

ECONOMIC GROWTH CENTER

YALE UNIVERSITY

Box 1987, Yale Station
New Haven, Connecticut

CENTER DISCUSSION PAPER NO. 383

AGE OF INDIVIDUALS AND FAMILY COMPOSITION AS FACTORS UNDERLYING
THE DISTRIBUTION OF PERSONAL INCOME

T. Paul Schultz

September 1981

Notes: The author has benefited from the comments of Allen Kelly, Simon Kuznets and Theodore Schultz on a draft of this paper presented at the IUSSP workshop in Hawaii, April 6, 1981. They have no responsibility of course, for what remains. This work is partially supported by the United Nations Fund for Population Activities. Thomas Frenkel provided valuable research assistance.

Center Discussion Papers are preliminary materials circulated to stimulate discussion and critical comment. References in publications to Discussion Papers should be cleared with the author to protect the tentative character of these papers.

An inverse association is generally observed between inequality in the size distribution of personal incomes and the level of per capita income in a country (Kuznets, 1955). Several aspects of the demographic composition of populations may account for this association between economic development and aggregate income inequality. In the long run, modern economic growth may contribute to changing death and birth rates and social organization, which in turn modify the age structure of the population and the composition and size distribution of families. If a substantial share of aggregate income inequality is linked to these demographic features of populations, a decomposition of income inequality based on these features might help to understand the proximate origins of trends and cross sectional patterns in aggregate inequality. Moreover, it may be argued that the inequality associated with certain demographic features, such as the age structure, is not necessarily an indication of the degree of lifetime inequality among persons, and may, therefore, be tentatively excluded from welfare comparisons of economic inequality. Thus, the causal origins and the normative significance of aggregate inequality may be clarified by such decompositions.

This paper reports two approaches to decompose income inequality, approximated by the variance of the logarithms of income (log variance), into components associated, first, with the age structure of individuals with income, and second, by the number of adults and children (per adult) in families.

First, data for three countries are used to illustrate how variation in age structures may help to account for observed patterns of aggregate income inequality. The data for Colombia are then analyzed further to explore the relationship between family composition and family income. Two elements of family composition are distinguished--fertility and the decision of adults to share living arrangements. The number of children that parents want may respond to incomes, relative prices, and wage rates of parents; the relationship between fertility and adult incomes can be interpreted, in this context, as a simple demand equation, albeit one that is subject to bias by the omission of other factors affecting reproductive demands that are probably correlated with adult incomes. The propensity of adults to live together may be interpreted similarly as a choice of adults that is conditioned by their economic resources. It may alternatively be viewed as a production relationship linking the productive contributions of adult workers, who contribute differentially with the growth of family scale, to total family income. These static decompositions of income inequality provide suggestive explanations for how economic development may affect over time the distribution of personal incomes, and how the path of demographic transition modifies the rate of income growth and its personal distribution across a society and within a society across generations.

AGE STRUCTURE AND THE DISTRIBUTION OF PERSONAL INCOME

The age structures of populations differ substantially from country to country and within a single country over time. These differences reflect the level of and predominantly recent changes in birth and death rates. High birth rates yield a younger age structure in the long run, and low birth rates an older structure. Recent sharp declines in mortality rates in low income countries have been larger among infants and children than they have been for adults. This has had a similar effect on the age structure of these populations as would an increase in fertility, namely, increasing the rate of growth of the youngest age groups relative to older age groups. Most low income countries, therefore, have experienced a shift in their age structure, after World War II, with the share of children increasing. These relatively large surviving birth cohorts from the post-war period have in the 1970s entered the labor force and begun to earn income. In those countries that have experienced declines in fertility, the proportion of the population in the youngest age groups has, conversely, fallen and in time the age structure of the labor force will tend to become older.

The secular decline in mortality rates in high income countries has exerted a less pronounced effect on the age structure of these populations, because mortality was already in this century at a lower level and the decline was more uniformly distributed across age groups. But notable long swings in birth rates occurred in some high income countries, such as the United States after the 1920s, and perturbed age-structures. The rise in birth rates following World War II created a relatively large birth cohort to be absorbed into the labor force in the 1970s, whereas the Great De-

pression produced a shortfall in births and thus a relatively small cohort of labor force entrants to satisfy the growth in labor demands in the 1950s. Given the current variation across countries in age structures and our capacity to project future swings in these structures, it would seem useful to assess how age structures affect, directly and indirectly, the distribution of personal income, and how aggregate economic developments and individual behavior respond to and modify these effects of the age structure on measured income "inequality."

The Logarithmic Variance of Personal Incomes and Its Decomposition by Age

Several measures of inequality can be decomposed into elements associated with a particular population characteristic; the log variance can be resolved, as any variance can, into between and within group variance components as reported below.¹ Such decompositions are insightful if they distinguish between different sources of inequality with different implications for economic welfare or policy and if they clarify empirical regularities that can be interpreted as causal relationships.

The analysis in this section of the paper focuses on individual money incomes. Our aggregate measure of economic inequality, the log variance, $V(y)$, is resolved into three portions associated with (1) the age structure of the income recipient population, (2) the profile of incomes received on average by persons of different ages, and (3) the income inequality within these different age groups.

$$\begin{aligned}
 V(y) &= \sum_{j1} (\bar{y} - y_{1j})^2 \\
 &= \sum_j \frac{n_j}{N} [\bar{y} - \bar{y}_j]^2 + \sum_j (\bar{y}_j - y_{1j})^2, \quad (1)
 \end{aligned}$$

where y_{1j} is the natural logarithm of the i^{th} individual in the j^{th} age group with a positive income in the preceding time period, $\bar{y}_j$ is the mean of logarithmic incomes in the j^{th} age group, $\bar{y}$ is the overall mean of logarithmic incomes, n_j is the number of persons of age j with a positive income, and thus $N = \sum_j n_j$.

The Age Structure

The first term, n_j/N , is the weight or relative frequency of the age groups in the population of income recipients and can modify measured aggregate income inequality without necessarily affecting lifetime income opportunities of individuals. Intertemporal or cross country comparisons may be confounded by differences in age structure, and few empirical studies of income inequality have attempted to isolate or remove this demographic source of measured inequality. As with index numbers, there is no straightforward method to normalize adequately for variation in quantity weights (i.e., age structure), because the other two components of income inequality are likely to differ across observations. The broad variations in income inequality that are empirically documented generally parallel variation in age structures; though many other factors are probably involved in generating these patterns in inequality, the effects of age structure warrant further quantitative analysis.

