R E P O R T   R E S U M E S

ED 016 287 EA 001 059

ANALYZING DIFFERENCES IN THE GROWTH OF TWO RATIOS.

BY- ZABROWSKI, EDWARD K.
NATIONAL CENTER FOR EDUCATIONAL STATISTICS (DHEW)

REPORT NUMBER TN-1 PUB DATE 24 MAY 66
EDRS PRICE MF-$0.25 HC-$0.48 10P.

DESCRIPTORS- *GROWTH PATTERNS, INCOME, *MODELS, *METHODOLOGY, *DATA ANALYSIS, *STATISTICAL STUDIES, DISTRICT OF COLUMBIA,

THIS NOTE OFFERS A MEANS FOR ANALYZING THE DIFFERENCES IN THE LEVELS OF TWO RATIOS OVER A PERIOD OF TIME WHEN THOSE RATIOS COME FROM EMPIRICAL DATA. THE PARTICULAR EXAMPLE USED PERTAINS TO WHITE-NONWHITE PERSONAL INCOME DIFFERENTIALS OVER A TIME INTERVAL, BUT IT IS NOTED THAT THIS TYPE ANALYSIS CAN BE USED IN OTHER CIRCUMSTANCES. (HW)

ED016287

Working Paper
Does Not Reflect
Official Policy of
The Office of Education

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

NATIONAL CENTER FOR EDUCATIONAL STATISTICS
Division of Operations Analysis

ANALYZING DIFFERENCES IN THE GROWTH OF TWO RATIOS

by

Edward K. Zabrowski

Technical Note
Number 1

May 24, 1966

EA 001 059

NATIONAL CENTER FOR EDUCATIONAL STATISTICS
Alexander M. Mood, Assistant Commissioner

DIVISION OF OPERATIONS ANALYSIS
David S. Stoller, Director

ERIC

ANALYZING DIFFERENCES IN THE GROWTH OF TWO RATIOS

Edward K. Zabrowski

Purpose and Background

This note offers a means for analyzing the differences in the levels of two ratios over a period of time, when those ratios come from empirical data. For example, in a study of white-nonwhite personal income differentials over a time interval, one might postulate that the differences between the respective levels are due to four factors:

1. Differences present at the beginning of the time interval (intercept-level difference);
2. Differences in the respective rates of growth of total income (numerator);
3. Differences in the respective rates of growth in the numbers in the cell (denominator);
4. Covariance in (2) and (3).

The solution to this problem was sought because it may be found necessary at some point in the future to attempt to put a dollar measure on white-nonwhite discrimination. The most probable course of data for such a study would be some type of per capita income, preferably adjusted for differences in productivity and differential living costs, at a minimum. Some work along this line has been undertaken by Becker, but the main reference work was not conveniently available.[1] Therefore it was decided to explore the problem directly.

[1] Becker /1/, pp. 96-100, presents a measure of white-nonwhite discrimination which is derived from relationships between the rates of return and costs of education for the two racial groups. A more complete discussion should be found in his Economics of Discrimination, Chicago, 1957.

The basis of the problem is shown in the chart below, with the following symbols representing the specific variables:

Y = numerator of the ratio of the variable of primary interest, e.g., total nonwhite income in a stratum;

N = the denominator of the ratio of the variable of primary interest, e.g., the total number of nonwhites in a stratum;

$\frac{Y_t}{N_t}$ = the actual level achieved by the ratio of primary interest after t periods, e.g., nonwhite per capita income in period t;

Y' = the numerators in the ratio of the "reference" variable, i.e., the variable against which the variable of prime interest is being compared: e.g., total white income in a stratum;

