

Development Discussion Papers

Ethnicity and Earnings in Malaysia

John Luke Gallup

Development Discussion Paper No. 593
July 1997

© Copyright 1997 John Luke Gallup
and President and Fellows of Harvard College

Harvard Institute for
International Development

HARVARD UNIVERSITY



ETNICITY AND EARNINGS IN MALAYSIA

John Luke Gallup*

Abstract

This paper considers the evolution of male earnings of the main ethnic groups in Malaysia. The relative earnings of the ethnic groups are especially at issue because of the ambitious anti-poverty and affirmative action programs that make up the New Economic Policy, inaugurated after bloody race riots in 1969. Data from the second Malaysian Family Life Survey (MFLS2), an important data source for recent studies of Malaysia, show that male Malay earnings steadily fell behind male Chinese earnings over the past thirty years. However, this is not consistent with the changes in large cross-sectional household surveys over the same time period. An explanation that could reconcile the two kinds of data is strong cohort effects, since the MFLS2 life history data has a different age and cohort structure than the cross-sectional data. The cohort hypothesis is rejected, which indicates there is probably recall bias in the reported work histories that varies with the ethnicity of the survey respondent.

John Luke Gallup is an Institute Associate at HIID and a Lecturer in the Department of Economics at Harvard University. His recent research includes work on migration in Malaysia, the determinants of fertility in Vietnam, and the role of endemic disease in economic development.

* I would like to thank Cassey Lee, Ron Lee, Bryan Lincoln, and Tey Nai Peng for helpful discussions. I also appreciate the support of a Mellon travel fellowship.

Ethnicity and Earnings in Malaysia

John Luke Gallup

I. Introduction

In Malaysia, the causes of ethnic differences in male earnings and their trends since the implementation of the New Economic Policy (NEP) in 1970 are important for assessing one of the most ambitious affirmative action programs in the world.

Malaysia has an ethnic mix of Malays who make up just over half the population, Chinese who make up a third more, and Indians who take up most of the rest. Mainly due to British colonial policies, the Malays are more likely to be poor, rural, and farmers, the Chinese richer, urban, shopkeepers and traders, and the Indians agricultural laborers and professionals, though these ethnic typecasts have been breaking down in the last twenty years. Since Independence in 1957, Malays have held most of the political power, while Chinese hold most of the economic power.¹

Race riots in 1969 gave rise to the NEP, which has two main prongs:

to reduce and eventually eradicate poverty, by raising income levels and increasing employment opportunities for all Malaysians, irrespective of race; [and] ... accelerating the process of restructuring Malaysian society to correct economic imbalance, so as to reduce and eventually eliminate the identification of race with economic function.

(from the Second Malaysia Plan, 1971, quoted in Cho, 1990, p.68)

The first goal, the elimination of poverty, was meant to remove an important source of the ethnic strife. The second goal of restructuring society was an ambitious affirmative action program for Malays meant to redistribute economic power between the ethnic groups. It consisted of preference for Malays in employment, education, government positions, and the goal of creating a Malay bourgeoisie.

The surprise when looking at male earnings in the MFLS2 sample is that not only have reported Malay earnings not caught up with the earnings of Chinese since the NEP, they have fallen further behind. The earnings of men in all ethnic groups have grown since Independence, but in this sample, the earnings gap between the Malays and the Chinese also grew, and the gap grew faster after the start of the NEP in 1970 (see Figure 1.1).

This paper will look for the source of this anomaly and consider whether it is representative of the experience of the Malaysian population. Changes in the characteristics of Malays will be examined to see if they account for Malay earnings falling behind Chinese earnings. Perhaps Malay education did not keep up with the education levels of Chinese and Indians, or employers discount Malay credentials and pay Malays less for the same educational credentials because of preferences for Malays in education. Malays may have been relegated to sectors and occupations where earnings have not grown.

¹For a discussion of the racial issues in Malaysia and the New Economic Policy, see Snodgrass (1980), Andaya and Andaya (1982), Jomo and Shari (1986), Jesudason (1989) and Wyznan (1990).

II. The Data

The Second Malaysian Family Life Survey (MFLS2), conducted in 1988-89, includes a retrospective work history for men covering every job they held at age fifteen or after. The survey is based on a representative sample of Malaysian women under age 50, and the male work history comes from the men married to these women.² Thus only men married in 1988 qualified for the survey, but a good fraction (37%) of the years covered by the survey's work history are for men who were single at the time. The sample nevertheless underrepresents young men in 1988. This survey is the only source of longitudinal data on the work experience of individuals in Malaysia for this time period.

The recorded job histories in MFLS2 surely contain some recall error, but the problem is not different in kind from most other surveys: Cross-sectional surveys contain recall data about completed education, year of birth, job tenure, many characteristics of children and parents, etc. However, answers to detailed questions about the circumstances of previous jobs may exhibit more error than questions about more important life transitions. The systematic biases often observed in recall data are the tendency to foreshorten the time of long past events or forget about them, and to heap reported events on round-numbered years or ages (see, for example, Ewbanks, 1981). As seen in Figure 2.1, there is little indication that the reported year of starting a new job in this survey is heaped around round-numbered years.³ The smoothness of the reported years of events (and the accuracy) may be enhanced by asking a complete life history instead of requesting the date of an isolated event in the past. When reporting a life history, the respondents are more likely to consider the internal consistency of the sequence of events job in the U.S. Consumer Price Survey, often has pronounced heaping with spikes at five-year intervals (see Ureta, 1992, p.324).

Compared to panel data, retrospective data has the advantage of avoiding both attrition and panel conditioning, where respondents change their responses over time due to repeated interviewing. For example, in the U.S. Panel Study of Income Dynamics, the response rate fell below 50% after 10 years into the study due to attrition, and several studies have found that the characteristics of those who dropped out were different from those who remained in the sample (see Randolph and Trzcinski, 1989).

Prices over the period covered in the work histories, as shown in Figure 2.2, were very stable which augurs well for the recall of earnings in past times. The average number of years between the recalled start of a job and the survey interview was 13.3 years. In a life history survey, even if the respondents make up a representative cross-section of the population at the time of the survey, data for the earliest years of the recall histories all come from men who were young at that time. The men covered in the job histories become progressively less representative of the general population as one moves backward in time, because the "surviving" men become younger and younger. However, longitudinal data are the only means of capturing the typical *life experience* of individuals at a point in time. By tracing individuals over time we

²This part of MFLS2 is called the "New Panel" by Rand.

³Only about half of the respondents reported the date when they started a new job. The others reported their age at the beginning of the job. The reported ages of beginning a job (not shown) also exhibit no heaping.

can avoid the fallacy of composition common in the use of aggregate data, where changes due to the evolving composition of the population may be attributed to changing individual experience.

For each job in the work history, the men reported their monthly earnings at the starting and ending⁴ years of the job as well as other job characteristics. Each job was converted into an annual earnings profile by quadratic interpolation of their pay at the start and end of the job. Starting and ending earnings were used to estimate the curvature of the age-earnings profile in the sample, and the individual earnings profile interpolations were constrained to have the same rate of curvature.⁵ Where earnings at the end of the job were equal to or lower than at the start of the job, earnings were interpolated linearly in the intervening years. Data on earnings at the start and end of a job give us considerably more information than just two years of earnings. We know a good deal about what probably happened in between, and the interpolation accounts for this. Interpolation is also necessary to study the determinants of labor market earnings rather than per-job earnings because if men hold several jobs, the jobs often begin and end at different times, so interpolation is needed to arrive at a figure for total labor earnings. The average length of the interpolation of earnings was 5.8 years (because of censoring at the time of the survey, the average length of uncensored jobs is longer). All of the results below have been checked using only initial and final earnings, and using a linear interpolation of earnings. The qualitative results remain unchanged. The annual earnings profiles of each job were summed to give total annual earnings and deflated by the consumer price index to give an estimated real earnings profile. Length of job tenure was taken from the job with the highest salary when the worker held more than one job at once. Fourteen percent of the men held more than one job in any given year. All earnings and income figures are reported in Malaysian Ringgit (2.7 Ringgit = 1 \$US in 1988).

The accuracy and representativeness of the MFLS2 sample have been investigated extensively by others. Sine and Peterson (1993) compare the responses of the wives of the men studied here with other Malaysian data sources. They considered reported births, infant deaths, birthweights, contraceptive acceptance, breastfeeding and women's education. They report good general agreement with independent data sources except for infant deaths which are usually underreported in retrospective surveys. In particular, reported education agrees closely with census data from 1980. Klerman (1993) is more pessimistic about reported breastfeeding duration in MFLS2, finding not only heaping on certain monthly durations, but that reported duration is often far from the true duration. Smith, Thomas, and Karoly (1992) compare the correspondence of retrospective migration histories between the 1976-77 Malaysian Family Life Survey (MFLS1) and the reinterview of the same respondents in 1988-89 (a part of MFLS2 not used here). Since there is a gap of 12 years between the two surveys, they can only assess the consistency of events reported to be at least 12 years earlier. The rate of matches between the

⁴The earnings received from current jobs (at the time of the interview) is not available in the job history, though starting and ending earnings in all previous jobs are available. Current job information was collected as part of the current sources of income for all the family members in another part of the survey, but inevitably the current job data did not always match the start of job data. Minor changes in occupation code accounted for most of the discrepancy, but despite hand culling, 36 men out of 1513 had to be eliminated because their current job information was not consistent in the two parts of the survey. An additional 43 men were eliminated because some of the work variables in the work history were missing. 22 men were eliminated for missing values in the migration or marriage histories, so the final number of men in the sample is 1412.

⁵The regression is $Wage = -384 + 49.6 \text{ age} - 0.5150 \text{ age}^2$ $R^2 = 0.06$. (t statistics are in parentheses.) Hence, the individual interpolations were constrained to have a second derivative of -0.5150.

surveys for reported state of birth and state of residence at age 15 were 98% and 95% respectively for men. The match rate for district at birth and age 15 were 90% and 89% respectively during a period of major changes in district boundaries. Of all the moves reported in the first survey, 61% of them were reported in the resurvey of the same men twelve years later, and 54% of the moves were reported in the same district. The match rates may suffer from problems in coding the MFLS1 male migration data. Smith *et al.* find that consistent reporting between the two surveys is positively correlated with age and education. For men, surprisingly, time since the reported event did not affect recall accuracy, but the length of stay after the move strongly improved recall.