For example, if a large fraction of young workers in a population increases measured aggregate income inequality, as appears to be the case in the United States, several empirical regularities first noted by Kuznets (1955, 1963) might be explained by variation in age structures. (1) In the advanced stages of industrialization and urbanization, particularly in the 20th century, a number of countries, including the United States, evidence declining income inequality. This pattern of change in inequality over time is consistent with the changes in age structure that accompanied the secular decline in fertility in these countries during this period. (2) Low income countries report today greater income inequality, by most summary measures, than do high income industrially advanced countries. Less developed countries have recently sustained higher levels of fertility than have the more developed countries and their consequent younger age structures could explain this cross country pattern in income inequality. (3) Some data suggest that inequality increased during the early stages of industrialization in the United States (Lindert and Williamson, 1976), and may also have increased recently in some low income countries, such as India. The earlier noted shifts in age structure in many low income countries stemming from the age pattern of mortality declines could account for some deterioration in measured income inequality in the current period. High fertility and immigration were sufficient in the United States to increase the ratio of men age 15 to 24 to all men 15 or older until about 1820. This ratio, corresponding to the proportion of youthful entrants to the labor force, has declined steadily since that time in the United States, from 38 percent in 1820 to 24 percent by 1940 (U.S. Bureau of Census, 1960).

The Age-Income Profile

The second component of income inequality in equation (1) is the difference between the age group logarithmic mean income and the overall population logarithmic mean income, squared. If equity is defined in terms of equality in the distribution of lifetime economic opportunities, appropriately discounted, then income differences by age need not represent inequitable variation in individual incomes, assuming of course that individuals experience the sequence of average incomes associated with each age interval in their lifetime. Individuals may decide to redistribute these earnings opportunities over their lifetime by means of investments in physical and human capital. According to this mechanism, age-earnings profiles are interpreted as a reflection of the level of schooling and post-school training and occupational experience that individuals acquire at an initial cost in anticipation of subsequent gain (Mincer, 1974).² Since these human capital investments tend to be concentrated at the outset of the life cycle, the greater the general level of these investments or the more highly educated the population, the more steeply upward sloping are age-earnings longitudinal profiles for a birth cohort. The time individuals allocate to earning income also varies systematically with age, displacing the life cycle profile of earnings from that of wage rates or the economic gains obtained per unit time worked.

But observations are not usually available on the income, earnings or wages of cohorts over their lifetime; rather, analysis typically relies on cross-sectional age groupings of a population at one time, from which a "synthetic" age-income profile is obtained. This cross-sectional (period)

age-earnings profile will differ from the longitudinal (cohort) age-earnings profile for two, possibly interrelated, reasons. First, different age groups in the cross section will tend to have different levels of education, and other productive qualifications. Younger age groups will in general have received more years of schooling than older age groups, with the consequence that cross sectional age-earnings profiles will tend to increase more slowly with age, peak earlier, and decline more rapidly than would the corresponding age-earnings profile from longitudinal data on individuals. In populations where the level of education has been increasing rapidly in recent decades, the covariance between age and education for workers will be large and negative. Cross sectional age-wage profiles for such populations will tend to be flatter at the younger ages than would be the case for a representative individual progressing through their life cycle in these populations.

Holding constant the educational qualifications and hours worked of the labor force, differences between longitudinal and cross sectional age-earnings profiles may remain. This residual may be attributed to omitted productive characteristics of the work force or secular growth (or decline) of labor productivity that workers capture due to physical capital accumulation and the growth in technical knowledge. If this residual effect on the productivity of labor is proportional in its impact on the earnings of workers at all ages, this age-neutral secular shift in productivity would contribute to further reducing the positive slope (or increasing the negative slope) of the age earnings profile as observed in the cross section.³

Within Age Group Inequality

The third component of income inequality in equation (1) is the within age group log variance. Some procedure is called for to summarize these measures of inequality over age groups to represent lifetime incomes. The cross-sectional decomposition suggests simply applying the current population structure, n_j/N , as weights, but this is inadequate if there are important sources of covariance between one time period and the next for individual incomes. Recent lifecycle econometric research has begun to estimate dynamic earnings models based on U.S. panel survey data. Persistent differences among individuals are characterized by permanent individual effects in these models, and transitory shocks to income are generally assumed to be serially correlated (e.g., Lillard and Willis, 1978). But shortage of panel data outside of the U.S. and the limited agreement on statistical specification for these dynamic models has slowed progress toward empirical generalizations. Only cross sectional static summarizations of lifetime inequality are within the scope of this paper.

Three Empirical Examples

Table 1 reports the empirical counterparts for this decomposition for the Netherlands in 1950 for individual annual income recipients, the United States in 1970 for all male annual income recipients, and Colombia in 1973 for all males with monthly money income. Column (1) reports the age structure of income recipients; column (2) the difference between the age group's average (log) income and that of the entire population, squared, with those age groups below the average showing a negative sign; and column (3) the within age group log variance. Column (4) presents the sum of the within and

Table 1

Components of the Variance of the Logarithms of Personal Incomes: Selected Countries

Country, Date and Age of Income Recipient (in years)	Proportion of Income Units	Squared Differ- ence Between Cohort and Population Mean Incomes	Within Cohort Income	Total Cohort Variance Components	Weighted Variance Share
	n_j/N (1)	$(\bar{y}_j - \bar{y})^2$ (2)	$\sum (\bar{y}_j - y_{1j})^2$ (3)	(2) + (3) = (4)	(1) x (4) = (5)
<u>Netherlands,</u> <u>1950, Persons:</u>					
15-20	.14	(-) 1.043	.181	1.224	.173
21-29	.22	(-) .047	.347	.394	.086
30-39	.19	.108	.357	.465	.087
40-49	.17	.220	.430	.650	.110
50-59	.13	.161	.506	.667	.090
60-69	.09	.013	.608	.622	.058
70 or more	.06	(-) .068	.628	.697	.040
Total	1.00	-	.644	-	.644
<u>United States,</u> <u>1970, Males:</u>					
14-19	.10*	(-) 3.24	1.04	4.28	.42
20-24	.11	(-) .15	.82	.97	.11
25-34	.20	.23	.46	.69	.13
35-44	.18	.66	.49	1.15	.19
45-54	.18	.37	.59	.96	.16
55-64	.14	.12	.77	.89	.12
65 or more	.09**	(-) .09	.64	.73	.06
Total	1.00	-	1.19	-	1.19
<u>Colombia, October,</u> <u>1973, Males:</u>					
10-19	.14	(-) .564	1.26	1.83	.252
20-24	.15	(-) .028	1.37	1.40	.214
25-29	.14	.027	1.55	1.58	.219
30-34	.12	.077	1.62	1.70	.209
35-44	.21	.079	1.81	1.89	.396
45-54	.14	.043	1.94	1.98	.275
55 or over	.10	(-) .025	2.33	2.36	.233
Total	1.00	-	1.80	-	1.80