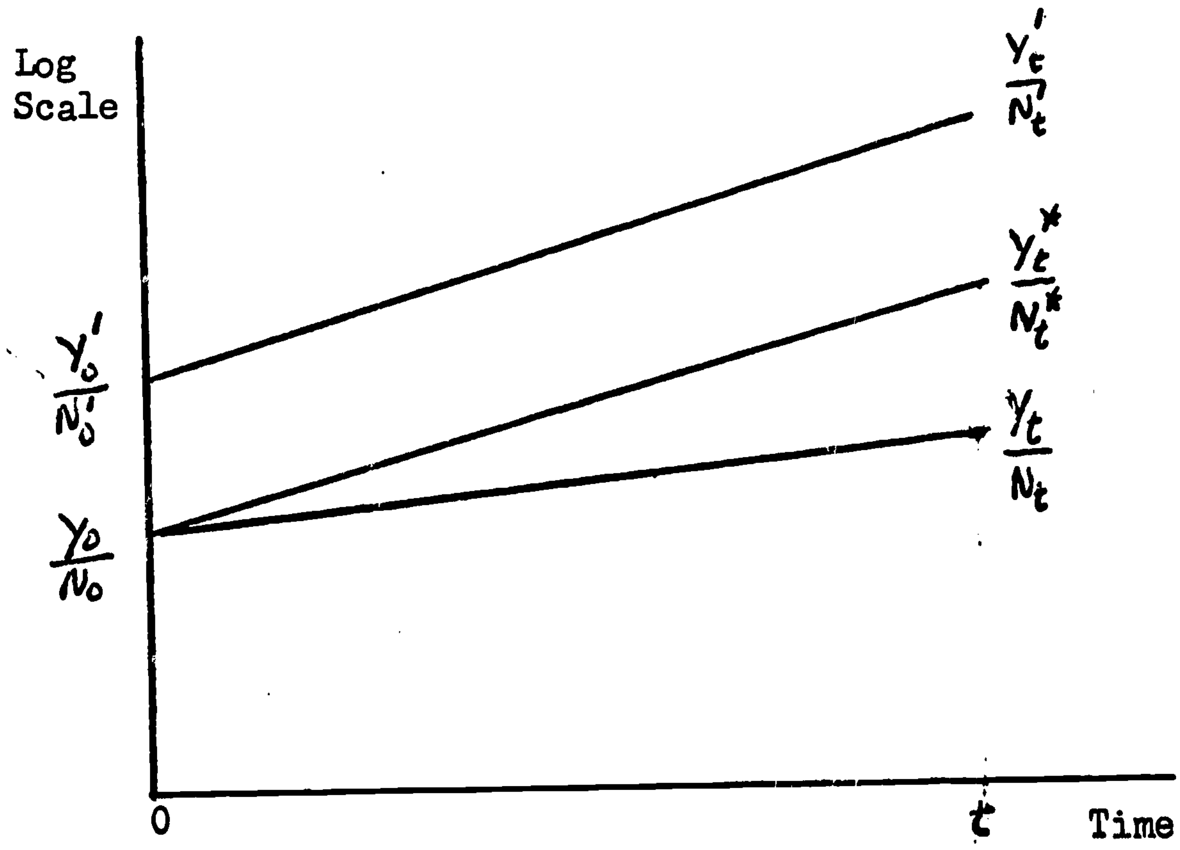
N' = the denominator corresponding to Y' in the reference ratio, e.g., the total number of Whites in a stratum;

$\frac{Y'_t}{N'_t}$ = the level of the reference ratio after t periods, i.e.,

$$\frac{Y'_t}{N'_t} = (1 + g)\left(\frac{Y'_o}{N'_o}\right);$$

$\frac{Y^*_t}{N^*_t}$ = the theoretical ratio representing the level of $\frac{Y_t}{N_t}$ had it grown at the same rate as did Y'_o/N'_o, i.e.,

$$\frac{Y^*_t}{N^*_t} = \frac{Y_o}{N_o}(1 + g).$$



From the chart it can be seen that, had $\frac{Y_o}{N_c}$ grown at the same rate as did $\frac{Y'_o}{N'_o}$, that is, from $\frac{Y'_o}{N'_o}$ to $\frac{Y'_t}{N'_t}$, the proportionate difference in their levels would have been the same as at t = 0, since the level would have been $\frac{Y^*_t}{N^*_t} = \frac{Y_o}{N_c}(1 + g)$. In that case, the difference in their levels would have been attributed to differences in their original levels, that is, to intercept differences.

In most instances, it is reasonable to assume that differences between two such time series will not be allocable solely to intercept differences. This is so because there is room for the numerators and denominators of the respective ratios to vary disproportionately, and there is also the possibility (in empirical series, at least) for a pair of ratios to be correlated with time, i.e., the ratios may covary to a significant degree, yielding a time trace such as that from Y_0/N_0 to Y_t/N_t.

In developing the simple method described below, it is necessary to consider all these factors and to combine them in one analytical framework that has a logical foundation. The foundation chosen has its roots in the equation of the total differential:

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy, \text{ or in empirical terms, } \Delta f = \frac{\partial f}{\partial x} \Delta x + \frac{\partial f}{\partial y} \Delta y + R.$$ [2/]

In nonmathematical terms, what this formula says is this: to find the change in a function of two variables, first find the change in the function contributed by each of the variables with the other held constant, then find the change contributed by their interaction. Conceptually, the total change in a function f caused by a variable, say x, independently of y is the rate of change in x (represented by the partial derivative $\frac{\partial f}{\partial x}$) times the amount of change in x (represented by dx).

