EDRP No. 149
B. Okun
15/11/68

Note: Economic Development Research Papers are written as a basis for discussion in the Makerere Economic Research Seminar. They are not publications and are subject to revision.

FACTORS AFFECTING FERTILITY RATIO DIFFERENTIALS IN THE UNITED STATES METROPOLITAN AREAS, 1950 and 1960*

I. Introduction:

United States data have revealed a number of associations between fertility and socio-economic variables. It has been observed that non-whites are more fertile than whites, that fertility varies inversely with the level of women's education, and that fertility varies inversely with family income (1).

Although these associations are of interest, they do not explain fertility behaviour. From these simple associations, no conclusions can be reached concerning the factors which affect fertility. Because the variables which may explain fertility are intercorrelated, a simple correlation between fertility and another variable may be spurious. Consider, for example, the aforementioned negative association between fertility and women's education. The educational attainment of wives is undoubtedly positively related to that of husbands, and this in turn is positively related to family income. Thus, if family income has a negative effect on fertility, then it is possible that the negative relation between fertility and women's educational attainment is attributable to the family income variable.

This paper employs multiple regression analysis in order to assess the net effect of each of several explanatory variables on fertility. The coefficient of multiple determination will indicate how much of the variation in fertility is explained by the combined effects of all the independent variables included in the regression equation.

The data underlying the analysis are census figures obtained from the United States Censuses of 1950 and 1960. Two regression equations are fitted with 1950 and with 1960 cross-section data. A comparison of the two sets of regression coefficients will indicate the degree of consistency of the findings at the two points in times. A third regression equation is fitted in an attempt to explain changes in fertility between 1950 and 1960. This is of interest because it is well known that factors which explain cross-section differences may not succeed in explaining changes over time.

For 1960, the 101 largest Standard Metropolitan Statistical Areas (SMSA) (2) serve as the spatial units of analysis. These are all the SMSA's with more than 250,000 inhabitants. For 1950, the regression equation is fitted

.....2.

with the data for the 100 largest SMA's (3).

Table I summarizes the information pertaining to the metropolitan areas chosen for 1950 and 1960. It is worth emphasizing that, as indicated in column 8, the metropolitan areas whose data are used in this study accounted for 86% of the SMSA population in 1960 and 90% of the SMA population in 1950. The overwhelming importance of the metropolitan areas as locations of population growth between 1950 and 1960 should also be noted. For example, based on the 1967 definition of an SMSA, 88% of United States population growth between 1950 and 1960 was accounted for by the population growth in the SMSA's. The rate of population increase was 26.5% in the SMSA's, and only 5.6% in the non-metropolitan territory (5).

A_i = the ratio of women aged 20 to 29 to women aged 15 to 49 in metropolitan area i. It is well known that women in the 20 - 29 age group have significantly higher fertility rates than the other women in the 15 to 49 age group, and it is therefore hypothesized that A_i will have a positive effect on F_i.

N_i = the ratio of non-white women aged 15 - 49 to total women aged 15 - 49 in metropolitan area i. Although fertility rates for non-whites have consistently been above those for whites in the United States as a whole, it doesn't follow that this differential prevails for metropolitan areas -- the country-wide racial differential might be mainly attributable to the high fertility rates of Negroes living in the rural parts of the Southern states. This study will seek to determine what effect, if any, N_i has on F_i, among the metropolitan areas (9).

III The Regression Results:

The regression equations are fitted by ordinary least squares. Table II presents the regression results for 1950 and 1960.

TABLE II

REGRESSION RESULTS FOR 1950 AND 1960

1950: (1) $F_i = 37.19 \underset{(.25252)}{-.06541\, S_i} \underset{(.00599)}{-.01967\, D_i^*} \underset{(.011835)}{-.004353\, Y_i} + \underset{(3.312)}{1.498\, E_i}$

$\underset{(222)}{+1176\, A_i^*} \underset{(39.02)}{-\, 97.51\, N_i^*}$

$R^2 = .43$

1960: (2) $F_i = 52.94 \underset{(.3117)}{-.5590\, S_i} \underset{(.002985)}{-.006687\, D_i^*} \underset{(.00777)}{+.01978\, Y_i^*} + \underset{(5.4843)}{2.2777\, E_i}$

$\underset{(201)}{+\, 1201\, A_i^*} \underset{(44.62)}{-\, 33.41\, N_i}$

$R^2 = .44$

* Based on a two-tail test at the 5 percent level of significance, the regression coefficients with the asterisk are significantly different from zero.