* proportion based on 2/3 of men age 14-19

** proportion based based on men age 65-74 only

between cohort components to the overall income inequality. If these values differ substantially by age group, then the earlier noted differences in age structures might help to explain variation in aggregate income inequality. Column (5) multiplies the age group's weight by its contribution to overall income inequality.

In the case of the Netherlands, the contribution to inequality of the youngest age group, age 15 to 20 is greatest; this group represents less than a tenth of the population but accounts for 27 percent of overall income inequality (i.e., $.173/.644 = .27$). The increase in the population share of this young group in the decade after 1950 contributed to increasing measured income inequality in the Netherlands by 1959 (Schultz, 1965).

The youngest age group is also the primary contributor to overall income inequality in the United States, constituting again a tenth of the estimated total number of males with income but contributing 35 percent of the log variance (i.e. $.42/1.19 = .35$). The source of this inequality due to the youngest labor force entrants differs in the two countries, however. In the Netherlands persons age 15-20 receive substantially lower than average incomes, but these incomes are relatively equally distributed within the age group, whereas in the United States the level of income for young men age 14-19 is not only lower but the variance within the young age group is also larger than any other age group. In either case, swings in the proportion of the population in the youngest age group could influence measured income inequality, without necessarily implying any change in inequality in lifetime economic opportunities of persons in these societies. From this evidence for two high income countries there is support for the hypothesis that the slow tendency for income inequality to diminish

in industrially advanced countries in this century could be partially explained by their aging population structures.

Monthly male income data from Colombia do not support the view that aggregate money income inequality is necessarily sensitive to changes in the age structure. The contribution to the total log variance attributable to each age group is nearly constant, as shown in Column (4). Though incomes are below average in the youngest age group, the difference is smaller than for the other two countries. Also, as in the Netherlands, the log variance of incomes within the youngest age groups is considerably smaller than within the older age groups.

In several respects the income data for Colombia differ from those available for the United States and the Netherlands. First, the Colombian sample is more restrictive with regard to employment status; domestic servants and unpaid family workers are excluded because these workers tend to receive all or a substantial fraction of their income in kind, e.g., room and board, and youth are often found in these employment groups in Colombia. Unpaid family workers with no income are also excluded from the U.S. data, but they represent a far smaller share of the U.S. population. Second, income is measured in the Colombian Census over the preceding month rather than over a year. This convention could affect measured inequality and the composition of the sample of income recipients, particularly before age 25 when young men are entering the labor force and terminating their education.⁴

Educational achievement in Colombia has increased, perhaps more rapidly than it has in the United States and the Netherlands. But holding constant for educational attainment of workers does not consistently increase the slope of the age-income profile, as postulated, because for the

two youngest age groups the men who report incomes tend to be those members of their cohort who have less than average levels of education. This selection bias leads to the result that the average education of the income recipients increases until age 25-29. Only after age 29 does the adjustment for the educational attainment of Colombian men increase the derivative of the income profile with respect to age.⁵

Extrapolations of Age Structure Effects on Income Inequality

A principle difference between the income data for the two high income countries and Colombia is the relative insensitivity of measured overall income inequality in Colombia to the age structure. To illustrate this difference between the U.S. and Colombia, assume that the age-income profile and within age group inequality did not change from the 1970-73 figures reported in Table 1. The actual change over time in the age structure of males in these two countries would then imply the estimates of the overall log variance of male incomes shown in Table 2. In the case of the United States, the gradual decline in fertility and decrease in immigration has had the effect of shifting the age structure of the population toward older ages, with the calculated effect of reducing the log variance of incomes from 1.57 in 1830 to 1.14 in 1950. Swings in birth rates since the depression have contributed subsequently to swings in the log variance of incomes, increasing ten percent from 1950 to 1960, decreasing 14 percent to 1980, and increasing again by ten percent by the year 2000. These are relatively large variations in measures of income inequality that are generally quite stable over time within a country (Schultz, 1975).

Table 2

Extrapolations of the Log Variance of Personal Incomes of Males
over Age 14 for the United States and Colombia,

Based on Changing Age Structures,
and Assuming that Variance Components of Age Groups Do Not Change

<u>Year of Age Structure from Census or Projection</u>	<u>United States 1970 Table 1, Col. (4)</u>	<u>Colombia 1973 Table 1, Col. (4)</u>
1830	1.57	
1870	1.47	-
1900	1.40	-
1940	1.32	(1938) 1.78
1950	1.14	(1951) 1.77
1960	1.25	(1964) 1.78
1970	1.19	(1973) 1.80
1980	1.07	-
1990	1.13*	(1993) 1.77**
2000	1.17*	-

* Census Bureau Series II that assumes cohort fertility rate stable at 2.1 children per woman.

** Projected by author assuming Coale and Demeny (1966) West level 18/19 tables applicable to males in two decades after 1973 Census and fertility continues to decline, but more slowly after 1973. Age group under 19 with income set equal to 2/3 of ages 15-19; age group with incomes over 55 set equal to ages 55-64 only.

Source: U.S. Bureau of the Census (1960) and subsequent CPS publications on projections. Colombian Office of Statistics, DANE (1977).

In the case of Colombia, however, despite the destabilizing effect of the demographic transition on the population's age structure since the 1940s, the estimated swings in the age structure from 1937 to 1993 have little effect on the calculated log variance of male incomes. If Colombia's age-income profile and within age group income inequality is closer to that in other low income countries than is that of the Netherlands and the United States, the change in age structures that has been produced by the demographic transition may not have been of itself a dominant factor in explaining variation and change in overall money income inequality in less developed countries among individuals.

