LBL-8782 c.2
UC-13

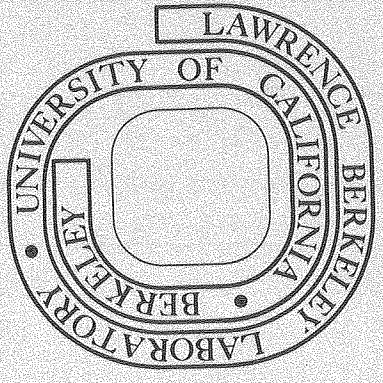
# PROJECTING POPULATION TO 1980

Esther Schroeder

February 1979

Prepared for the U. S. Department of Energy
under Contract W-7405-ENG-48





## DISCLAIMER

LBL-8782

# PROJECTING POPULATION

to 1980

Esther Schroeder
Computer Science and Applied Mathematics
University of California
Berkeley, California, 94720

Phone - (415) 486-5306
FTS 451-5306

February 1979

TABLE OF CONTENTS 21 FEB 79

Over the past few years, the Lawrence Berkeley Laboratory and the U. S. Department of Labor - Employment and Training Administration have developed a projections procedure, the Labor Market Projections Model, (LMPM), to aid local planners. LMPM was designed to project population, labor force, and unemployment at the local level. Both the labor force and unemployment projections rely heavily on the results of the population projections. For more information on LMPM and its background, please refer to Population, Labor Force and Unemployment Projections by Esther Schroeder and Jim Woods, March 1978, LBL-7547.

This past year considerable effort has been put into modifying and improving the population projections. The projections are now calculated for three races (white, black and other), instead of just two (white and nonwhite). Submodels to take into account the special populations - the college students and the military - were included. Net migration totals for 1970-1975 from the Revenue Sharing tapes were used to improve the net migration rates for 1960-1970 previously used. Lastly, as the projections are now being done from 1970 to 1980, the model does this in two steps - projecting first to 1975 and then to 1980. With these changes, the population projections model has been strengthened considerably. As a result, the labor force and unemployment projections are similarly improved.

LMPM has been designed to aid local planners by providing "current" data at the local level. The resultant projections are done for all states, SMSAs and prime sponsors. A prime sponsor is a local jurisdiction of at least 100,000 people with the responsibility of providing job training and employment opportunities for the economically disadvantaged and the unemployed in its area. Prime sponsors cover the United States and vary widely in size and composition from the City of Berkeley, California, to New York City to the State of Vermont, for example. When considering possible improvements to the model, one must first insure that the necessary data exists with which to implement the improvement.

As implemented in the fall of 1976, the population projections are calculated using a refined cohort component method. This involves the separate projections of births, deaths and net migration to yield population by race by sex by five year age cohorts for a particular geoarea. (Footnote 1) Using total population figures for the target year as received from the Bureau of the Census, these cohort projections are then normalized to add up to a control total for that area. For more detail on the actual steps involved, please see the write-up previously cited.

The data sources used to calculate the population projections were

(1) 1970 population by race by sex by age from Tabulation 17 of the 1970 Fourth Count Census,

(2) 1960-1970 net migration rates by race by sex by age cohort at the county level (Footnote 2),

(3) National Survival rates for 1970-1975 by age by sex (Footnote 3),

(4) National Fertility rates by age by race (Footnote 4), and

(5) Projections of Resident Population for the target year for states, counties, MCDs in New England and cities with 50,000+ in 1973 (Footnote 5).

## 3.1 BETTER RACIAL DETAIL

In the past, the projections were calculated only for whites and nonwhite. Now the racial breakdown is white, black and other. This added information can be of great use in some geoareas such as New York City. This detail can be obtained from Tabulation 17 of the 1970 Fourth Count Census. The rest of the calculations follow as before. Since net migration rates and fertility rates are available only for white and nonwhite, the nonwhite rates are now used for both the racial categories - blacks and others.

For geoareas that are comprised of many small pieces (New England areas in particular), disclosure suppression frequently occurs in Tabulation 17. The projections are still calculated for the three races, but a message appears at the bottom of each table where disclosure suppression occurred stating that there is disclosure suppresion in the original 1970 population data from the Census. Also, the projections are then recalculated for just two races - white and nonwhite.

## 3.2 PROJECTING IN TWO STEPS

The population projections from 1970 to 1980 are now being done in two steps. First the projections are done to 1975 with the race, sex and age detail normalized to a control total obtained from the 1975 Revenue Sharing Tapes (Footnote 6). These population figures are then projected to 1980 with the race, sex and age detail normalized to a control total provided by the states or by the Bureau of the Census (Footnote 7).