Population Size and Density

A comparison of the regression coefficients reveals that for four of the six independent variables, the 1950 and 1960 findings are consistent. Looking first at the consistent findings, we observe that D_i has a significantly negative effect on F_i whereas no significant relationship between S_i and F_i is detected. For the economic reasons mentioned above, greater population density

.....6

appears to depress fertility. The frequently observed negative association between population size and fertility may be largely explained by the positive correlation between population size and population density. (10) There is no valid reason why the population size of a place, if the effect of density is "held constant", should have a significant effect on fertility.

Age-distribution of Women 15 - 49 Years Old

Another consistent finding for both equations is that A_i has a significantly positive effect on F_i. This is because women aged 20 - 29 have a significantly higher fertility rate than the other women aged 15 - 49. There is relatively greater fecundity among the 20 to 29 age-group, and this is coupled with a preference for bearing children during the 20 - 29 age-period. Thus, the finding of a significantly positive relationship between A_i and F_i is easily explained.

Level of Women's Education

The fourth instance of concurrence between the 1950 and 1960 regression equations is with regard to the education variable. No significant relationship between E_i and F_i is revealed, though both equations yield positive regression coefficients: thus, the hypothesis that E_i has a negative effect on F_i is clearly refuted. It is important to note that for this study's sample metropolitan areas, the simple coefficient of correlation between E_i and F_i is also positive and low, with $r = +.14$ for 1950 and $r = +.35$ for 1960.

It will be recalled that United States sample data indicate that there is a negative association between fertility and the level of women's education. Table III presents the United States pattern that prevailed in 1960.

TABLE III

Children Ever Born per 1000 Women 15 to 44 Years Old, Standardized for Age, 1960 (11)

Years of School Completed	Children Ever Born per 1000 Women aged 15 - 44, Standardized for Age
Elementary: Less than 8 years	2,316
8 years	2,088
High School: 1 to 3 years	1,982
4 years	1,625
College: 1 to 3 years	1,454
4 years or more	1,162

The reader will note that both the education variable and the fertility variable in Table III differ from E_i and F_i, and this may partly

.....7

account for the differences in findings. Table III suggests that the relationship between women's education and fertility is non-linear. Women with 1 to 3 years of high school have a fertility rate which is 86% of that of women with less than 8 years of elementary school education: thus, over this wide educational range, the variation in fertility is rather small. But, women with 4 years or more of college have a fertility rate which is only 59% of that for women with 1 to 3 years of high school, and 50% of that for women with less than 8 years of elementary school education. There are especially sharp declines in fertility between successive educational levels in two instances: where women pass from 1 to 3 years of high school to 4 years of high school (18% fertility decline), and where women pass from 1 to 3 years of college to 4 years or more of college (20% fertility decline). Perhaps the securing of a high school diploma in the first instance and the securing of a college degree in the second instance may cause a significant increase in the opportunity cost of bearing and of rearing children (where this results in the woman's exit from the labour force) compared to the opportunity cost for women in the immediately lower educational categories; this may contribute to the comparatively sharp declines in fertility for women with 4 years of high school and for women with 4 or more years of college education. In the case of the women with 4 or more years of college education, postponement of marriage may also be an important factor contributing to lower fertility. This discussion points to the conclusion that the relationship between women's education and fertility is rather complex and non-linear: thus, our attempt to find a linear relationship between fertility and median years of school completed for females 25 years old and over may have resulted in no significant regression coefficient because the true relationship may be non-linear and because our education variable is too crude.

Another factor that may partly explain the discrepancy between our findings and those in Table III relates to the choice of sample. Women living in non-metropolitan areas tend to have relatively high fertility and relatively low educational levels. Thus, they contribute to the negative association between fertility and women's educational levels in the United States as a whole. As this study is based on data taken from the larger metropolitan areas, any contribution that non-metropolitan living makes to the United States negative association between women's education and fertility does not affect our results. This may partly account for the absence of a negative correlation between F_i and E_i in this study, despite the probable existence of a negative correlation between E_i and F_i for the United States as a whole.

Median Male Income

Passing on to the relationship between median income for males 14 years old and over and fertility, we find that for 1950 the regression coefficient

.....8

is negative but not significant, whereas for 1960 the regression coefficient is significantly positive. These findings do not permit us to reach a firm conclusion as to the relationship between Y_i and F_i: the findings are inconsistent, and for 1960, the significance test is passed by a narrow margin. One may speculate that the swing to a positive regression coefficient between 1950 and 1960 resulted from a relative improvement in the capability of lower income families to control family size and reduce excess fertility. However, given the nature of the statistical evidence, it seems appropriate to reserve judgement on the relationship between Y_i and F_i until the forthcoming 1970 census data are available and analyzed.