Large residual differences across countries remain in within age group income inequality. At about age 30, when continuing education and early labor force investments should be least important, the log variance of incomes is about .35 in the Netherlands, .45 in the United States among males, and 1.6 in Colombia among males. If these differences are not due to differences in statistical sources, these are indeed large differences in lifetime inequality, as are those extrapolated for the United States over the century 1830 to 1950. Summing within age group inequality with population structure weights, i.e., $\Sigma \text{Col. (1)} * \text{Col. (3)}$, this static measure of lifetime income inequality is .40 in the Netherlands, .65 in the United States, and 1.68 in Colombia. Conversely, lifecycle variation in income inequality approximated by deviations of the age-income profile from the average, i.e., $\Sigma \text{Col. (1)} * \text{Col. (2)}$, account for 50 percent of the overall measure of income inequality in the U.S., 37 percent in the Netherlands, but only 7 percent in Colombia.

Age Structure Effects on the Equilibrium Components of Income Inequality

The form of static decomposition performed above focuses on only the direct effect of changes in age structure on income inequality. But the age structure also indirectly affects measured inequality by changing the age-income profile and by influencing within age group income inequality. Estimation of the incidence and magnitude of these indirect demographic effects calls for economic analysis of time series. The former case is precisely the demographic-economic mechanism that Ronald Lee (1977) has explored to explain long swings in the relative income status of a sequence of U.S. birth cohorts of differing size. Relatively large (small) cohorts are expected to depress (inflate) their longitudinal path of earnings and thus distort the cross sectional age-income profile.

The institutional and technical mechanisms determining the adjustment of cohort earnings to cohort size remain unclear; do adjustments across age groups occur in wage rates or in hours worked; does the latter adjustment come about through change in labor force participation rates, unemployment rates or average hours worked by the employed? Recent evidence for the United States indicates that most of the adjustment of the labor market to the relatively large cohort entering the labor force in the 1970s occurred through adjustment of wage rates by age, but age-specific unemployment rates and labor force participation rates also reacted to cohort size (Freeman, 1979; Welch, 1979). Further research is required to clarify whether cohort relative size imparts a persisting lifetime effect to the level of the cohort's longitudinal age-income profile, or whether cohort size primarily influences the cohort's starting wage, and that this initial effect subsequently wears off as members of the cohort obtain more job-related experience.

Macro economic indicators of the tightness of labor markets and the effect of such tightness on inflation were reconsidered in the 1970s as the labor force grew more rapidly and its age-sex composition changed. Real wage rates tended to deteriorate for youth with limited experience, and yet wages increased for older, more experienced male workers. The overall unemployment rate increased, but this development did not curb inflationary pressures from some segments of the labor market.

Finally, within age group income inequality may be affected by cohort relative size, other things being equal. Since tight labor markets are generally associated with diminished income inequality, within age group inequality is likely to diminish for relatively small cohorts, and widen for large cohorts. But evidence of the effect of cohort size on the cross sectional age-income profile and on within age group inequality is no more than suggestive at this time. Firm conclusions as to the magnitude and persistence of these effects of cohort size on the structure of income inequality in high and low income countries must await further research and probably analyses of longer time series than have been available to date.

When younger groups in the labor force increase more rapidly than do others, the effect of increased cohort size of the new entrants is to augment overall measures of income inequality. This will occur until the rapidly growing age group's contribution to the overall log variance of incomes is no longer greater than average (Col. 4). This may occur between about age 25 and 35, depending on the slope of the age-income profile. As the growth of the labor force entering cohorts falls below the average for the population of income recipients, and the most rapid growth occurs in the middle age groups, the indirect effects of relative cohort size are likely

to reduce overall income inequality. The precise timing of this reversal depends on several as yet unquantified offsetting factors.

In conclusion, the static decomposition illustrates how recent changes in the age structure of high income countries may explain secular trends and recent cycles in their measured income inequality. The importance of these direct age structure effects may be less marked in low income countries, at least this appears to be the case for Colombia in 1973. But the data for Colombia may overstate the relative income position of youth, because unpaid family workers and other low productivity groups that are numerically important in Colombia are not observed as individual income recipients. Improved income data, corrections for selection bias, and further analysis of the family as the production unit may clarify some of these issues in low income countries.

Conversely, patterns of part-time employment and the inclusion of students and unemployed may understate the relative income position of youth in the United States and thus exaggerate the importance of youth in overall measures of inequality. If the U.S. decomposition of log variance is repeated for 1967, when earnings are reported from the U.S. Current Population Survey for full-time year-round working males, the age pattern and level of inequality is different from that for all males with income, but the finding stressed in this paper of the overall sensitivity of measured inequality to the age structure does not change. The relative income status of men age 14-19 improves (about doubles) and the within age group inequality diminishes, particularly for men age 20-24. But since the overall log variance of full-time year-round earnings is 60 percent less than that for all income recipients, the

total cohort contribution of youth age 14-19 (column 4, Table 1) remains about four times the overall average log variance, similar to that reported in Table 1 for all income recipients in 1970. (See summary and sources of data for 1967 in Schultz, 1975.) Although there are reasons to prefer measures of inequality based on wage rates or restricted to comparisons of persons working in the labor market the same amount of time, this is not a common practice and has not been the empirical basis for the widely observed relationship between economic development and income inequality (Kusnic and DaVanzo, 1980; Schultz, 1981, Section D).

Indirect dynamic effects of cohort relative size on cohort earnings and on within cohort inequality, possibly associated with the demographic transition, may also be responsible for increasing measured income inequality. Since these latter two effects of relative cohort size on income inequality have a clear bearing on inequality in lifetime income opportunities, they warrant more explicit study in which the direct effects of age structure are held constant. Data examined here relate to only three countries, each at only one time period. They do not provide more than an illustration of the proposed decomposition methodology. They do suggest, however, that there are substantial differences between countries in within age group inequality. They also imply that at least in high income countries changes in overall inequality may be strongly influenced by age structure shifts, both secular trends and long swings. Many standard interpretations of patterns in overall inequality may need to be revised when these salient effects of age structure are suitably identified and removed from the data.