To assess whether the male earnings data in MFLS2 represents the Malaysian population, household income in MFLS2 in the year of the survey for those households with a male labor history is compared to household income data in published sources, since comparable earnings data are not available. The distribution of per capita income in MFLS2 households in 1988 is quite close to the World Bank's income distribution data for 1987 (see Table 2.1).⁶ Household income in the MFLS2 is probably somewhat less equally distributed because the sample excludes households of men and women who are not currently married, including female-headed households,⁷ and also because MFLS2 probably has a more ample measure of monetary income.

The MFLS2 data are compared with the Malaysian government's Household Income Survey (HIS) for 1987 (expressed in 1988 Ringgit) in Table 2.2. The total household income is close, given that the Malaysian GDP per capita grew by 8% in 1987-88.⁸ The ethnic composition comes out rather differently. MFLS2 estimates a higher household income for the Chinese, and a lower income for the Malays, and especially for the Indians compared to total income in the HIS. Both HIS and MFLS2 household income estimates are much lower than GDP per capita figures, as is common with income surveys. Mean household income *per capita* in MFLS2 was 2762 Ringgit in 1988, or 1017 US\$, while the World Bank (1990, p. 179) estimate for GDP per capita in Malaysia in 1988 was 1940 US\$.

The MFLS2 data *for 1988* seems generally representative of the Malaysian population when compared with other data sources. What this means in terms of the retrospective work history for men is that the work history ends up in about the same place as other data sources at its termination point in 1988. In terms of trends in earnings by ethnic group, though, the work history is quite different from trends in cross-sectional surveys. The trends in average earnings by ethnic group from the male labor history in MFLS2 are shown in Figure 2.3. Several parts of this picture are unexpected. Earnings in 1960 are relatively equal for all ethnic groups. Chinese earnings pulled away from the earnings of the other ethnic groups just around 1970 when the New Economic Policy was introduced with the purpose of *helping* poorer Malays. Indians fared neither worse nor better than the Malays in these trends. Are these trends to be believed?

Figure 2.3 could simply reflect that the Chinese earnings profile over the course of their working life was steeper than the profiles of Malays and Indians. Since these trends come from recall data, the average age of the respondents in the sample was rising (see Figure 2.4). This

⁶Household income includes earnings of all household members as well as rent payments received on land and buildings, and vehicles, dividend and interest receipts, and various benefits. Gifts from family members are available in MFLS2, but are not included.

⁷The MFLS2 includes households consisting of, or headed by unmarried women over the age of 18, but they are not included in this assessment of the representativeness of male wages data.

⁸All of the households used here (the "New Panel" for males), were interviewed in 1988.

caused the gap between Chinese and Malay earnings to widen as the sample got *older* on average, but it doesn't necessarily imply that the gap for a constant-age cross section of Malays and Chinese was widening over *time*. However, when we look at the profile of earnings by age for specific cohorts of Malays, Chinese and Indians, although the earnings profiles for Chinese were steeper, they rose over time, so that age composition is not the whole story (see Figures 2.5, 2.6, and 2.7). For instance, for the cohort of workers which entered the labor force in the years 1969-72, the earnings of Chinese workers rise faster with age than the earnings of Malays. However, the earnings of Chinese workers who entered the labor force between 1977-80 were higher at all ages than the 1969-72 cohort, while the earnings of Malays were roughly the same by age for the 69-72 and 77-80 cohorts.

The growing discrepancy between Malay and Chinese earnings shown in Figures 1.1 and 2.3 does *not* agree with data from other sources on trends in the income of different ethnic groups. Two series of cross-sectional household surveys have been pieced together. Figure 2.8 shows the trend in a group of Malaysian government surveys and also the cross-sectional current household income results in the year of survey from MFLS1 and MFLS2. The Malaysian government series is comprised of the 1957-8 Household Budget Survey (HBS), the 1970 Post-Enumeration Survey (PES) of the census, the 1973 Household Expenditure Survey (HES), and the 1984 and 1987 Household Income Surveys (HIS). The 1957-8 HBS surveyed 2,760 households, the 1970 PES 28,000 households, the 1973 HES over 7,000 households, and the 1984 and 1987 HIS each surveyed about 40,000 households. MFLS1 had 1,262 households and MFLS2⁹ has 3,541 households. All the surveys except for the two HIS and the two MFLS used somewhat different definitions of income (see Snodgrass, pp. 70-76, and Bhalla and Kharas, p. 43).

The surveys making up the series in Figure 2.8 are not perfectly comparable, but given their sample sizes, they are probably roughly accurate, even piecing together a trend. The cross-sectional results from the two MFLS agree with the improvement in Malay/Chinese earnings since the mid-1970s, although the 1976-77 MFLS1 has a slightly lower estimate of Malay/Chinese income than the 1976 HIS (42% vs. 44%), and the 1988-89 MFLS2 has a considerably lower estimate of relative income than the 1987 HIS (52% vs. 61%).

How can the trend in Figure 2.8 coexist with the trend in the earnings histories of men in MFLS2 in Figure 1.1? Before 1988, the MFLS2 labor history had a younger sample than the general population. This made Malay earnings closer to Chinese earnings in the labor history trend because Malays had flatter age profiles of earnings than the Chinese. However, the steepness of the age profiles does not explain the whole trend as seen in Figures 2.5 and 2.6. Chinese age profiles shifted up over time, while Malay age profiles did not. This means that things are worse for younger Malay cohorts compared to the Chinese than it was for older cohorts in 1988. For this deteriorating cohort trend to be consistent with the improving population trend, as indicated in Figure 2.8, the later cohorts of Malays who were under fifty in 1988 must have had a greater disparity with the Chinese than the over-fifty workers in 1988. Without the older cohort counteracting the deteriorating trend in the under 50 cohorts, the population trend could not have improved.

⁹The MFLS2 sample used in Figure 2.2 (from Lefranc, 1992, p.61) includes all households, not just the households with a husband under 50 reporting a wage history as in the rest of this paper

If the relationship between the trends in the labor history data and in the population are to be made consistent, they would appear as in Figure 2.9. The ratio of Malay to Chinese earnings in the MFLS2 male labor history converged to about 50% in 1988 which was the ratio of incomes in the MFLS2 household cross-section, when the composition of the men in the labor history should have been most representative of the population. In the labor history, however, the Malay to Chinese earnings ratio started at about 90% in 1960, while in the cross-sectional surveys, the Malay to Chinese income ratio started at about 40% in 1960. In the "Cohort" figure below, the trapezoid ABCD represents the labor history sample. The darker the diagonal band in the figure, the more equal are Malay and Chinese earnings. In 1960, the Malay/Chinese earnings ratio *in the labor history sample* was very equal (90%) in its cross-section AB. In the population, though, the relative earnings could be quite unequal (40% ratio) if unobserved older cohorts in 1960 had unequal earnings. The cross-sections of the labor history become less and less equal as the disparate younger cohorts became part of the sample, but the population average could nonetheless improve, since the unequal older cohorts were leaving the population, with both samples ending up with somewhat unequal relative earnings (50% ratio).

The diagram shows how it is possible for both the household income surveys and the labor history data to be correct, and it may also give us a glimpse at the future prospects for the population as a whole. If workers are more or less locked in to an earnings profile when they enter the labor market (which will be examined below), the relative earnings experience of young cohorts today foreshadows the state of the population tomorrow. If the deteriorating cohort trend for Malays continues, with young Malays entering the labor force with the same age profile of older Malays, but Chinese entering the labor force with higher age profiles than older Chinese, then the disparity between Malays and Chinese will get worse in the future, even though it appears to have narrowed in the population since the mid 1970s. For the relative earnings of Malays and Chinese to continue to converge in the scenario in Figure 2.9, entering cohorts must have closer relative earnings than the exiting cohorts in the population, which would be a sharp break with the trend of recent young cohorts. What is more, older cohorts of male workers have a greater weight in measures of household income *per capita* (and welfare of the population) than younger cohorts since they have more dependents that share their income. So new younger cohorts of Malays and Chinese would have to be closer to parity than exiting cohorts to counteract the affect on relative welfare of the aging of currently unequal young cohorts.

In the next section, the hypothesis that the trend in relative ethnic earnings in the MFLS2 retrospective work history is due to cohort effects will be evaluated with the data.

III. Earnings estimation

The log-linear earnings regression is one of the more successful specifications in empirical econometrics. As noted by Willis (1986), the mystery is why this specification works so well in so many different contexts. One reason is that the log-normal distribution "fits actual earnings distributions quite well, perhaps better than any other rather simple distribution, [although] it is generally unable to predict the relatively numerous occurrences of very high earnings." (Berndt, 1990, p.161) Log-linear earnings functions may be viewed as an approximation of an individual's intertemporal wealth maximization where education and on-the-job training are human capital investments under some more or less reasonable assumptions (see Willis (1986), Schultz (1988a), and Mincer (1974)). The earnings function can also be viewed as the result of

an individual's intertemporal wealth maximization where education makes no contribution to students' productivity, but only provides a "signal" of their inherent abilities to employers, and job experience also makes no contribution to productivity, but earnings rise with experience to reward workers that firms have learned are more productive (a "matching" model). It is not necessary to resolve whether education and job experience contribute to an individual worker's productivity because even if they don't, they are still a determinant of his or her earnings.