## 3.3 SPECIAL POPULATIONS

Special populations, such as the military and college, create difficulties for population projections models. Usually, the special populations are comprised of people within a narrow age range and their behavior is quite different from that of the resident population. Those areas that contain a college or military barracks are apt to have an unusually large proportion of younger people. These people usually leave the area once they finish college or leave the military.

To deal with these special populations, it is best to separate them from the resident population of the geoarea in 1970 and project them separately. Ideally, they should be subtracted from the 1970 population before the projections are made and then the correct student and military populations of 1990 added back in at the end. From the 1970 Fourth Count Census, it is possible to get some information on the student populations and those in military barracks in 1970: however, no comparable data exists for 1980. Due to the lack of data, it is being assumed that the special populations have not changed since 1970, i.e., what is subtracted out in 1970 will be added back in, in 1980. However, since these projections are geared towards estimating the civilian labor force and unemployment in the civilian labor force, the military population will not be added back in, in 1980. Thus the 1980 population projections are a projection of the civilian population. As will be shown later, it is not possible just to add back in the students to the correct age cohorts at the end.

## MILITARY

From the 1970 Fourth Count Census, it is possible to get a count of those people in the military. Tabulation 51 contains the number of people living in military barracks by race. There is no sex or age breakdown. In order to get an estimate of the age and sex breakdown, tabulations of several states were run on the 1970 one-in-a hundred Public Use Sample, based on the 15% sample. Those states with a large military population were selected. Please look at Appendix A for the tabulations of people in military barracks in some states, as obtained from the Public Use Sample.

After looking at these tabulations, several facts become apparent.

1) In 1970, the people in military barracks were essentially male.

2) The sample is too small to give an adequate age distribution by race. There are not enough blacks or others to generate a meaningful distribution.

3) There are not as many in the military in age cohort 15-19 as one would have thought. The bulk of the military is in the age cohort 20-24.

These tabulations were used to estimate the percentage of the military in the age cohorts from 15-19 through 40-45. Although the age distribution did vary some from state to state, the following percentages seemed to approximate the "average" age distribution for the military. These percentages were used for all three races - white, black, and other.

| Age Cohort | Percentage |
|---|---|
| (15-19) | 25. |
| (20-24) | 57. |
| (25-29) | 09. |
| (30-34) | 04. |
| (35-39) | 03. |
| (40-44) | 01. |

In each geoarea, the following pattern was used to estimate the race by age breakdown of any men in military barracks.

1) For each race, determine if there are some people in the military.

2) If there is essentially no military, ignore them.

3) If there are some in the military, break them down into an age distribution using the percentages derived from the Public Use Sample.

4) Compare the estimated military population with that of the area. For some small areas, the estimated military conflicts with the actual population. If possible, the program redoes the age breakdown so that it is compatible with the population of the area. If the population of the area is very sparse, the program may not be able to determine a satisfactory age breakdown for the military. If this is the case, the military of the race in question are ignored and the program prints out a message to that effect.

If the program feels that it has successfully estimated the military in a given race, these estimates are subtracted from the corresponding race sex and age cohorts of the population of the geoarea to yield an estimated civilian population in 1970. This civilian population is then projected to 1980 yielding a projected civilian population by race, sex, and age.

## STUDENT POPULATION

Tabulation 39 of the 1970 Fourth Count Census contains a count of students by age cohort and by race. The relevant age cohorts in Tab 39 are (16-17), (18-19), (20-24), and (25-34). I have assumed that 16-17 year olds are in high school and living at home, while the 18-19 year olds are in college and are part of the special student population. However, there is no sex breakdown in this tabulation. In order to get an estimate of the sex breakdown in each age race cohort, tabulations of several states were run on the 1970 one-in-a hundred Public Use Sample, based on the 15% sample. Those states with a large student population were selected. Please look at Appendix A for the tabulations of the student populations in some states, as obtained from the Public Use Sample.