FAMILY COMPOSITION AND INCOME INEQUALITY

Two approaches to income inequality are found in the economics literature: one emphasizes the distribution of endowments and productive opportunities among individuals over their lifetimes, while the other treats income per family, adjusted somehow for its current consumption needs. The former is oriented toward understanding the determinants of earnings of productive factors and their personal distribution, whereas the latter is concerned with the distribution of consumption and economic welfare.

As an income recipient unit, families differ in size and composition, and some studies suggest that family composition responds to the economic endowments and opportunities of its potential members. Whereas the age structure of a population was previously interpreted as given and thus exogenously affecting the distribution of income across individuals, it is not always reasonable to assume that the size and composition of families is exogenously affecting the size distribution of family incomes.

The second half of this paper explores how family income inequality might be approached with decomposition methods to clarify two distinct demographic processes that modify family size and composition: the propensity of adult to share living arrangements, and the level of surviving fertility per adult.

One common procedure to normalize the distribution of income across families for family composition is to express the income of the family (or unrelated individual) in per capita terms.⁶ This per capita family income measure of economic welfare, or consumption opportunities is adopted here for simple illustrative purposes.

A Framework for Study of Demands Underlying Family Composition

Two sources of variation in family size are conveniently distinguished: the number of adults and number of children in the family. The number of children (under age 15) per family is expressed per adult (age 15 or over) in the family, and may be viewed as an index of net reproduction, which embodies the impact of both fertility and the incidence of child (and adult) mortality on the rate of population growth. The first component of family composition is defined for our purposes as the natural logarithm of one plus this index of surviving fertility:

$$f = \ln(1 + n_c/n_a),$$

where the number of persons in the family, N , is simply divided between children and adults, $N = n_c + n_a$.

The second component of family composition is the logarithm of the number of adults living together in the family:

$$a = \ln(n_a)$$

The logarithm of family income is then the final or third component of our measure of personal economic welfare, the logarithm of family per capita income:

$$\ln\left(\frac{Y}{N}\right) = \ln(Y) - \ln(n_a) - \ln(1 + n_c/n_a)$$

which is rewritten as follows:

$$y_n = y - a - f.$$

The variance of the logarithms of family per capita income or income "inequality" can then be decomposed into three variance and three covariance terms as follows:

$$V(y_n) = V(y) + V(a) + V(f) - 2C(y,a) - 2C(y,f) + 2C(a,f) \quad (2)$$

where $V(.)$ and $C(.,.)$ represent the variance and covariance of the respective argument(s). Because the adult size and fertility index contribute to a reduction in family per capita income, the first two covariance terms involving adult size and the fertility index with family income are subtracted from the sum of the three component variances.

The covariance between family income and adult family size and between family income and fertility can be economically interpreted as the responsiveness of these aspects of family composition to income. But economic analyses of the demand for children and marriage rely heavily on relative price variation across populations that is captured in the differences in the shadow value of time (or wage rates) of men and women. Furthermore, if investments in children, such as schooling, are substitutes for numbers of children, differences in fertility may parallel inversely investments in population quality. Hence, the next step in elaborating this framework is to distinguish between the potential earned income or wage rate of adult males and females, with the expectation that female potential income will be inversely related to fertility due to the predominance of own-price effects, whereas male income will be weakly related to fertility, positively or negatively, depending on a variety of factors that determine the magnitude of offsetting income and price effects (Schultz, 1976).

It is not uncommon to focus attention on the covariance between family income and total family size, but, as suggested above, this procedure may conceal the more interesting relationships between income and the subcomponents of family size that represent distinct demands.

Empirical Illustration: Colombia 1973

Table 3 reports the means, variances, and covariances of measures of unrelated individual and family income and family composition for Colombia, stratified by age of head of household. The data are, as before, from a four percent public use sample of the noninstitutional questionnaires of the October 1973 Population and Housing Census of Colombia (DANE, 1977). Only units reporting some income in the month before the Census are considered.

Several empirical regularities may be noted. The number of adults per family increases steadily with the age of head, from 2.1 at ages 15-19 to 3.7 at ages 50-64.⁷ The (surviving) fertility index (i.e., children per adult) increases from .38 at ages 15-19 to a maximum 1.14 at ages 35-39, and thereafter falls to .45 in the last age group. The total number of persons per family therefore increases rapidly from 2.9 at ages 15-19 to 5.6 at ages 35-39, and is more or less constant thereafter, as the share of adults in the family slowly increases. Family income also increases with the age of its head, but at a slower rate than does family size, peaking at ages 45-49. Consequently, family per capita income reaches its largest value in this cross section for young families whose heads are ages 20-24, and declines thereafter until ages 45-49.

ANEXO
Componentes of the Logarithmic Variance of Family and Unrelated Individual Incomes, by Age of Head:
Colombia, 1973

Means and Variances	Age of Head of Household							Total 15-64
	15-19	20-24	25-29	30-34	35-39	40-44	45-49	
1. Log. Income $y = \ln(Y)$	6.67 (1.05)	6.87 (1.06)	7.03 (1.23)	7.11 (1.39)	7.07 (1.44)	7.13 (1.44)	7.17 (1.51)	7.08 (1.45)
2. Log. Per Capita Income $y_n = \ln(Y/N)$	5.61 (1.29)	5.69 (1.28)	5.64 (1.47)	5.50 (1.58)	5.34 (1.61)	5.33 (1.58)	5.36 (1.55)	5.45 (1.66)
3. Log. No. Adults $a = \ln(n_a)$	.736 (.311)	.793 (.216)	.840 (.188)	.883 (.177)	.962 (.196)	1.13 (.234)	1.25 (.275)	1.30 (.303)
4. Log. Fertility Index $f = \ln(1 + n_c/n_a)$	.323 (.145)	.386 (.131)	.551 (.165)	.716 (.187)	.762 (.190)	.673 (.176)	.554 (.157)	.373 (.123)
<u>Covariances</u>								
5. C(Y,a)	-.175	.140	.157	.179	.172	.193	.257	.297
6. C(Y,f)	-.0337	-.0791	-.125	-.140	-.109	-.103	-.0979	-.0704
7. C(a,f)	.0272	-.0069	-.0333	-.0525	-.0534	-.0541	-.0442	-.0144
<u>Slope Coefficients</u>								
8. B_1	.564	.649	.836	1.01	.877	.825	.933	.982
9. B_2	-.0322	-.0745	-.102	-.101	-.0757	-.0714	-.0649	-.0414
10. $dn_c/(dy/y)$	-.0928	-.242	-.410	-.500	-.424	-.431	-.391	-.220
Sample Size (Number of Families)	1,514	6,727	12,008	13,955	14,535	13,859	11,519	23,442
								97,559

Family income inequality, or the variance of the logarithms of family incomes, increases monotonically with age from 1.05 at age 15-19 to 1.70 at age 50-64. Family per capita income inequality tends to be larger and also increases with age until the age group 35-39, when the proportion of children in the family peaks; inequality then varies within a narrow range across subsequent age groups. Comparison with column (3) of Table 1 indicates that the within cohort variance of log incomes of individual men varies by age at a slightly higher level than does family per capita income inequality by age in Table 3.