The log-linear earnings equation specifies the relationship between individual characteristics and labor market earnings as in Equation 3.1:

$$\ln y_i = \ln y_0 + \mathbf{b}_1 s_i + \mathbf{b}_2 X_i + \mathbf{b}_3 X_i^2 + \mathbf{b}_4 t + \mathbf{Z}_i' \boldsymbol{\gamma} + \mathbf{e}_i \quad (3.1)$$

$\ln y_i$ is the natural logarithm of mean real monthly labor market earnings for individual i , $\ln y_0$ is the log of a constant earnings that all workers receive regardless of their characteristics, s_i is years of education, X_i is years of labor market experience, t is time in years, and \mathbf{Z}_i is a vector of other individual characteristics. $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$, and $\boldsymbol{\gamma}$ are parameters, and \mathbf{e}_i is a normally distributed random error.¹⁰ Under certain assumptions, for the wealth-maximizing individual \mathbf{b}_1 is the internal rate of return to an additional year of education (see Willis, 1986, pp. 529-532 for a derivation).

The first regression in Table 3.1 shows the growth rate of earnings in the MFLS2 sample controlling for the changing experience composition (or age composition) and race. The growth rate of real earnings was 0.8%, clearly lower than the earnings growth rate for the Malaysian population as a whole over the period covered by MFLS2 life histories. From Independence in 1957 to 1988, the average real growth rate of GDP per capita was 3.8%.¹¹ The low growth rate of earnings in the sample may be an indication of recall bias in the interview responses. However, the more usual pattern for recall error is to foreshorten the time since past events, which would tend to inflate the growth rate of earnings (lower earnings in the past would tend to be reported more recently than they actually occurred). Malays in the MFLS2 sample earned on average 69% of the earnings of Chinese, the excluded ethnic group in the regression.¹² Indians earned 76% of Chinese earnings over the sample period.

The log-linear earning regression in column 2 of Table 3.1 relates earnings to individual and job characteristics. The earnings profile grew by 12.5% per year of job experience at the start of working life and peaks at 18 years of experience. Job tenure, the amount of time spent in a particular job,¹³ has a much weaker effect on earnings than experience. Tenure has a small negative affect on earnings for the first ten years of a job. If tenure is considered to be

¹⁰ Note that in Equation 3.1 the effect of experience on earnings peaks at age $-\mathbf{b}_2 / 2\mathbf{b}_3$, \mathbf{b}_4 is the annual growth rate of earnings over time not due to other characteristics, and if an element of \mathbf{Z}_i , say Z_{i1} , is a dummy variable, then a person with characteristic Z_{i1} has earnings $e^{\mathbf{g}_1}$ times the earnings of a person without characteristic Z_{i1} where \mathbf{g}_1 is the element of $\boldsymbol{\gamma}$ that corresponds to Z_{i1} .

¹¹ Calculated from International Monetary Fund (1992).

¹² There is a small number of people with other ethnicities who are included with the Chinese in the earnings regressions for the whole sample, since they have characteristics closest to the Chinese. The characteristics of this group is very diverse, and it is probably mostly made up of poor Thais and wealthy Europeans. These "other" people made up 0.7% of the population of Peninsular Malaysia in the 1980 census (Information Malaysia Yearbook, 1992-93, p.76) and 1.6% of the men in the MFLS2 sample used here. These men are not included in the race-specific earnings regressions below.

¹³ In MFLS2, a change of job requires that the occupation of the worker change (see Davanzo, *et al.* 1993, pp. 22, A-58). Job tenure might better be called industry tenure.

endogenous in the decisions of workers and firms, a negative correlation of tenure and earnings implies that better workers, who get paid more, are more likely to change jobs. This is not consistent with a simple description of job-specific human capital investment. Education has a 7.6% rate of return, which is low for Asia. Psacharopoulos (1985, p.588) reports an average rate of return of 11% for Asia, 13% for Africa, and 14% for Latin America in studies using the Mincer-type earnings equations used here. The inclusion of other characteristics, especially education, mitigates the gap between the Malays and Indians on one hand and the Chinese on the other. Even with differences in experience, tenure, and education accounted for, Malays and Indians still earned only 76% and 83%, respectively, of Chinese earnings.

The time trend in Regression 2 peters out entirely, becoming slightly negative, but not significantly different from zero. The "Business Cycle" is the deviations from trend in GDP, and it varies positively with earnings as expected.¹⁴ Like most other earnings data, the MFLS2 earnings history data are highly correlated over time. The first order autocorrelation (AR(1)) coefficient, ρ , is 0.94, and strongly significant. The autocorrelation is probably accentuated by the interpolation of the earnings data. When interpolating between two points that are each subject to random shocks, a portion of the shocks will be transmitted to the interpolated points between the two endpoints, causing some spurious autocorrelation. ρ and the variable coefficients are estimated with a panel data AR(1) by maximum likelihood. The fit, as indicated by the adjusted R^2 , improves considerably from Regression 1 to Regression 2, mainly because of the estimation of autocorrelation.

Column 3 in Table 3.1 explores some additional labor market variables. Working in an urban area (large town or city, as reported by the survey respondents) did not imply much higher earnings. The earnings of urban workers were 104% those of rural workers with similar characteristics. This seems to indicate there is not a large cost of living difference between the country and the city. Hirschman (1984) argues that the rural population was extensively involved in the market economy, and most transactions were monetized, which makes large urban-rural price differences less likely. Payment of a part of earnings in kind had a slight negative effect on money earnings, as one would expect since in-kind goods substitute for monetary payment. Holding at least one job part-time is more likely for lower paid workers, who only earned 80% of the earnings of workers who worked exclusively full-time. Only 10% of the jobs in the sample were part-time. Conditional on part-time status and the other independent variables, the number of jobs held slightly increased the earnings of Malaysian men. The additional labor market variables used in this regression had a slight effect on earnings except for part-time work, and together they had almost no influence on the relationship between other variables and earnings.

Table 3.2 considers the determinants of earnings separately for Malays, Chinese, and Indians. Chinese had a steeper and more concave experience profile than the Malays and the Indians. All three groups received their highest earnings at 18-19 years of work experience, but peak Chinese earnings were 5.6 times their initial earnings while peak Malay and Indian earnings were 2.3 and 2.9 times their initial earnings, respectively. Earnings were convex in tenure,

¹⁴ The Business Cycle variable is made up of the residuals of the following regression of the natural logarithm of GDP on time in years (t statistics in parentheses):

$$\ln \text{GDP} = -105.28 + 0.066 \text{ Time}, \quad R^2=0.98$$

(34.67) (42.684)

The GDP data are from the IMF's International Financial Statistics.

falling initially for Chinese and Indians, but tenure has a much smaller effect on earnings than job experience.

Malays had the highest rate of return to a year of schooling at all education levels during a period of rapidly rising Malay education relative to other ethnic groups. Within this sample, Malay education had caught up with and overtaken Chinese and Indian education (see Figure 3.1). The flattening out of education trends in the 1980s in Figure 3.1 was at least partly an artifact of having few new "entrants" to the recall data in the 1980s (i.e. most earnings history respondents entered the labor force before the mid-1980s so the educational composition of the recall sample did not change much after then). The timing of these ethnic trends in education is exactly what one would expect from an effective affirmative action program coming on line in the 1970s. Changes in government education policy in the 1960s and early 1970s, an irritant to non-Malays which contributed to the May 13, 1969 race riots, coincided with a stagnation in Chinese and Indian educational progress, and an acceleration in Malay attainment. In 1965, lower secondary school was made accessible to all students, but since government-supported Chinese-language secondary schools were abolished in 1961, this helped the Malays more than the Chinese. Starting in 1971, English-language schooling was phased out one school level per year until all public school classes were taught in Malay. Malays are given preference among university applicants.¹⁵

The pattern of rates of return in Table 3.2 between primary, secondary, tertiary education are unusual. For each race, the rates of return were lowest for primary school and highest for tertiary schooling, whereas in most other countries, the highest rate of return is in primary school (see Psacharopoulos, 1985). Smith (1991) also found that returns to education were highest in primary school and lowest for tertiary schooling using the 1976 MFLS1. This may indicate good access to primary schools and difficult access to higher education in Malaysia, or it may mean that the economy has shifted its labor demand to higher skilled workers faster than the population is increasing its education levels.

The rates of return to education are also unusual because they are all so low by international or Asian standards. An unhappy possibility is that Malay rates of return are highest because Chinese and Indian rates of return were depressed by the switch to Malay-language instruction. Smith (1991) found rates of return to education in Malaysia of 12.7%, 7.5%, and 10.6% for Malays, Chinese, and Indians, respectively, for all education levels combined using the MFLS1, which mostly covered men educated before the 1960s. In the late 1960s and early 1970s, there was a shortage of qualified Malay teachers and Malay textbooks (Snodgrass, 1980, p.250), and non-Malay students were required to learn in a second language, not their native tongues. The New Economic Policy's goals in education seem to have been successful in that Malay education levels have been boosted and their rate of return to education in the labor market is higher than for other ethnic groups, so that the preferential access for Malays does not seem to have caused employers to discount the value of Malay educational credentials.¹⁶ However, the progress may have been at the absolute expense of the other groups, not just a relative gain. It could also be that education had a low rate of return for all groups because of low educational quality or good access to schools, and Malay rates of return were higher because

¹⁵See Andaya and Andaya (1982) pp. 278-280, and Snodgrass (1980), Chapter 9 for analysis of changes in education policy.

¹⁶This may indicate that education is more human capital than signal.

the government and private firms wanted to hire educated Malays to fulfill affirmative action goals.

The time trends for earnings of the ethnic groups separately in Table 3.2 are very low like the time trend for the whole sample. During a period when the incomes of all Malaysians grew rapidly and Malays seem to have caught up with Chinese incomes, the earnings of Malays in the MFLS2 sample are estimated to fall during the sample period by 1% per year, other things being equal. One *would* take other variables in the regression to be equal, except for education, if one wants to look at the whole Malaysian population. There is no reason to expect big changes in average job experience, tenure, or the business cycle in the Malaysian economy over this period. Education levels have not risen sufficiently to entirely offset the estimated decline in Malay earnings in the regression during a period when we know from other data sources that Malay real incomes have grown rapidly.