These tabulations were used to estimate the sex breakdown of college students for the relevant age cohorts as well as the breakdown of those student aged (25-34) into two age cohorts, (25-29) and (30-34). Note that the sex breakdown varies considerably with race. The following percentages were used to estimate the sex breakdown for the college population in each geoarea.

| Age Cohort | White | Black | Other |
|---|---|---|---|
| (15-19) | | | |
| Percentage Male | 53. | 48.5 | 58. |
| Percentage Female | 47. | 51.5 | 47. |
| (20-24) | | | |
| Percentage Male | 57.5 | 49. | 58. |
| Percentage Female | 42.5 | 52. | 42. |
| (25-34) | | | |
| % Male (25-29) | 47. | 42. | 40. |
| % Female (25-29) | 20.5 | 24. | 22. |
| % Male (30-34) | 21.5 | 22. | 25. |
| % Female (30-34) | 11. | 12. | 13. |

Essentially, in those areas with students, the above percentages are used to estimate the sex breakdown of the student population. This estimated student population is then compared with the total population of the corresponding race, sex and age cohorts. When the estimated student population conflicts with that of the area, the estimates are modified to be compatible with the population of the area.

The Public Use Sample is used to provide the first estimate of the student population. By comparing this estimate with that of the area, the program can make improvements when necessary.

In greater detail, the actual steps taken, are as follows.

1) For the relevant age cohorts in each race, determine if college students are present, ie., compare the student population of each race age cohort with the total population of the same race age cohort.

2) If the number of students is less than 5% of the population of the corresponding race age cohort, they will be ignored.

3) If the number of students is greater than 5% of the population of the corresponding race age cohort, estimate the sex breakdown by using the proper percentages as derived from the Public Use Sample.

4) Compare the estimated student populations with the total populations for the corresponding race sex and age cohorts. For some geoareas, the estimated breakdown conflicts with the actual population. This sometimes occurs in small areas and in the racial categories black or other. For example, the area in question may have very few black residents but a college has attracted many from outside the area. The program then redoes the sex breakdown of the student population to be compatible with the population of the race sex and age cohorts involved.

5) Subtract the estimated student population from the total population of the relevant race sex and age cohorts. Project this adjusted population to 1975.

   Before normalizing this projected population to the desired control total, the student population must be added back in as the student population is included in the control total. Although the student population is assumed to be of a constant size from 1970 on, it is not possible to add back in, in 1975 what was subtracted out in 1970. The difficulty is due to the fact that some of the students at college in this

geoarea come from outside the area whereas some were already residents of the area. For example, a college student aged 18 in 1975 would have been 13 in 1970 and in high school. If he was in high school in the area in 1970, he is already counted in the (15-19) age cohort in 1975 and should not be added in again. If however, he came from outside the geoarea, he is not yet counted in the (15-19) age cohort in 1975 and should be added in.

By looking at ratios of students in adjoining age cohorts, the program determines what proportion of the students come from outside the area. (For details on how this is done, please see Appendix B.) The estimated "nonlocal" student population is added back into the relevant race age sex cohorts and these figures are then normalized to the desired 1975 control total.

6) The above process from step 4 on is repeated in projecting from 1975 to 1980, ie., the estimated student population is subtracted out, the resultant adjusted population is projected, the estimated "nonlocal" students are added back in, and the resultant population figures are normalized to the desired 1980 control total.

## 3.4 MIGRATION RATES

The most current source of county migration by race sex and age cohort is the 1960-1970 data provided by Gladys Bowles (Footnote 8). However, the 1975 Revenue Sharing tape contains estimates of net migration from 1970 to 1975 by state and by county (Footnote 9). This data is more up-to-date than the Bowles data but is lacking the desired race sex and age breakdown. By using the plus-minus procedure (Footnote 10), the net migration rates by race, sex, and age for 1960-1970 can be adjusted to give the desired total net migration for 1970-1975. The amount of the adjustment is distributed among the items in proportion to their absolute values. This technique does not change the sign of any migration rate but does alter the magnitude to get the desired result. For example, if the net migration desired for 1970-1975 is greater than that obtained by using the rates from 1960-1970, the positive rates will be increased and the negative rates will be decreased. These adjusted net migration rates were used in projecting from 1970 to 1975 and again from 1975 to 1980.

As the 1975 Revenue Sharing tapes contain net migration figures only at the state and county levels, it was not possible to use the above technique for those geoareas that are not integral sums of counties. Furthermore, in some cases where the above modifications was tried, it broke down due to a large change in migration patterns. The plus-minus technique does not work when the net amount of adjustment required exceeds the sum of the absolute values in the distribution. When this occurred, a message was printed out warning the user that migration patterns have changed considerably recently, and that the newer patterns have not been incorporated into the model.