Adult Family Size and Income

The first covariance term in the full decomposition, $C(y,a)$, represents the association between family income and the number of adults living together in a family. In a simple regression of the log of family income on the log of the number of adults in the family:

$$\ln Y = \alpha_1 + \beta_1 \ln n_a$$

where β_1 in row 8 of Table 3 is the estimate of the elasticity of family income with respect to number of adults in the family.⁸ If the propensity of adults to live together were not correlated with their potential income contribution, their own and that of others in their family, and our measure of family income captured fully the sum of the potential income of its adult members, then this elasticity would be approximately one. Family incomes would then tend to increase in proportion to the number of adults in the family.

The Colombian relationship between family income and number of adults per family implies approximately a unitary elasticity from age 25-29 to 50-64, varying from .83 to 1.01. But several factors could explain departures of this elasticity from unity. First, there may be economies of scale in household production and thus gains from specialization within the family-household that encourages some degree of combination and coordination of adult activities. Also the production and rearing of children at certain stages of the life cycle is an important factor in the combination of adults into families in most societies. Technology of production, firm-specific training, and information costs of monitoring activities may work to extend further the productive limits of the family beyond the nuclear childrearing unit. The production determinants of adult family composition is a largely unexplored field for theoretical speculation and empirical study (Rosenzweig and Wolpin, 1979).

The effect of numbers of adults on family income may also be distorted by imperfect measures of family income or production. Goods and services produced and consumed within the family are often omitted from measures of family income. The proportion of nonmarket income in the family's real total income may vary with number of adults, e.g., in two-adult families compared with one-adult "families." The proportion of total income consumed in nonmarket forms may also vary systematically over certain periods of the life cycle, such as during the early years of childbearing before the first offspring begins to contribute importantly to family income. This hypothesis could explain the markedly lower values of the adult-income elasticity for families whose heads are between the ages of 15 and 24.

Just as the composition of income between market and nonmarket sources may be affected by the number of adults in the family, the potential lifetime wealth or permanent income of individuals may influence the demand for goods and services that are more economically produced in larger (or smaller) units, or in the market or nonmarket sectors (Kusnic and DaVanzo, 1980). The nuclear family is thought to facilitate childbearing and the transmission of productive skills and culture to the young. As more of these functions are performed outside of the family, and the share of adult lifetimes devoted to childbearing decreases, the need for a permanent nuclear family may diminish, or at least that is concluded from some studies of modern industrialized societies. As the wages for men and women in the marketplace approach equality in some high income countries, the marriage gains from specialization in market and nonmarket production are reduced and the opportunity cost of single household "privacy" and mobility may decrease. The recent increase in the proportion of single person households in many high income countries may be attributed to the high income elasticity of demand for this form of "privacy" (Michael, et al., 1980).

In sum, the proportionate relationship between family income and number of adults in the family implies that the combination of adults into family units is not associated in Colombia with augmenting or diminishing appreciably the inequality of family per capita income. This is a relatively neutral factor of family composition, except during the early childrearing years, ages 15 to 25. During these years parents produce nonmarket income in the form of child care services that are excluded from personal income accounts. A more comprehensive measure of family income that included household nonmarket production and child care services might, therefore, increase the estimated income-adult elasticity for these younger age groups, and probably also increase slightly the estimated elasticity at later ages.⁹

Fertility and Family Income

Fertility and family income tend to be inversely related, with the associated covariance obtaining a maximum value in the Colombian case at ages 30-34 (Table 3), and decreasing thereafter slowly to the oldest age group. Here the causal relationship is thought to operate primarily from the level of family income to the level of fertility, and to be achieved through voluntary choice rather than any form of biological predisposition. Of course, children may also contribute by their efforts to family income, though this positive effect should not be substantial until a child is about ten years of age, and, by our measuring convention, this child upon reaching age 15 is counted as an adult even though the average age at first marriage in Colombia is now eight years later, at age 23. The observed negative association at all ages suggests the level of surviving fertility systematically declines with increases in family income. Thus, differences in surviving fertility across families in Colombia increases inequality in family per capita income. The negative covariance between income and fertility never reaches, in absolute value terms, the magnitude of the positive covariance between income and adult family size, but remains substantial. The collapse of income-fertility differentials would reduce overall inequality in family per capita income by 12 to 18 percent from ages 20-24 and 45-49.

In the regression of the logarithm of the fertility index on family income:

$$\ln(1 + n_c/n_a) = \alpha_2 + \beta_2 \ln Y,$$

the income-fertility elasticity, β_2 , is reported in row 9, Table 3, and converted in row 10 into the derivative of change in the number of children

per household with respect to a proportionate change in income, evaluated at the sample logarithmic means. At ages 30-34 the relationship suggests a doubling of family income is associated with a decrease of .50 in n_c or an average reduction of one-half of a child from the sample mean of 2.5 children per family. This fertility derivative remains in the vicinity of -.4 from age 25-29 to 45-49, suggesting a strong inverse relationship from income to fertility at all ages and not one that is limited to the initial timing of childbearing.¹⁰

Again, one suspects that nonmarket income is greater in higher fertility families, but this fact is unlikely to change the conclusion that differential fertility patterns by income classes in Colombia add to the inequality in per capita economic resources in the hands of families. This finding would be less pronounced if children were weighted less heavily than adults by our "per capita" normalization of family income, but the empirical regularity and the direction of its effect on inequality would not thereby be changed.

