 1)$ are small in relation to the respective terms in the numerator and the denominator of (4.38), the invariant property of $\dfrac{m'_j - m'_i}{m'_\ell - m'_i}$ would remain and m'_j and m'_ℓ could be compared in the ordinal sense.

Strictly speaking, α's for two color-sex categories are not comparable even in the ordinal sense, due to change of origin in each case, since the estimated values come from two different regressions and each α' has its own specific invariant function. Under certain conditions, however, if the Z's between states are in fact similar, the comparisons between α's for two color-sex categories are valid.

Appendix I. Race-Sex Discrimination index, by State, based on cross-section analyses of net outmigration data for Non-metropolitan State Economic Areas, 1950-60 decade.

Region/State	White Male	White Female	Nonwhite Male	Nonwhite Female	$H_{1.} \overset{D}{=}$ (2)-(3)	$H_{2.} \overset{D}{=}$ (4)-(5)	$H_{.1} \overset{D}{=}$ (2)-(4)	$H_{.2} \overset{D}{=}$ (3)-(5)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
New England					(2, 0)			
Maine	.0088	-.0316			1			
New Hampshire	.0226	-.0272			1			
Middle Atlantic					(1, 1)			
New York	-.0140	-.0163			1			
Pennsylvania	-.0115	-.0042			0			
East No. Central					(2, 1)			
Indiana	-.0216	-.0269			1			
Illinois	-.0249	-.0155			0			
Wisconsin	-.0002	-.0137			1			

Appendix I. (Continued)

Region/State	White		Nonwhite		$H_{1.}^{D=}$	$H_{2.}^{D=}$	$H_{.1}^{D=}$	$H_{.2}^{D=}$
	Male	Female	Male	Female	(2)-(3)	(4)-(5)	(2)-(4)	(3)-(5)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
West No. Central					(2, 4)	(2, 0)	(0, 1)	(1, 1)
Minnesota	-.0050	-.0100	.0400	-.1003	1	1	0	1
Iowa	-.0120	-.0060			0			
Missouri		-.0133	.0913	-.0031		1		0
North Dakota	-.0321	-.0205			0			
South Dakota	-.0146	-.0109			0			
Nebraska	-.0244	-.0008			0			
Kansas	-.0057	-.0067			1			
South Atlantic and D. C.					(2, 3)	(3, 3)	(3, 2)	(3, 2)
Virginia	-.0100	-.0056	.0031	.0015	0	1	0	0
West Virginia	-.0040	-.0051	-.0147	-.0172	1	1	1	1
North Carolina	-.0153	.0071	-.0170	-.0132	0	0	1	1
South Carolina	-.0240	-.0113	-.0450	-.0351	0	0	1	1
Georgia	-.0097	-.0121	-.0021	-.0064	1	1	0	0
Florida			-.0161	.0064		0		
East So. Central					(1, 1)	(2, 2)	(1, 1)	(1, 2)
Kentucky			-.0161	-.0080		0		
Tennessee		-.0085	-.0157	-.0785	1			
Alabama	-.0149	-.0155	-.0710	-.0047	1	0	1	0
Mississippi	-.0236	-.0069	.0091	.0071	0	1	0	0
West So. Central					(1, 3)	(1, 3)	(2, 2)	(3, 1)
Arkansas	-.0363	-.0362	-.0288	-.0200	0	0	0	0
Louisiana	-.0110	-.0115	-.0425	-.0316	1	0	1	1
Oklahoma	-.0255	-.0104	-.1548	-.0405	0	0	1	1
Texas	-.0453	-.0095	-.0222	-.0577	0	1	0	1
Mountain					(1, 3)			(1,0)
Montana	-.0212	-.0196		-.0680	0			1
Idaho	.0275	.0057			1			
Wyoming	-.0605	-.0135			0			
Colorado	-.0350	.0579			0			
Pacific					(0, 1)			
Washington	-.0166	-.0111			0			
Oregon	4.5990							
United States Total					(12, 17)	(8, 8)	(6, 6)	(8, 7)