The above outlined modifications have greatly strengthened the population projections model; the projections are now less dependant upon the actual geoarea involved, whether a populous state, a college town, or a rural county. However there is much more work that should be done. The submodels dealing with the military and with the college populations need more work. It may also be possible to introduce a submodel to take into account people living in institutions. The present submodels dealing with the college students and the military could be improved if better and/or more up-to-date sources of data were found.

Work should be done on the fertility rates used. The further the projections get from the base year, the more critical the birth rate becomes. Furthermore, fertility patterns have been changing greatly. After the baby boon in the 1950's and early 1960's, the birth rates declined sharply. They appear to have reached a minimum about 1975 and are now leveling off and even rising slightly. It would be most helpful if more current age specific fertility rates could be introduced at the local level.

A cohort component projections model is very dependant upon the migration rates used. However, not only is it difficult to get data reflecting recent migration patterns, it is almost impossible to predict future migration patterns. Although this past year, migration totals for the years 1970-1975 were introduced where possible, this area needs more work. For those areas that could not be normalized to the 1970-1975 migration figure, it may be beneficial to introduce migration rates dependant upon the type of area involved, urban, suburban, rural, etc.

1) Shyrock, Henry, Siegel, Jacob and Associates, 1973, The Methods and Materials of Demography , U. S. Department of Commerce

2) Bowles, Gladys K., "Net Migration of the Population by Age, Sex and Color", U. S. Department of Agriculture, Economic Research Service, University of Georgia, Institute for Behavorial Research, and National Science Foundation, Research Applied to National Needs, cooperating

3) Current Population Reports, December 1972, "Population Estimates and Projections Series", P-25, No. 493, table B-5

4) Vital Statistics of the United States, 1971, Volume 1 - Natality, Table 1-6, U. S. Department of H. E. W.

5) Gibson, Cam, 1976, "Projections of Resident Population for April 1, 1985 for the Manpower Administration, Bureau of the Census

6) Current Population Reports, April 1977, "Population Estimates and Projections", Series P-25

7) Gibson, ibid

8) Bowles, ibid

9) Current Population Reports, April 1977, ibid

10) Shyrock, ibid, pp705-707

Tabulations were run on the 1970 one-in-a hundred Public Use Sample, based on the 15% sample, for those states with a large student and/or military population. The author used these tabulations to estimate age and sex breakdowns for the student and military populations. This Appendix contains tabulations for several states.

## CALIFORNIA

### Student Population

| Males<br>Age Cohort | Total | White | Black | Other |
|---|---|---|---|---|
| 16-17 | 346500 | 310600 | 23900 | 12000 |
| 18-19 | 212000 | 187400 | 13700 | 10900 |
| 20-24 | 252800 | 225600 | 10200 | 17000 |
| 25-29 | 113700 | 101500 | 7600 | 4600 |
| 30-34 | 58100 | 51100 | 3900 | 3100 |

| Females<br>Age Cohort | Total | White | Black | Other |
|---|---|---|---|---|
| 16-17 | 325100 | 288400 | 24700 | 12000 |
| 18-19 | 178800 | 155900 | 14900 | 8000 |
| 20-24 | 164200 | 140400 | 12200 | 11600 |
| 25-29 | 56300 | 49200 | 4400 | 2700 |
| 30-34 | 36400 | 31500 | 2900 | 2000 |

### Population in Military Barracks

| Males<br>Age Cohort | Total | White | Black | Other |
|---|---|---|---|---|
| 15-19 | 58100 | 53000 | 4300 | 800 |
| 20-24 | 157800 | 144200 | 9400 | 4200 |
| 25-29 | 38600 | 34800 | 2400 | 1400 |
| 30-34 | 31700 | 27300 | 2700 | 1700 |
| 35-39 | 28300 | 24200 | 2200 | 1900 |
| 40-44 | 12900 | 10900 | 1000 | 1000 |
| 45 and over | 11500 | 10800 | 1500 | 200 |

| Females<br>Age Cohort | Total | White | Black | Other |
|---|---|---|---|---|
| 15-19 | 800 | 600 | 200 | 0 |
| 20-24 | 2600 | 2100 | 400 | 100 |
| 25 and over | 2500 | 2100 | 300 | 100 |