The Analytical Model

The model chosen is stated as follows:

$$\frac{Y}{N} - \frac{Y^*}{N^*} = \left(\frac{Y}{N} - \frac{Y^*}{N}\right) + \left(\frac{Y}{N} - \frac{Y}{N^*}\right) + R = 1/N \ (Y - Y^*) + Y\ (1/N - 1/N^*) + R,$$

[2/] Most textbooks on differential calculus omit the final remainder term "R" in presenting the total differential because the true differential (dx say) is discussed instead of the difference value. Without this "R" term, the total differential for empirical purposes is a linear approximation which is considered to be suitable as long as Δx and Δy remain "small". Cf. /2/, p. 82, and /3/, pp. 161-164. The term is of interest in this problem because it is analogous to the covariation in the ratios.

that is, the difference between the ratio of the variable of interest and the theoretical or "expected" ratio (analogous to Δf) is heroically asserted to be due to the difference in the numerators with the denominator held constant (analogous to $\frac{\partial f}{\partial x} \Delta x$), plus the difference in the denominators with the numerator held constant (or $\doteq \frac{\partial f}{\partial y} \Delta y$), plus some remainder R, which in this case reflects covariance between the variables. This completes the weak analogy to the equation of the total differential.

The real key to the utility of this formula lies with R, since if R is large relative to the other two sources of difference, most of the potential advantages of partitioning will be lost. To test the sensitivity of the model, it is necessary to solve the equation shown above for R. With some wiggling around of the symbols, R can be shown to be:

$$R = (Y^* - Y)\left(\frac{1}{N} - \frac{1}{N^*}\right).$$

Thus, as N (and N*) get larger, R will tend to decrease. Similarly, the smaller the discrepancy between the actual and theoretical values of the numerators Y and Y*, the smaller the value of R, and hence, the more accurate will be the factoring of the differences into sources due to numerators and those due to denominators.

Going back to the chart, then, it can be seen that the difference between $\frac{Y'_t}{N'_t}$ and $\frac{Y_t}{N_t}$ is due to (a) the original difference multiplied by a factor (1 + g) which is shown by the point $\frac{Y^*_t}{N^*_t}$; and (b) a remaining portion which can be explained by deviations in the numerator and denominator from "expected" values, plus (c) a remainder representing covariance, as discussed above.

Within any particular time frame, of course, the difference between $\frac{Y'_t}{N'_t}$ and $\frac{Y^*_t}{N^*_t}$ is not relevant to many types of analysis as, for example in the case of discrimination. This is both an advantage and disadvantage. It would be particularly knotty in the case of discrimination, because while we have measured the amount that has taken place during the time interval, we could not state the level of entire discrimination at time t without outside information as to how to partition the difference between $\frac{Y'_t}{N'_t}$ and $\frac{Y^*_t}{N^*_t}$ into discriminatory and other factors.

Uses

When might this type of analysis be used? The most obvious application would be when the item in the denominator of the variable of primary interest exhibits a chronic relationship to the denominator of the reference variable. For example, assume that the number of employed Negroes is in nearly constant proportion to employed whites in periods of "normal" activity. Now assume that the number of Negroes employed suddenly increases. If the hypothesis can be entertained that the Negroes will be remunerated in accordance with the "expected" numbers, then it is possible to use this technique to factor the differences in Negro per capita income into that which is due to "pure" discrimination as reflected in income, and into that which is caused by variation in the numbers in the group. Another area where this technique may be useful is in input analysis, where, for example, different levels of skills may be used in more or less fixed proportions by employers of these skills. You may think of others.

Final Caveat

No claim to originality is made concerning this technique. No literature search has been conducted, since the potential sources of publishers of such a technique are very numerous. Because this type of derivation represents rather low-level creativity, it is highly probable that it has been done elsewhere at some time.

ERIC

References

1. Becker, Gary S.; Human Capital. New York: Columbia University Press, 1964.

2. Yamane, Taro; Mathematics for Economists. Englewood, California: Prentice-Hall, Inc., 1962.

3. Taylor, Angus E.; Advanced Calculus. Boston: Ginn and Company, 1955.

ERIC