The time trends do fit the time trend in relative Malay/Chinese earnings in Figure 1.1. From Table 3.2, Chinese earnings are growing 2.1% faster than Malay earnings in the sample. If Malay earnings were 90% of Chinese earnings in 1960 in the sample, the estimated time trends imply that they should have fallen to 50% of Chinese earnings in 1988, as occurs in Figure 1.1.

A possible cause of stagnating Malay earnings is the inability of Malays to break into the more lucrative occupations. If Malays are unable to move from rural agriculture to urban clerical, professional, or managerial work, their earnings profiles would remain low and flat over time. In Figure 3.2, Malays have a lower representation in the higher paying occupations than the Chinese, but they have moved in large numbers from agriculture into the military and police, clerical and sales jobs, and white collar jobs. So occupational segregation of Malays is not a likely explanation for the failure of Malays to catch up to Chinese earnings in the MFLS2 sample. Chinese occupations have been more stable, with some movement out of agriculture, into blue collar jobs, out of clerical and sales jobs, and into white collar jobs. Indians have become scarcer in agriculture, and more numerous in blue collar jobs. Jomo and Ishak (1986, pp. 83-90) argue that in Malaysia as a whole, the New Economic Policy has largely achieved its employment restructuring goals of giving Malays more access to higher paying urban occupations.

Cohort effects, as proposed in the previous section, could reconcile the decline in Malay earnings in the sample with a rise in Malay earnings in the population, and the divergence of Malay and Chinese earnings in the sample with the convergence of Malay and Chinese earnings in the cross-sectional surveys. As seen in Table 3.3, though, cohort effects do not resolve the apparent inconsistency. If they did, the time trend for Malays should become positive and larger than the time trend for Chinese, and the cohort effect for Malays should be negative, more negative than the cohort effect for Chinese. Then Malays could have caught up with Chinese earnings and had a positive time trend in the population (due to pre-sample Malay cohorts with low earnings relative to the Chinese), but Malay earnings would decline in the sample due to the declining trend in Malay/Chinese earnings by cohort. Malays, though, have no estimated cohort effect, and the Chinese have a negative effect. The rate of return to education for Malays appears to have risen by cohort, and the rate of return for Indians was falling, which is alarming because it was not very high to begin with.

The permanence of cohort effects on earnings profiles is not as important an issue when there are no cohort effects, but in Table 3.4 it is clear that experience profiles are not locked in for each cohort. Time has a stronger effect on the shape of the earnings-experience profile than

cohort does, indicating that earnings profiles are more responsive to aggregate demand trends that affect all experience levels rather than cohort vintage effects.

The cohort variable used in the regression analysis could lack sufficient independent variability to be accurately estimated. If experience were defined as age minus education minus 6, as it often is in cross-sectional analysis, cohort, experience and time would not all be identifiable, because cohort would be a linear combination of other included variables: cohort would equal time - experience - education - 6. Cohort is identified in MFLS2 data because experience counts the actual years of paid work - it is not a linear combination of age and education. Cohort is identified by the variability across individuals in years between school and work, spells of unemployment, and periods out of the labor force. These gaps could be insufficient, though, to get a good estimate of cohort effects separate from the effects of time, experience and education. However, if time is dropped from the regression, there is no identification problem, and cohort effects should appear in the regression if they exist. The cohort effects are still insignificant for Malays and Indians, and negative for Chinese.

IV. Conclusion

The exploration of the MFLS2 data demonstrates that retrospective data should not be used uncritically to reconstruct what has happened to the population over time. However, if the survey data *do* represent the typical experiences of people alive at the time of the survey, it can be of more interest than what happened to the population over time. It is often because people do not distinguish between what has happened to people alive now, and what has happened to everyone over time, that they are interested in population trends.

There are three general explanations for the difference in the trends in the MFLS2 labor history data and the trends in large cross-sectional surveys. The first is that the retrospective life history data are correct, but the life history is not representative of the population in the past. The second is that there are errors in the life history data that make it unrepresentative of the experience of the people surveyed. The third is that the cross-sectional surveys are not accurate. This last seems unlikely because of the large sample sizes and independent sources of the cross-sectional data.

The cohort effect hypothesis is an example of the first kind of explanation. It presumes that the sample is representative of the cohorts which it covers and successive cohorts have systematic differences, but this hypothesis is rejected by the data. Although the men and the families in MFLS2 appear to have been broadly representative of Malaysia in 1988, the men's life histories may not be representative for other reasons. An implicit selection process determines who is around to represent the characteristics and events of people in the past in a retrospective survey. Most obviously, the people chosen for a retrospective survey don't accurately represent the age of the population in the past, but they may be unrepresentative in other respects as well.

An explanation of the second kind for the flat time trend in Malay earnings in the sample and the divergence of Malay and Chinese earnings is recall bias. The more commonly recognized recall bias is the foreshortening of the recalled time since past events occurred. Foreshortening the time or age of reported earnings would tend to increase the measured growth rates when earnings are growing. Malays have to be *exaggerating* the time since the start of past jobs in this scenario, and the more wealthy and urban Chinese must be foreshortening the time

since their jobs started, which is usually associated with people less involved in structured society.

A different sort of recall bias would be to take a rosy view of the past, or to be pessimistic about one's progress. The time trends in the sample could be explained by Malays and Indians exaggerating their past earnings while reporting their current earnings correctly. If the Chinese behaved differently and reported past earnings as low as they really were, the data would make the Malays and Indians appear to lose ground to the Chinese over time.

It is important to distinguish whether the MFLS2 earnings history data are accurate but unrepresentative of the population in the past, or if they are subject to recall bias, in order to draw conclusions about trends in the labor market. This could be investigated by matching the MFLS1 male earnings history with the resurvey of the same men in MFLS2. If the MFLS2 earnings data are accurate, the time trends in earnings are cause for concern. For a good-sized sample that seems to be a representative cross-section of the Malaysian population in 1988, earnings have been stagnating or falling for Malays and Indians, despite Malay progress in education, and entering higher paying occupations. Another worrisome finding is the low rates of return to education, especially for Indians.

References

- Andaya, Barbara Watson, and Leonard Y. Andaya (1982). *A History of Malaysia*. London: MacMillan.
- Berndt, Ernst R. (1990). *The Practice of Econometrics: Classic and Contemporary*. Reading, MA: Addison-Wesley.
- Bhalla, Surjit, and Homi Kharas (1992). "Growth and Equity in Malaysia: Policies and Consequences," in Teh Hoe Yoke and Goh Kim Leng, ed., *Malaysia's Economic Vision: Issues and Challenges*. Petaling Jaya, Malaysia: Pemanduk Publications.
- Cho, George (1990). *The Malaysian Economy: Spatial Perspectives*. London: Routledge.
- Da Vanzo, Julie, John G. Haaga, Tey Nai Peng, Ellen H. Starbird, and Christine E. Peterson (1993). *The Second Malaysian Family Life Survey: Survey Instruments*. Santa Monica, CA: Rand Corporation.
- Ewbanks, D.C. (1981). *Age Misreporting and Age-Selective Underenumeration: Sources, Patterns, and Consequences for Demographic Analysis*. Washington, D.C.: National Academy Press.
- Hirschman, Charles (1984). "The Society and its Environment," in F.M. Bunge, ed., *Malaysia: A Country Study*. Area Handbook Series. Washington, D.C.: Department of the Army. pp. 67-128.
- Information Malaysia Yearbook 1992-93* (1992). Kuala Lumpur: Berita Publishing.
- International Monetary Fund (1992). *International Financial Statistics*. Washington, DC: International Monetary Fund.
- Jesudason, James V. (1989). *Ethnicity and the Economy: The State, Chinese Business, and Multinationals in Malaysia*. Singapore: Oxford University Press.
- Jomo, K.S., and Ishak Shari (1986). *Development Policies and Income Inequality in Peninsular Malaysia*. Kuala Lumpur: Institute of Advanced Studies, University of Malaya.
- Klerman, Jacob Alex (1993). *Heaping in Retrospective Data: Insights from Malaysian Family Life Surveys' Breastfeeding Data*. Rand Labor and Population Program Working Paper 93-21, Santa Monica, CA, July.
- Lang, Kevin (1994). "Does the Human Capital/Educational-Sorting Debate Matter for Development Policy?" *American Economic Review* 84(1):353-358, March.
- Lefranc, Christophe (1992). *Demographic Influences on Income Distribution: A Study of Household Survey Data from Peninsular Malaysia*. Ph.D. Thesis, Department of Demography, University of California at Berkeley.
- Malaysia (1989). *Yearbook of Statistics Malaysia 1988*. Kuala Lumpur: Department of Statistics, July.
- Malaysia (1991). *Yearbook of Statistics Malaysia 1990*. Kuala Lumpur: Department of Statistics, July.
- Mincer, Jacob (1974). *Schooling, Experience, and Earnings*. New York: NBER.
- Psacharopoulos, George (1985). "Returns to Education: A Further International Update and Implications," *Journal of Human Resources* 20(4):583-604, Fall.
- Randolph, Susan and Eileen Trzcinski (1989). "Relative Earnings Mobility in a Third World Country," *World Development*. 17:4, pp. 513-524.
- Sine, Jeffrey and Christine E. Peterson (1993). *The Second Malaysian Family Life Survey: Quality of Retrospective Data for the New Sample*. Santa Monica, CA: Rand Corporation.
- Smith, James P. (1991). "Labor Markets and Economic Development in Malaysia," in T. Paul Schultz, ed., *Research in Population Economics*. Greenwich, CT: JAI Press.
- Smith, James P., Duncan Thomas, and Lynn A. Karoly (1992). "Migration in Retrospect: Differences between Men and Women," manuscript, Rand Corporation, April.
- Snodgrass, Donald R. (1980). *Inequality and Economic Development in Malaysia*. Kuala Lumpur: Oxford University Press.
- Ureta, Manuelita (1992). "The Importance of Lifetime Jobs in the U.S. Economy, Revisited," *American Economic Review*. 82:1, pp. 322-335, March.