The positive family size-family income relationship is insufficient to prevent family per capita income from declining in larger family units. This observation has been stressed in the recent writings of Kuznets (1976, 1978). In the case of Colombia in 1973, the inequality increasing effect of the distribution of families by size is due almost entirely to different levels of fertility (or the proportion of children among all persons in the household) by income level and is hardly affected at all by the size distribution of numbers of adults living together.

Adult Family Size and Fertility

The third covariance term in our decomposition of per capita family

income inequality is less readily ascribed an economic interpretation and, given its modest size, it will be discussed only briefly here. Large collections of adults may be synonymous with extended families. It has been hypothesized that extended family structures may lower the cost of children and contribute to higher fertility. With more adults to coordinate and specialize in home child care activities, for which there may exist economies of scale, the opportunity cost of children may be reduced. But there is no evidence in support of this hypothesis from these data; extended families with more adults are associated with somewhat lower levels of fertility per adult. The covariance between the logarithm of the fertility index and logarithm of the number of adults in the family, $C(f,a)$ in row 7, is about $-.03$ to $-.05$ from ages 25-29 to 45-49.

CONCLUSIONS

It is essential that we get behind overall measures of individual and family income inequality and identify regularities among subcomponents that have economic and behavioral meaning. The first part of this paper analyzed individual income inequality in three countries, and sought to distinguish between inequality directly associated with the age structure of the population, that associated with the cross sectional age-income profile, and that remaining within age groups. The first two sources of aggregate inequality warrant more study, but the issues of equity and economic inequality are perhaps most clearly associated with the third component of inequality, that which arises within a birth cohort.

The larger the share of youth, age 14-19, among income recipient units the greater is the aggregate log variance in individual incomes in the Netherlands and in the United States. Given the currently documented

pattern of income inequality within and across age groups, the secular trend toward an older age composition in the United States could directly account for a decrease of one-third in U.S. individual income inequality since the Civil War, as found by Lindert and Williamson (1976). Indirect economic effects of changes in age structures should reinforce this extrapolated trend in income inequality based on direct effects of compositional changes. Age structure differences between Colombia and the United States do not explain, however, the much greater overall inequality in Colombia. Within age group inequality is more than twice as large in Colombia as in the United States, while inequality is a third less in the Netherlands than in the U.S., independent of age structure. Comparing income inequality by age across populations is complicated, however, by the different employment opportunities open to youth in countries at different levels of development, and in particular, family unpaid jobs. The empirical data considered here suggest that the widely observed relationship between modern economic growth and decreasing aggregate income inequality may be exaggerated by differences in age structures across contemporary populations and over time within more industrially and demographically advanced countries.

The second part of this paper sought to divide family composition into two distinct behavioral elements associated with fertility and the propensity of adults to live together. The latter adult aspect of family structure may in some circumstances respond to income opportunities across the population. But in the case of Colombia in 1973, the elasticity of family income with respect to number of adults in the family was nearly one, indicating that this process was not a major source of inequality in per capita family income. Fertility, on the other hand, was distinctly higher in low income

families, adding to the inequality in per capita family income in Colombia. Replication of this simple decomposition analysis in other countries at different stages in the demographic transition and at different income levels might clarify how fertility by family income varies with particular patterns of economic growth and with different emphases on education, public health and family planning activities.

The avenues open to research are many. To better understand family income inequality, behavioral and institutional causal interpretations are needed of component regularities. Micro economic theory, standard techniques of statistical decomposition and estimation, and common procedures of age and sex stratification may all be useful in advancing this goal. The growing public availability of large household surveys and samples of censuses for many countries and time periods provides the opportunity to proceed in a variety of directions as explored here, without being limited to standard tabulations and income accounting frameworks. Age structures of the income recipient population, fertility (and mortality) differences by family income, and the market labor supply behavior of women, appear to be essential parts to this puzzle. The parts must fit together and add up to a consistent whole. The framework that takes form from this research should facilitate more precise and meaningful measurement of income inequality than has been seen in the past. It should also suggest new approaches to assembling these components into a integrated two-level household-aggregate model of economic demographic development, one that has been sorely needed since Malthus's grand design went wide of the mark a hundred years ago.

FOOTNOTES

¹See for example Schultz (1965), Pyatt (1966), and Fei, Ranis and Kuo (1978). The log variance of incomes assigns greater weight to inequalities among the poor than among the rich, in contrast with the Gini coefficient and the coefficient of variation, which assign equal weights to the same absolute differences in income between rich persons and between poor persons. Although rankings of inequality across countries or socioeconomic groups tend to be relatively insensitive to which of these alternative measures of inequality is adopted (Atkinson, 1970), the empirical conclusions reported here may not hold for other summary measures of inequality.

²See Rosen (1977) for discussion of an earnings function as a structural equation and as a reduced-form equation. The distinction does not seem paramount in this context but is important for the economic interpretation of the earnings function and its parameters.

³Analysis of the divergence of longitudinal age-income profiles from cross-sectional age-income profiles for individuals in the United States from 1947 to 1970 suggests that secular productivity gains have been roughly age-neutral for males age 25-54, for whom the average annual hours of work have changed least (Schultz, 1975).

⁴In one way the measurement of income during the last month helps to standardize income for the duration of time worked, and provides a better approximation for the wage rate. Consequently, students who would work for pay only during summers would be included with

artificially low incomes in the Dutch and U.S. statistics, but are probably excluded from the Colombian Census sample. On the other hand, workers who were entirely unemployed last month with no other sources of income would be unavoidably excluded from the Colombian sample and might be included in the U.S. and Dutch data if they found any employment or received any welfare-transfer income during the preceding year. Yet unemployment in Colombia is not frequent by conventional standards according to the 1973 census: two to five percent of men ages 10 to 24 are unemployed.

Our reliance on the logarithmic variance of incomes to summarize income inequality does not permit the retention of persons with no income in our sample. But if persons with no income are to be included, the universe of income recipient units must be defined on new criteria such as the individual's labor force status. The disabled, housekeepers, pensioners, and discouraged workers who report themselves as being outside of the labor force are thereby arbitrarily excluded. It would be preferable to use only exogenous characteristics for determining the study population, such as sex and age. If men of a specific age were considered, additional problems arise of placing a value on outputs of (or inputs into) all household production, schooling and training activities. To define a measure of "full" income comprehensively as both money market income and the cash value of all services and goods consumed or invested in kind involves one in many more unresolved conceptual and empirical problems (Kusnic and DaVanzo, 1980) .