Despite the inconclusiveness of our findings, they do indicate that the relationship between Y_i and F_i is not significantly negative. Thus, the negative association between Y_i and F_i revealed by the United States sample data, referred to earlier in the paper, is not upheld in the regression analysis, where other variables are taken into account and where the data are taken from the larger metropolitan areas. (12)

The Racial Factor in Fertility

The last independent variable that remains to be considered is N_i, the proportion of women aged 15 - 49 which is non-white. In both equations, the regression coefficient of N_i is negative, though it is significantly negative only for 1950. The simple coefficient of correlation between N_i and F_i, equal to -.05 for 1950 and -.01 for 1960, reveals no association between these two variables. Although non-white fertility in the United States exceeds that for whites, this racial differential may not exist for the larger metropolitan areas, especially after other factors are taken into account. As was suggested earlier, the high fertility rate of Negroes living in the rural South contributes to the country-wide racial differential in fertility, but has no bearing on racial differentials in metropolitan areas.

But what explains the negative effect of N_i on F_i for 1950? Within metropolitan areas, non-whites are relatively more concentrated in the central cities whereas whites live in relatively larger numbers in the suburban outer ring. As we have found that greater density has a negative effect on fertility among metropolitan areas, it is also reasonable to assume this relationship applies within metropolitan areas. This would tend to depress non-white fertility within metropolitan areas, and may partly account for the negative effect of N_i on F_i. Also, as survival rates of children under 5 are somewhat lower for non-whites, this would bias the non-white fertility ratio downward. Taking account of these qualifications as well as the absence of a significant relationship between N_i and F_i for 1960, we conclude that the racial composition of women of child-bearing age is not a significant factor in explaining fertility differentials among metropolitan areas.

.....9

IV Summary of Cross-section Findings:

The two linear regression equations, fitted with cross-section metropolitan area data for 1950 and 1960, reveal that:

(1) population size has no significant effect on fertility;

(2) population density has a significantly negative effect on fertility ;

(3) the age-composition of women of child-bearing age, as measured by the ratio of women aged 20 - 29 to women aged 15 - 49, has a significantly positive effect on fertility;

(4) no significant relationship exists between fertility and median years of school completed for women 25 years old and over; however, given the possibility that the relationship between women's education and fertility is non-linear, the education variable employed in this study may be too crude to detect the true relationship, if any exists;

(5) median income of males 14 years old and over has no significant effect on fertility for 1950, and a significantly positive effect for 1960, although it is possible that a fundamental change has taken place in the relationship between income and fertility, it is suggested that the forthcoming 1970 census data be analyzed before a firm conclusion is reached on this question;

(6) the racial composition of women of child-bearing age, as measured by the ratio of non-white women aged 15 - 49 to total women aged 15 - 49, has a significantly negative effect on fertility for 1950 and no significant effect on fertility for 1960; this study concludes that the racial factor is unimportant in explaining fertility differences among metropolitan areas, given the inconsistency of the above findings and the qualifications that non-white fertility may be biased downward because of lower survival rates of non-white children under 5, and may be further depressed because among the metropolitan area inhabitants, non-whites live in more densely populated conditions than whites.

It is interesting to note that the associations between fertility and other variables, based on United States sample data and referred to in our introductory remarks, provide no clues, or rather misleading clues, as to the factors which affect fertility differences among metropolitan areas. This emphasizes the point that simple correlation analysis is worthless in explaining fertility. Furthermore, when multiple regression analysis is employed, it is important to either select data from similar kinds of spatial units (e.g., metropolitan areas) or to include among the independent

....10

variables, some variable which explicitly measures the impact on fertility of place of residence, such as metropolitan versus non-metropolitan living.

V Changes in Fertility Between 1950 and 1960:

It has already been noted that factors which explain cross-section differences in fertility may not explain changes in fertility over time. The data on 1950 and 1960 fertility ratios permit a comparison to be made over time between 1945 to 1950 and 1955 to 1960. In the metropolitan areas selected for this study, the arithmetic mean fertility ratio is 399 for 1950 and 475 for 1960, resulting in a 19% increase in/fertility between the two time periods. (13) Between 1950 and 1960, mean population density increased by 40% and the mean value of the ratio of women aged 20 - 29 to women aged 15 - 49 decreased by 17%. Thus, changes in the mean values of D_i and A_i are not useful in explaining the increases in average fertility. Median income of males 14 years and over increased by 65% in current dollars. A comparable figure in constant dollars is not available, but is probably in the neighbourhood of about half the 65% increase. If the increase in average density contributed to rising rents which tend to reduce fertility, then the rise in real income may have been an offsetting factor. Percentage changes in the mean values of the other independent variables are small, and are probably not important in explaining the increase in average fertility.