- Willis, Robert J. (1986). "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions," in O. Ashenfelter and R. Layard, eds., *Handbook of Labor Economics*, Vol. 1. Amsterdam: Elsevier.
- World Bank (1990) *World Development Report 1990*. Oxford: Oxford University Press.
- WYZAN, MICHAEL (1990). "ETHNIC RELATIONS AND THE NEW ECONOMIC POLICY," IN MICHAEL WYZAN, ED., *THE POLITICAL ECONOMY OF ETHNIC DISCRIMINATION AND AFFIRMATIVE ACTION*. NEW YORK: PRAEGER.

TABLE 2.1 SHARE OF PER CAPITA HOUSEHOLD INCOME BY QUINTILE

	MFLS2	World Bank
Lowest 20 percent	4.4	4.6
Second Quintile	8.3	9.3
Third Quintile	12.5	13.9
Fourth Quintile	20.4	21.2
Highest 20 percent	54.3	51.2
Highest 10 percent	38.7	34.8

Source: Author's calculations, and World Development Report, 1990.

Table 2.2 MFLS2 versus Household Income Survey

	Mean Annual Household Income 1988 Ringgit			
	Malays	Chinese	Indians	Total
MFLS2 (88)	11,449	23,324	12,974	14,976
HIS (87)	10,682	17,599	13,402	13,217
Ratio (%)	107	133	97	113

Source: Author's calculations, and Wyzan, p.60.

TABLE 3.1 EARNINGS REGRESSIONS - WHOLE MFLS2 SAMPLE

	1	2	3
Constant	5.80 (386.58) **	5.00 (111.38) **	5.00 (109.27) **
Experience	0.089 (41.05) **	0.125 (33.13) **	0.121 (31.85) **
Experience ²	-0.0028 (35.42) **	-0.0034 (29.46) **	-0.0032 (28.12) **
Tenure		-0.015 (8.38) **	-0.012 (6.75) **
Tenure ²		0.0007 (6.74) **	0.0006 (5.36) **
Education		0.076 (18.62) **	0.074 (17.79) **
Malay	-0.37 (38.26) **	-0.28 (8.66) **	-0.26 (7.99) **
Indian	-0.28 (18.18) **	-0.20 (3.96) **	-0.19 (3.78) **
Time	0.008 (10.39) **	-0.002 (1.03)	-0.002 (0.76)
Business Cycle		0.113 (2.25) *	0.111 (2.22) *
Urban			0.04 (4.94) **
In Kind			-0.01 (1.25)
Part Time			-0.23 (16.15) **
Number of Jobs			0.03 (5.66) **
ρ		0.94 (384.82) **	0.94 (385.64) **
Adjusted R ²	0.16	0.72	0.71

(t statistics are in parentheses) * significant at the 5% level

** significant at the 1% level Time = 0 in 1960

TABLE 3.2 EARNINGS REGRESSIONS BY ETHNICITY

	1 Malay	2 Chinese	3 Indian
Constant	5.07 (64.65) **	4.76 (34.08) **	5.17 (44.61) **
Experience	0.096 (19.72) **	0.188 (22.39) **	0.107 (15.26) **
Experience ²	-0.00270 (17.55) **	-0.00510 (21.74) **	-0.00273 (12.58) **
Tenure	0.002 (0.95)	-0.034 (10.21) **	-0.026 (7.85) **
Tenure ²	0.00044 (2.91) **	0.00103 (4.85) **	0.00079 (3.45) **
Primary Education	0.056 (3.97) **	0.030 (1.16)	0.036 (1.71) *
Secondary Education	0.061 (8.46) **	0.044 (3.17) **	0.035 (3.35) **
Tertiary Education	0.096 (14.69) **	0.080 (6.59) **	0.075 (7.54) **
Time	-0.010 (3.66) **	0.011 (1.99) *	0.001 (0.19)
Business Cycle	0.134 (1.96) *	0.002 (0.02)	0.176 (1.85) *
ρ	0.93 (256.40) **	0.95 (216.28) **	0.94 (178.43) **
Log Likelihood	-334	-587	511
Adjusted R ²	0.76	0.62	0.78

(*t* statistics in parentheses) * significant at 5% level ** significant at 1 level

TABLE 3.3 EARNINGS REGRESSIONS BY ETHNICITY WITH COHORT EFFECTS

	1 Malays	2 Chinese	3 Indians
Constant	4.81 (55.33) **	4.77 (32.59) **	4.94 (32.32) **
Experience	0.082 (11.33) **	0.154 (14.07) **	0.071 (5.81) **
Experience ²	-0.00267 (17.28) **	-0.00506 (21.71) **	-0.00264 (12.17) **
Tenure	0.002 (0.85)	-0.034 (10.11) **	-0.024 (7.46) **
Tenure ²	0.00045 (2.98) **	0.00102 (4.81) **	0.00067 (2.91) **
Education	0.105 (10.50) **	0.063 (3.92) **	0.084 (5.49) **
Time	0.003 (0.51)	0.043 (4.84) **	0.035 (3.23) **
Business Cycle	0.136 (1.98) *	0.004 (0.04)	0.178 (1.87) *
Cohort	-0.000 (0.05)	-0.044 (2.87) **	-0.010 (0.68)
Cohort x Education	0.00187 (2.51) **	0.00042 (0.29)	-0.00348 (3.31) **
ρ	0.93 (258.79) **	0.94 (213.03) **	0.94 (176.21) **
Adjusted R ²	0.76	0.63	0.79

(t statistics are in parentheses) * significant at the 5% level
 ** significant at the 1% level Cohort = year of birth - 1940

TABLE 3.4 EARNINGS REGRESSIONS: EXPERIENCE PROFILE BY ETHNICITY

	1 Malay	2 Chinese	3 Indian
Constant	4.68 (50.42) **	4.51 (28.86) **	4.69 (29.51) **
Experience	0.160 (121.33) **	0.281 (15.06) **	0.193 (9.68) **
Experience ²	-0.00709 (8.94) **	-0.01094 (9.67) **	-0.00950 (8.03) **
Tenure	0.001 (0.39)	-0.036 (11.07) **	-0.026 (7.95) **
Tenure ²	0.00052 (3.39) **	0.00114 (5.55) **	0.00072 (3.12) **
Education	0.105 (10.47)	0.074 (4.50)	0.083 (5.47) *
Time	0.013 (1.74) *	0.086 (8.02) **	0.034 (2.71) **
Business Cycle	0.134 (1.95) *	0.068 (0.68)	0.195 (2.06) *
Cohort	-0.003 (0.27)	-0.065 (3.89) **	0.005 (0.32)
Cohort x Education	-0.00187 (2.49) **	-0.00045 (0.30)	-0.00345 (3.28) **
Cohort x Experience	0.00021 (0.28)	0.00543 (4.79) **	-0.00193 (1.71) *
Cohort x Experience ²	0.00001 (0.56)	-0.00010 (2.49) **	0.00001 (0.47)
Time x Experience	-0.00387 (6.13) **	-0.01165 (13.63) **	-0.00341 (3.50) **
Time x Experience ²	0.00019 (10.44) **	0.00042 (16.37) **	0.00023 (9.24) **
ρ	0.93 (258.86) **	0.95 (213.15) **	0.94 (175.37) **
Log Likelihood	-306	-408	564
Adjusted R ²	0.76	0.64	0.79

(*t* statistics are in parentheses) * significant at the 5% level ** significant at the 1% level