Notes: An explanation is necessary for columns (6) through (9). Let a'_{11}, a'_{12}, a'_{21}, and a'_{22} refer to the values of race-sex discrimination index a', relating to white male, white female, nonwhite male and nonwhite female categories, respectively. (The first subscript refers to race, 1 for whites and 2 for nonwhites; the second subscript refers to sex, 1 for males and 2 for females.) Let us define:

$$H_{1.}\underset{=}{D} \quad a'_{11} - a'_{12} = 1 \text{ if } a'_{11} > a'_{12}$$
$$= 0 \text{ if } a'_{11} < a'_{12}$$
$$H_{2.}\underset{=}{D} \quad a'_{21} - a'_{22} = 1 \text{ if } a'_{21} > a'_{22}$$
$$= 0 \text{ if } a'_{22} < a'_{22}$$
$$H_{.1}\underset{=}{D} \quad a'_{11} - a'_{21} = 1 \text{ if } a'_{11} > a'_{21}$$
$$= 0 \text{ if } a'_{11} < a'_{21}$$
$$H_{.2}\underset{=}{D} \quad a'_{12} - a'_{22} = 1 \text{ if } a'_{12} > a'_{22}$$
$$= 0 \text{ if } a'_{12} < a'_{22}$$

It will be observed that $H_{1.}$ involves a comparison between the sexes among whites. $H_{1.}$ gets value 1 if a' for white males is greater than a' for white females. Value 1 for $H_{1.}$ signifies that the autonomous part of net migration pull or push, as the case may be, is greater for white males than for white females in the particular state. Similarly, $H_{2.}$ involves a comparison between the sexes in the case of nonwhites; $H_{.1}$ involves a comparison between males among the races and $H_{.2}$ involves a comparison between white females and nonwhite females.

Footnotes

*Important results based on data for metropolitan state economic areas (MSEA) as the units of study were reported in a paper read at the 1970 Detroit meetings and published in the Proceedings Volume of the Social Statistics Section.

Acknowledgment

**Main work on this research was done while the author worked at the N. C. State University on a National Science Foundation Project. "Area Population Adjustments in Relation to Economic Activity." The author gratefully acknowledges the help he received from C. E. Bishop, G. S. Tolley, C. H. Hamilton, and F. E. McVay. Results reported here are contained in author's report to NSF. However, work on the interpretation and significance of race-sex discrimination effect component of internal net migration has been done during author's stay at Western Michigan University.

[1] I am indebted to my colleague Louis Junker for suggesting that the variable m_i may represent and capture age-related factors of "race-sex discrimination" and that "race-sex discrimination" index need not necessarily be totally described by component α'.

[2] Since the estimated values of k'_i in Cases I and II leading to modified versions of Johnston's model be different, the problem of distinguishing as between them will arise. Hence, in Case II, we denote the k parameters by the letter m.

Principal References

1. Johnston, W. E. and Tolley, G. S., 1968. "The Supply of Farm Operators." Econometrica Vol. 36, No. 2 (April, 1968), p. 365-382.

2. Lyttkens, Ejnar, 1965. 'On the Fixed-Point Property of Wold's Iterative Estimation Method for Principal Components,"-- a paper presented at the International Symposium on Multivariate Analysis at University of Dayton, Dayton, Ohio, June 14-19, 1965.

3. Wold, Herman, 1965. "Nonlinear Estimation by Iterative Least Squares;" unpublished paper, Institute of Statistics, University of Uppsala, Sweden.

4. U. S. Department of Agriculture, 1965. States, Counties, Economic Areas and Metropolitan Areas, Volume I. Population-Migration Report; Net Migration of the Population, 1950-60, by Age, Sex, and Color. U. S. Government Printing Office, Washington, D. C.