⁵This adjustment is performed by regressing the natural logarithm of individual income on a series of dummy variables defined over the age intervals included in Table 1. By including in this regression a series of dummy variables for four levels of education, or a continuous variable

for years of education of the individual, we "adjust" the age profile estimates for education. The squared deviations of the log income profile from the overall mean become: $(-).200$; $(-).005$; $.027$; $.110$; $.216$; $.203$; $.038$, respectively for the age groups. The deviation effects are increased for age groups beyond age 30, and decreased for those below age 25.

⁶ An alternative approach would divide family income by a weighted sum of family members, where the weights assigned to different types of persons in the family would be dictated by the purposes of the analysis, such as the study of consumption or production. For example, it has been argued that consumption requirements of a person vary by age and sex. Real income available for consumption may tend to be overstated by the per capita normalization of family income in families with a relatively large number of children, and conversely understated in families with disproportionate numbers of adults. An alternative normalization scheme could also be examined that assigns weights to children which are some variable fraction of those assigned to adults, reflecting crudely the lower production potential or consumption requirements of children than of adults.

⁷ References are to antilogs of the means of logarithmic variables reported in Table 3. For example, the number of adults per unit, whose head is age 50-64 is $3.7 = \exp(1.3)$.

⁸ The simple regression coefficient in this case, β_1 , is equal to the covariance of $\ln Y$ and $\ln n_a$, divided by the variance of $\ln n_a$. For example, in a family whose head was between the ages of 30 and 34, $\beta_1 = C(y,a)/V(a) = .179/.177 = 1.01$.

⁹One may wonder, however, whether the single loglinear relationship estimated here does not embody several distinct relationships for different family sizes and production technologies. More research might isolate whether the relationship that holds from one to two adult families continues to fit the data for three, four and five adult families, and whether the relationship between family income and number of adults per family is the same in populations where the household head is a rural landless worker, owner-operator-farmer, urban self employed, or urban wage-salary employee.

¹⁰The fertility demand equation could also be viewed as conditioned on family income per adult or an average "wage rate", since the elasticity of family income with respect to adult size is approximately one.

REFERENCES

- Atkinson, A.B. (1970) "On the Measurement of Inequality," Journal of Economic Theory, Vol. 2, No. 3, September, pp. 244-263.
- Coale, A.J. and P. Demeny (1966) Regional Model Life Tables and Stable Populations, Princeton: Princeton University Press, 1966.
- DANE (Departamento Administrativo Nacional de Estadística) 1977, La Poblacion en Colombia 1973, Muestra de Avance, Bogota, Colombia.
- Fei, J.C.G., G. Ranis, and S.W.Y. Kuo (1978) "Growth and the Family Distribution of Income by Factor Components," Quarterly Journal of Economics, February, pp. 17-53.
- Freeman, R. (1979) "The Effect of Demographic Factors on the Age Earnings Profile in the U.S., " Journal of Human Resources, Summer, 1979.
- Kuznets, S. (1955) "Economic Growth and Income Inequality," American Economic Review, Vol. 45, No. 1, March, pp. 1-28.
- _____ (1963) "Quantitative Aspects of the Economic Growth of Nations: VIII, Distribution of Income by Size," Economic Development and Cultural Change, Vol. 11, No. 2, Pt. II, January.
- _____ (1978) "Size and Age Structure of Family Households: Exploratory Comparisons," Population and Development Review, Vol. 4, No. 2, June, pp. 187-224.
- _____ (1980a) "Size of Households and Income Disparities," Yale Economic Growth Center Discussion Paper No. 342, March, New Haven, Conn.
- _____ (1980b) "Inequality in the Size Distribution of Households: Differences and Trends," Yale Economic Growth Center Discussion Paper No. 343, March, New Haven, Conn.

- Kusnic, M. and J. DaVanzo (1980) Income Inequality and the Definition of Income: The Case of Malaysia, Santa Monica, Calif.: The Rand Corporation, R-2416, June.
- Lee, Ronald (1977) Fluctuations in U.S. Fertility, Age Structure and Income, NICHD research report, Population Studies Center, Ann Arbor, Michigan, University of Michigan, July.
- Lillard, L. and R. Willis (1978) "Dynamic Aspects of Earnings Mobility," Econometrica, Vol. 46, No. 5, September.
- Lindert, P.H. and J.G. Williamson (1976) "Three Centuries of American Inequality," Research in Economic History, Vol. 1, September.
- Michael, R.T., V.R. Fuchs and S.R. Scott (1980) "Changes in the Propensity to Live Alone," Demography, Vol. 17, No. 1, February, pp. 39-56.
- Mincer, J. (1974) Schooling, Experience and Earnings, New York: Columbia University Press, 1974.
- Pyatt, G. (1976) "On the Interpretation and Disaggregation of Gini Coefficients," Economic Journal, Vol. 86, No. 342, June, pp. 243-255.
- Rosen, S. (1977) "Human Capital: A Survey of Empirical Research," in R.G. Ehrenberg (ed.) Research in Labor Economics, Vol. 1, Greenwich, Conn.: JAI Press.
- Rosenzweig, M.R. and K. Wolpin (1979) "An Economic Analysis of the Extended Family in a Less Developed Country," Yale Economic Growth Center Discussion Paper No. 317, August, New Haven, Conn.
- Schultz, T.P. (1965) "The Distribution of Personal Income: Case Study of the Netherlands," unpublished Ph.D. dissertation, MIT, Cambridge, Mass.

- Schultz, T.P. (1975) "Long Term Change in Personal Income Distribution: Theoretical Approaches, Evidence, and Explanations," in D.M. Levine and M.J. Bane (eds.), The Inequality Controversy, New York: Basic Books.
- _____ (1976) "Determinants of Fertility: A Microeconomic Model of Choice," in A.J. Coale, (ed.), Economic Aspects of Population Growth, London: MacMillan.
- _____ (1981) "Individual and Per Capita Family Income Inequality, and Family Market and Full Income," paper presented at Honolulu IUSSP seminar, April.
- U.S. Bureau of the Census (1960) Historical Statistics of the United States, Colonial Times to 1957, U.S. Government Printing Office, Washington, D.C.
- Welch, F. (1979) "Effects of Cohort Size on Earnings," Journal of Political Economy, Vol. 87, No. 5, Pt. 2, pp. 565-598.