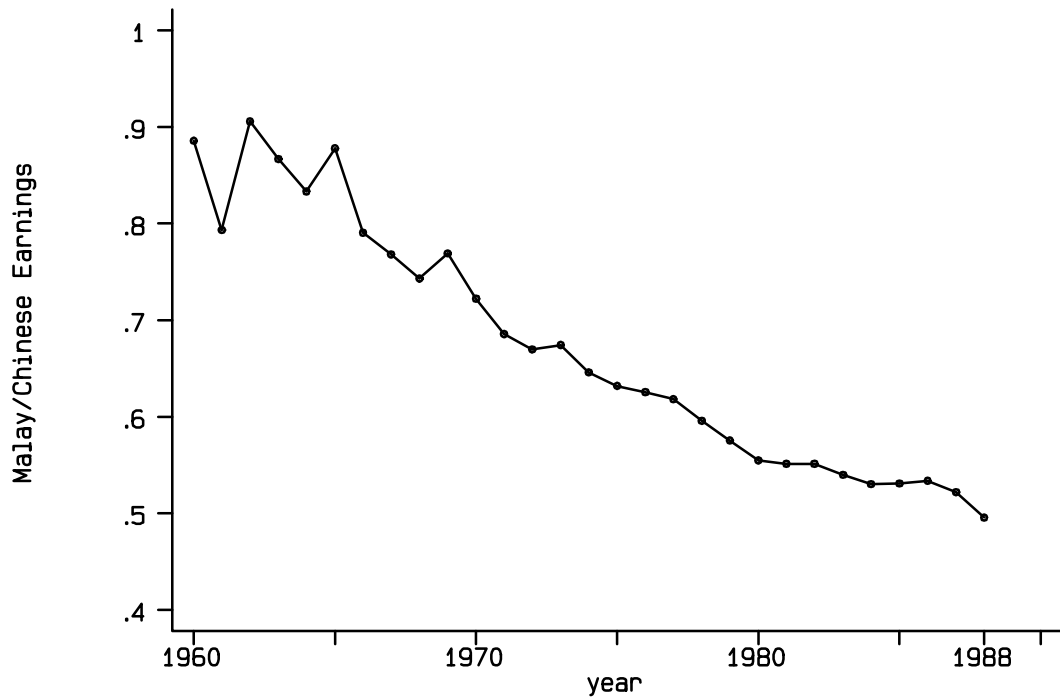


FIGURE 1.1 TREND IN MALAY/CHINESE EARNINGS IN MFLS2

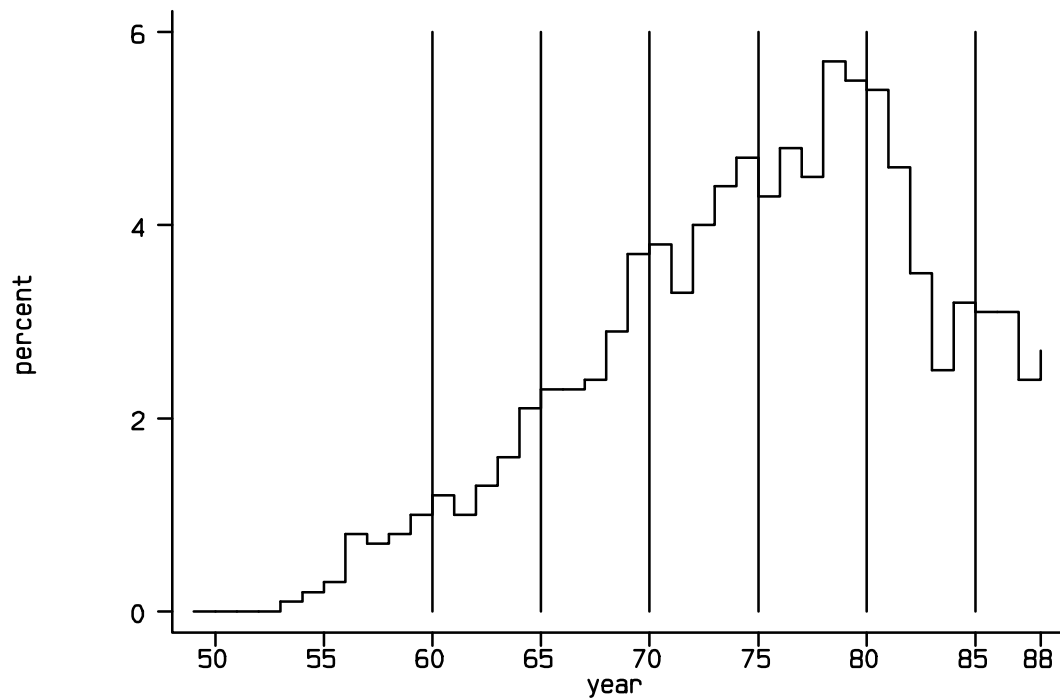


FIGURE 2.1 REPORTED YEAR OF STARTING JOB IN MFLS2

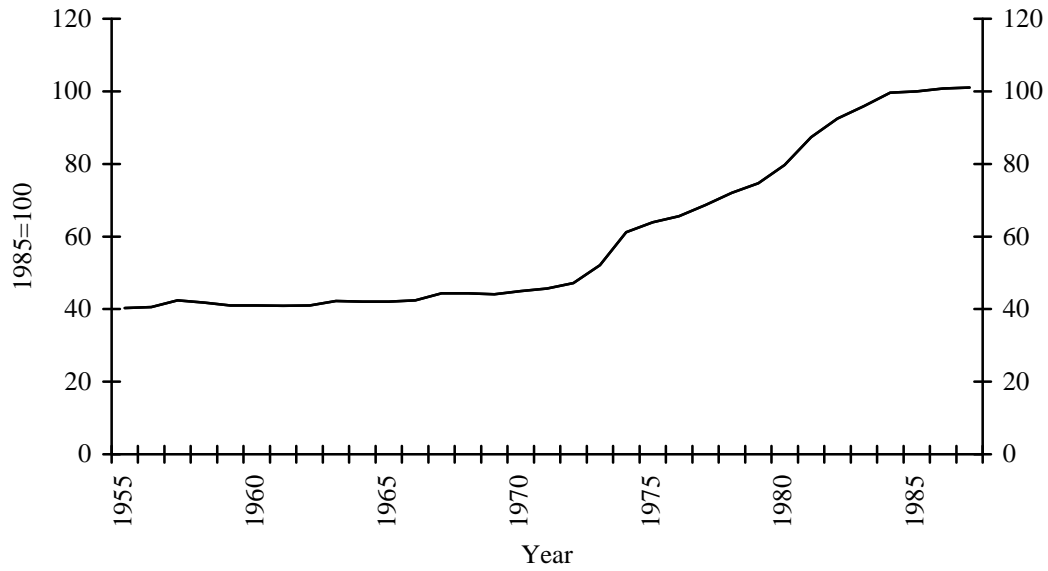


FIGURE 2.2 MALAYSIAN CONSUMER PRICE INDEX
 Source: IMF International Financial Statistics



FIGURE 2.3 TRENDS IN AVERAGE ANNUAL EARNINGS BY RACE IN MFLS2 (1985 RINGGIT)

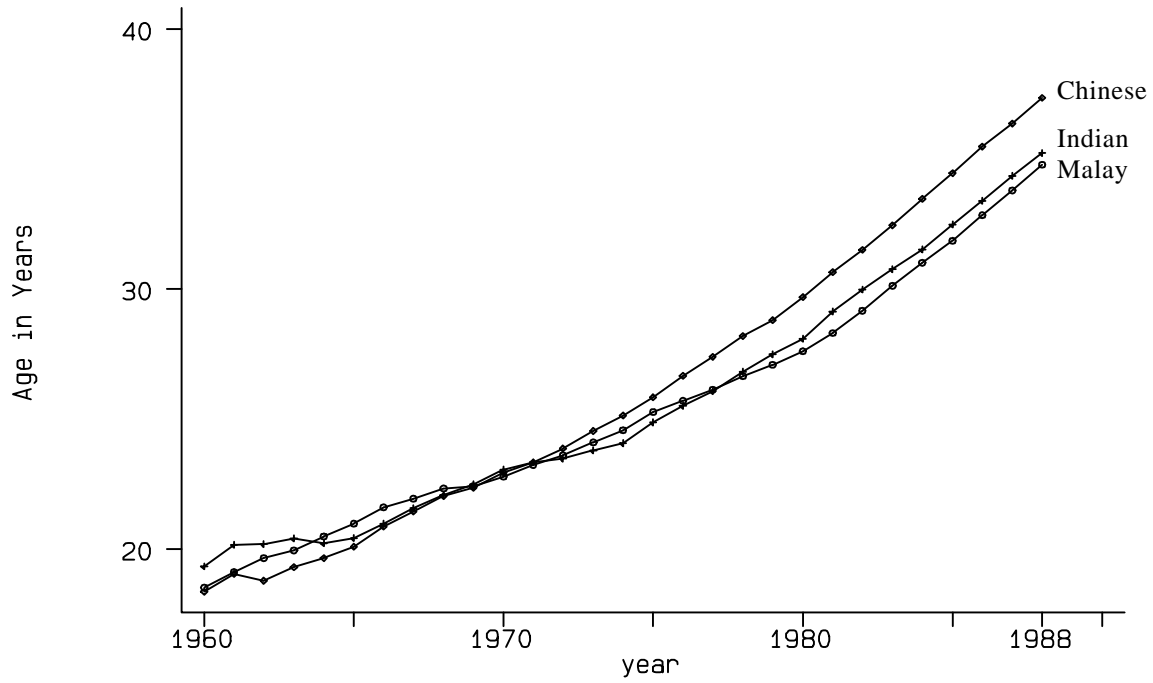


FIGURE 2.4 AGE TRENDS IN MFLS2

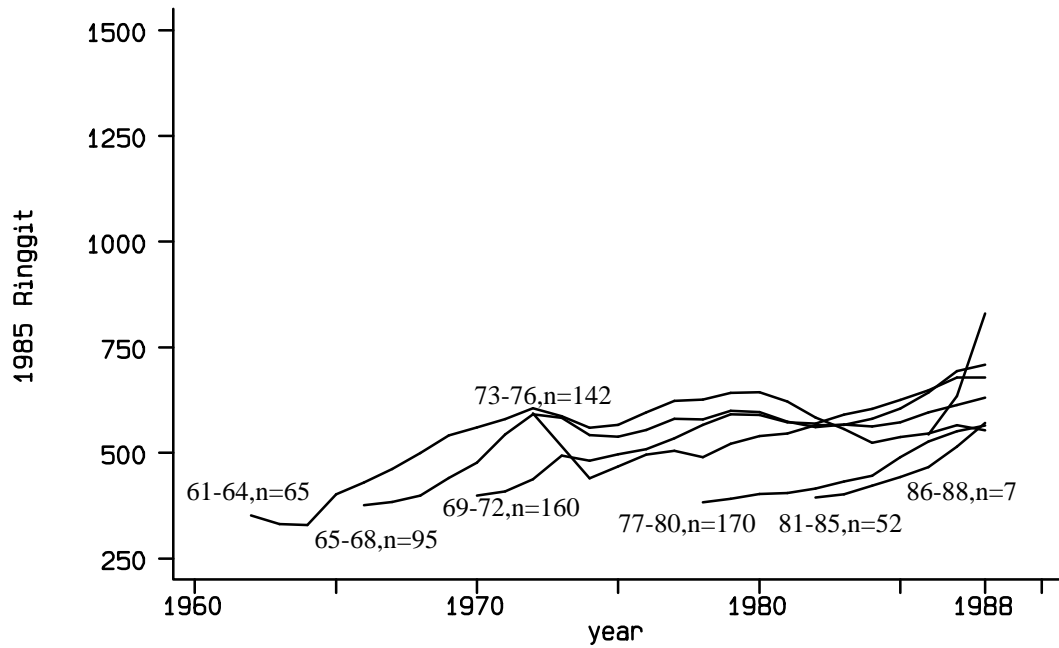


FIGURE 2.5 TRENDS IN AVERAGE COHORT EARNINGS FOR MALAYS

Each cohort is labeled by the year of entry into the labor force and the number of observations