With data from 89 metropolitan areas, an equation is fitted regressing absolute changes in fertility against absolute changes in the values of D_i, A_i, E_i, Y_i and N_i. (The change in S_i is not included because it is too closely related to the change in D_i, since land areas remained the same for most metropolitan areas.) We obtain a very low coefficient of multiple determination -- only 12% -- and none of the regression coefficients is significant except for the change in A_i which has a positive effect on F_i. It is clear that the set of independent variables which this study has found to be partly successful in explaining cross-section differences in fertility, has failed to explain the increases in fertility ratios between 1945 to 1950 and 1955 to 1960. Clearly, further research is required to explain the temporal changes in fertility rates. It will be especially important to explain why United States fertility rates in the second half of the 1950's were higher than in any period from 1920 to the present. (14)

* The author expresses thanks to Ronald Needleman for his valuable assistance in collecting and processing the data.

1) "Fertility of the Population; June 1964 and March 1962", Current Population Reports, Population Characteristics, (U.S. Department of Commerce: Washington, D.C., January, 1966), pp 1-4.

2) In the 1950 Census, these were referred to as Standard Metropolitan Areas (SMA), In the 1960 Census, the criteria for designating metropolitan areas were revised and the areas were designated as standard metropolitan statistical areas (SMSA). "The population residing in SMSA's constitutes the metropolitan population. Except in new England, an SMSA is a county or group of contiguous counties which contains at least one city of 50,000 inhabitants or more, or "twin cities" with a combined population of at least 50,000. In addition to the county or counties, containing such a city or cities, contiguous counties are included in an SMSA if, according to certain criteria, they are essentially metropolitan in character and are socially and economically integrated with the central city. In New England, SMSA's consist of towns and cities, rather than counties. East SMSA must include at least one central city, and the complete title of an SMSA identifies the central city or cities." Quote taken from Current Population Reports, Technical Studies, Series P - 23, No. 23, October 9, 1967, page 3.

3) Based on the 1950 definition of an SMA.

4) Sources: "Standard Metropolitan Statistical Areas in the United States as Defined On May 1, 1967 , with Population in 1960 and 1950", Current Population Reports, Series P - 23, No. 23, October 9, 1967, page 2.

5) Ibid., page 1.

6) Note that the fertility ratio, although a widely used measure of fertility, may not perfectly reflect cross-section differences in fertility because it is affected by: (1) differentials in mortality rates of children under 5 and of women aged 15 - 49, among the selected metropolitan areas; (2) differentials in undernumeration of children aged 0 - 4 years and of women aged 15 - 49, among the selected metropolitan areas; (3) migration of children aged 0 - 4 and women aged 15 - 49 during the 5 year period preceding the census. See B. Okun, Trends in Birth Rates in the United States since 1870 (Baltimore, John Hopkins Press, 1958), pp 20-25.

7) For the metropolitan areas selected in this study, the correlation coefficient between S_i and Y_i is -.34 in 1950 and -.24 in 1960. For an earlier period, Warren S. Thompson showed that in 1920 differentials in ratios of children to women were very great, and that the fertility ratio tended to vary inversely with the size of the community. See Warren S. Thompson, Ratio of Children to Women (Washington: U.S. Government Printing Office, 1931).

8) See, for example, Gary Becker, "An Economic Analysis of Fertility", Demographic and Economic Change in Developed Countries, Princeton University Press, 1960, pp 209-231.

9) Sources of data are from U.S. Census of Population Volumes. For 1950, see Volume I, Number of Inhabitants, Table 28, for data on S_i and Table 29 for data on D_i. See Volume II, Characteristics of the Population, Part I, Table 181 for data on E_i and Table 185 for data on Y_i. See Volume II, Parts 2-50, Table 33, for data on A_i, N_i, and F_i.

For 1960, see Volume I, Part A, Table 36 for data on S_i and Table 34 for data on D_i. See Part 1, Table 149 for data on Y_i and Table 288 for data on E_i. See Parts 2-52, Tables 114 and 20, for data on A_i, N_i, and F_i.

10) In this study, the coefficient of correlation between size and density is .55 and .40 for 1950 and 1960, respectively.

11) Source: "Fertility of the Population: June 1964 and March 1962", op.cit., p.15.

12) This study yields coefficients of correlation between Y_i and F_i equal to -.13 for 1950 and +.12 for 1960; each of the coefficients does not differ significantly from zero.

13) For the United States as a whole, there was an increase in fertility of 13% between 1946 to 1949 and 1956 to 1959. This is based on average annual fertility rates for the two periods, where fertility is measured by the number of births per 1000 women aged 15 - 44 years. See Natality Statistics Analysis, United States - 1963, National Centre for Health Statistics, Series 21, Number 8, March 1966, Table 1, page 2.

14) Ibid.