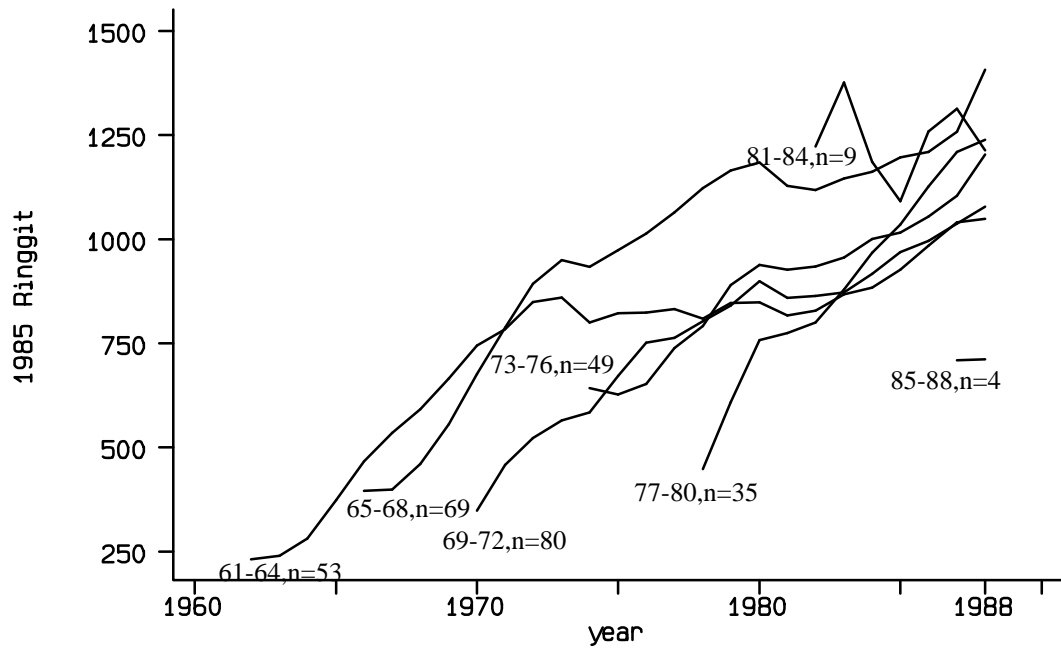


FIGURE 2.6 TRENDS IN AVERAGE COHORT EARNINGS FOR CHINESE

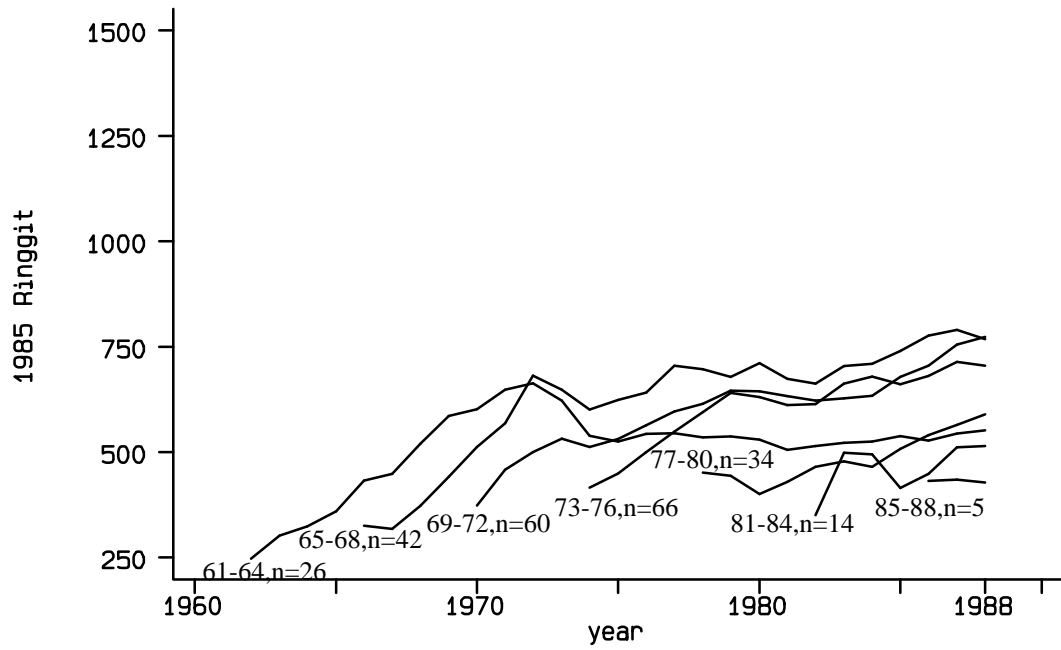


FIGURE 2.7 TRENDS IN AVERAGE COHORT EARNINGS FOR INDIANS

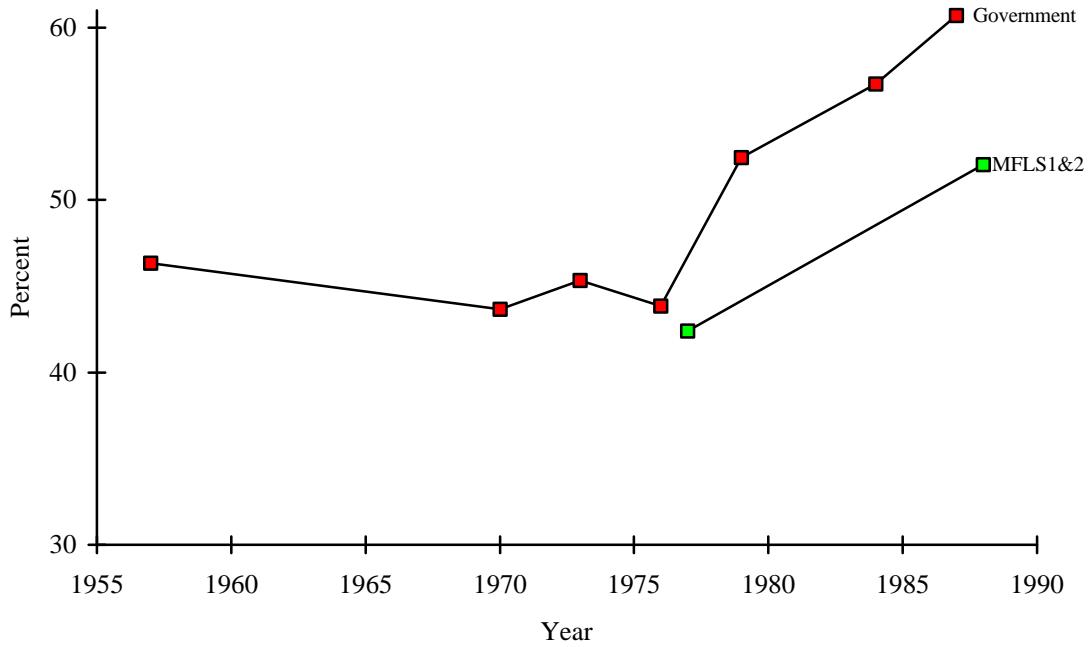


FIGURE 2.8 TREND IN MALAY/CHINESE INCOME FROM MALAYSIAN HOUSEHOLD SURVEYS
 Sources: Wyznan (1990), p.60, Snodgrass (1980), p.83 for 1970 PES, and Lefranc (1992), p.61 for MFLS1&2.

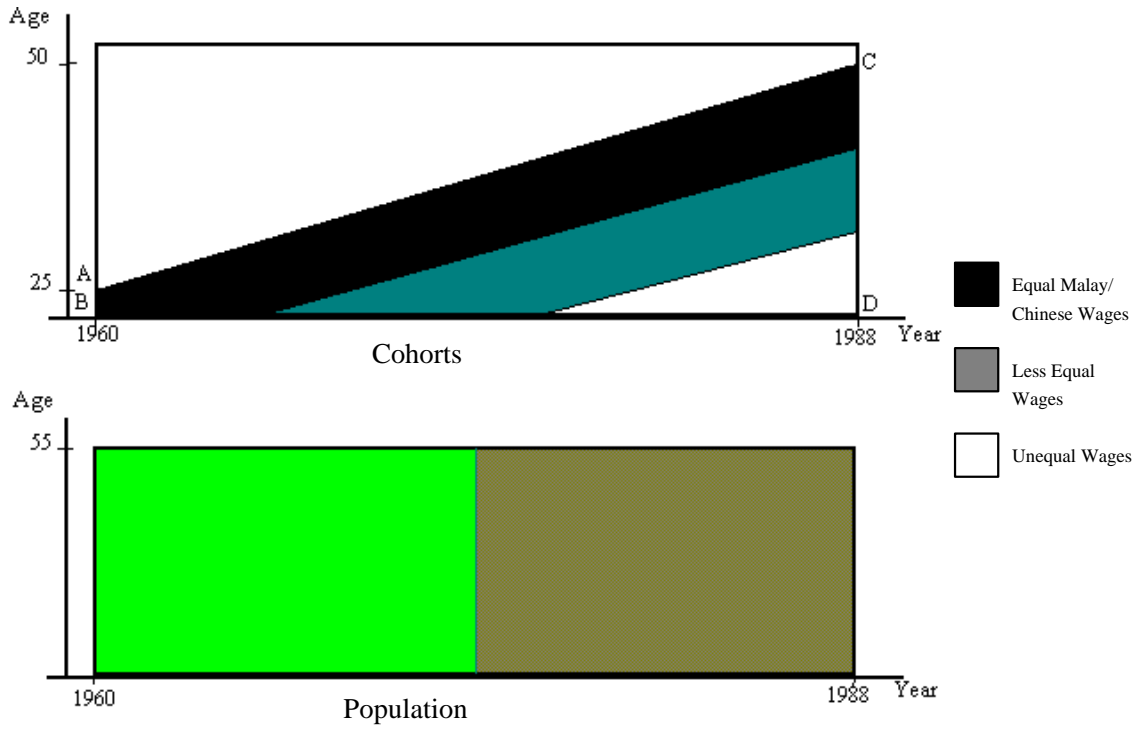


FIGURE 2.9 COMPOSITION OF LABOR HISTORY AND POPULATION IN INCOME TRENDS

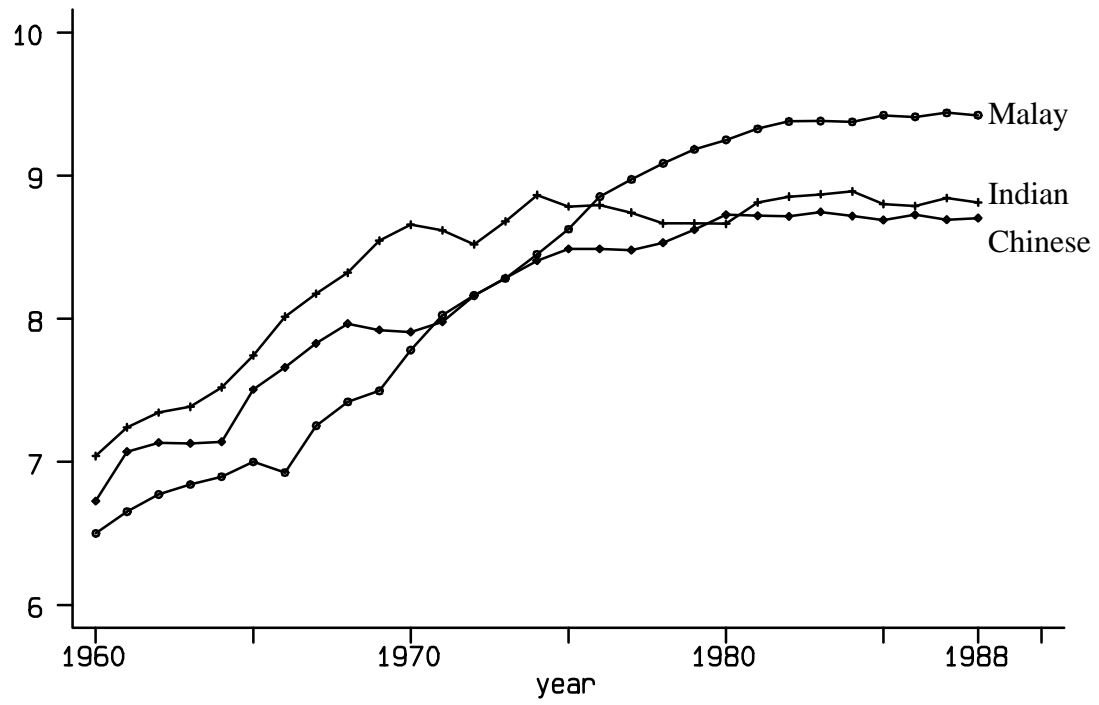
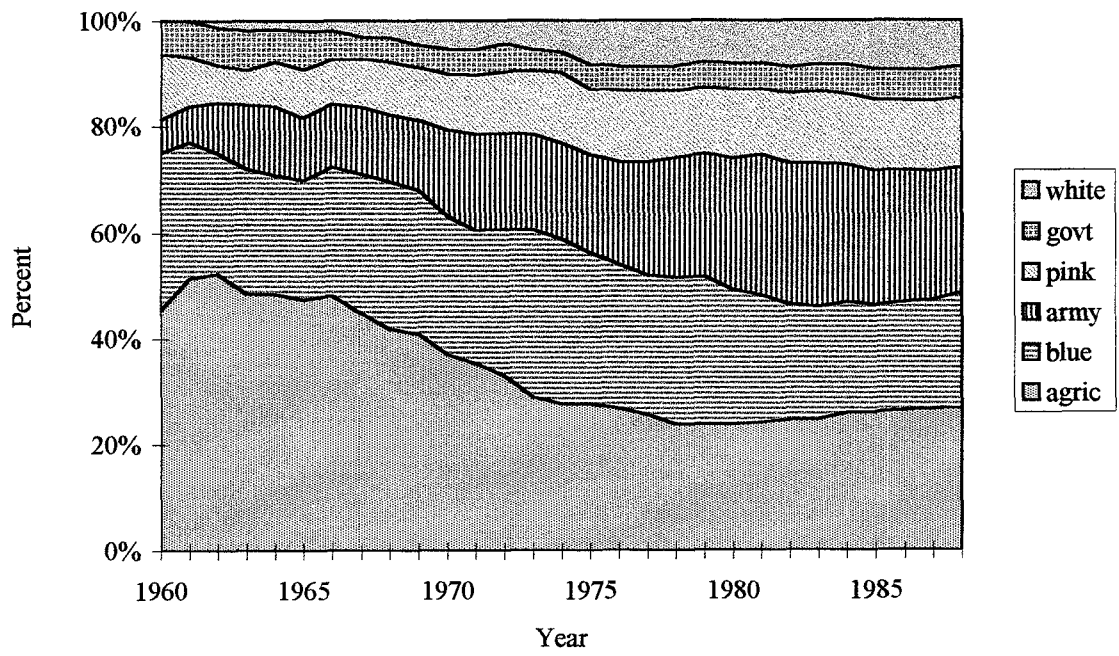
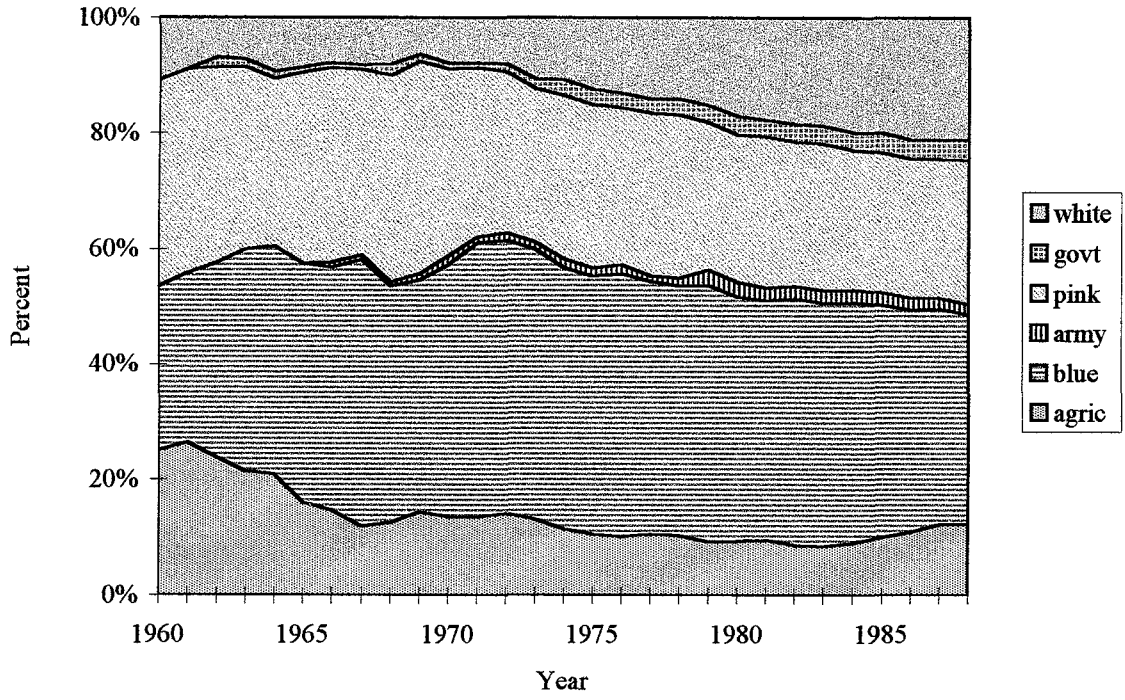


FIGURE 3.1 TRENDS IN AVERAGE EDUCATION BY RACE (YEARS OF EDUCATION)

Occupations of Malays



Occupations of Chinese



Occupations of Indians

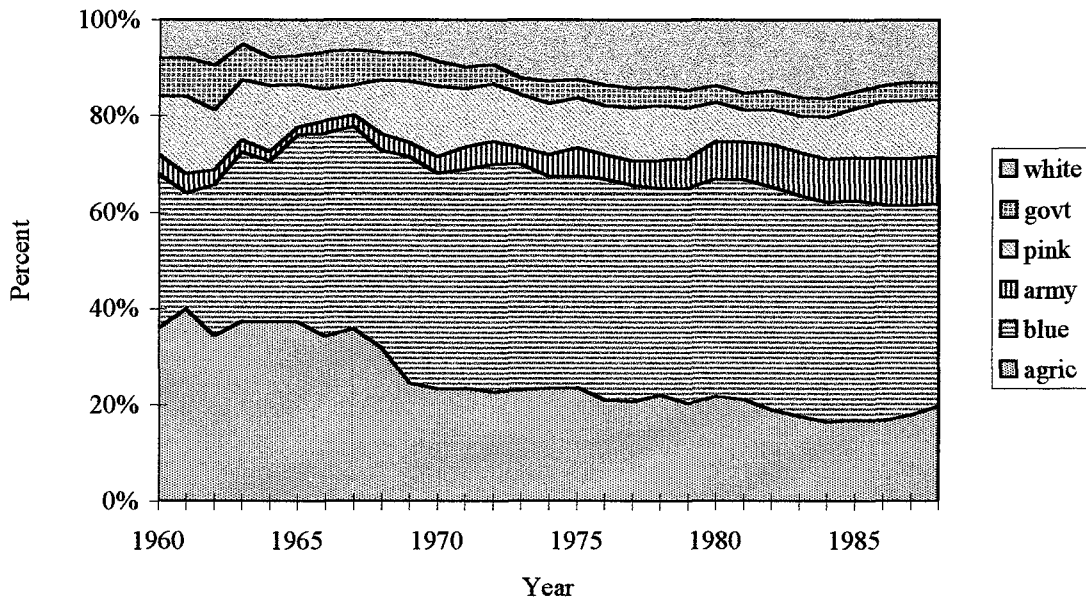


FIGURE 3.2 OCCUPATION SHARES BY RACE

White: "White collar" - managers and professionals

Govt: government employees except Army

Pink: "Pink collar" - sales and clerical workers

Army: military, police and fire persons

Blue: "Blue collar" - manufacturing, trades, transportation and construction workers

Agric: agricultural workers and farmers