# TEXAS

## Student Population

| Males Age Cohort | Total | White | Black | Other |
|---|---|---|---|---|
| 16-17 | 196900 | 168300 | 27400 | 1200 |
| 18-19 | 122800 | 109700 | 12300 | 800 |
| 20-24 | 116200 | 107200 | 7800 | 1200 |
| 25-29 | 36900 | 33100 | 3200 | 600 |
| 30-34 | 16300 | 15300 | 900 | 100 |

| Females Age Cohort | Total | White | Black | Other |
|---|---|---|---|---|
| 16-17 | 178200 | 155000 | 22400 | 800 |
| 18-19 | 101900 | 89200 | 12200 | 500 |
| 20-24 | 71300 | 63000 | 7900 | 400 |
| 25-29 | 12200 | 10100 | 2000 | 100 |
| 30-34 | 9500 | 8100 | 1400 | 0 |

## Population in Military Barracks

| Males Age Cohort | Total | White | Black | Other |
|---|---|---|---|---|
| 15-19 | 18700 | 16900 | 1600 | 200 |
| 20-24 | 47700 | 41200 | 5900 | 600 |
| 25-29 | 5700 | 4600 | 1000 | 100 |
| 30-34 | 1300 | 900 | 400 | 0 |
| 35-39 | 1100 | 600 | 500 | 0 |
| 40-44 | 600 | 300 | 300 | 0 |
| 45 and over | 300 | 300 | 0 | 0 |

| Females Age Cohort | Total | White | Black | Other |
|---|---|---|---|---|
| 15-19 | 900 | 700 | 200 | 0 |
| 20-24 | 1100 | 900 | 200 | 0 |
| 25 and over | 100 | 100 | 0 | 0 |

# ILLINOIS

## Student Population

| Males Age Cohort | Total | White | Black | Other |
|---|---|---|---|---|
| 16-17 | 192500 | 166500 | 25100 | 900 |
| 18-19 | 113300 | 100000 | 12600 | 700 |
| 20-24 | 110500 | 102800 | 6600 | 1100 |
| 25-29 | 37900 | 34800 | 1800 | 1300 |
| 30-34 | 16000 | 13700 | 2100 | 200 |

| Females Age Cohort | Total | White | Black | Other |
|---|---|---|---|---|
| 16-17 | 182200 | 155200 | 26200 | 800 |
| 18-19 | 93500 | 80900 | 12100 | 500 |
| 20-24 | 78000 | 70400 | 6700 | 900 |
| 25-29 | 17400 | 15200 | 1900 | 300 |
| 30-34 | 11700 | 10100 | 1300 | 300 |

## Population in Military Barracks

| Males Age Cohort | Total | White | Black | Other |
|---|---|---|---|---|
| 15-19 | 15600 | 14500 | 900 | 200 |
| 20-24 | 12700 | 11300 | 1300 | 100 |
| 25-29 | 800 | 700 | 100 | 0 |
| 30-34 | 700 | 600 | 100 | 0 |
| 35-39 | 500 | 400 | 0 | 100 |
| 40-44 | 200 | 200 | 0 | 0 |
| 45 and over | 200 | 200 | 0 | 0 |

| Females Age Cohort | Total | White | Black | Other |
|---|---|---|---|---|
| 15-19 | 500 | 500 | 0 | 0 |
| 20-24 | 300 | 300 | 0 | 0 |
| 25 and over | 0 | 0 | 0 | 0 |

# HAWAII

## Student Population

| Males Age Cohort | Total | White | Black | Other |
|---|---|---|---|---|
| 16-17 | 12700 | 5000 | 100 | 7600 |
| 18-19 | 7000 | 2700 | 0 | 4300 |
| 20-24 | 8200 | 3000 | 100 | 5100 |
| 25-29 | 3300 | 1800 | 0 | 1500 |
| 30-34 | 1900 | 900 | 0 | 1000 |

| Females Age Cohort | Total | White | Black | Other |
|---|---|---|---|---|
| 16-17 | 13200 | 5000 | 0 | 8200 |
| 18-19 | 7300 | 2500 | 0 | 4800 |
| 20-24 | 8700 | 2100 | 0 | 6600 |
| 25-29 | 1500 | 800 | 0 | 700 |
| 30-34 | 1000 | 300 | 0 | 700 |

## Population in Military Barracks

| Males Age Cohort | Total | White | Black | Other |
|---|---|---|---|---|
| 15-19 | 3400 | 2800 | 600 | 0 |
| 20-24 | 16700 | 14900 | 1500 | 300 |
| 25-29 | 2000 | 1800 | 0 | 200 |
| 30-34 | 1000 | 500 | 100 | 400 |
| 35-39 | 500 | 300 | 0 | 200 |
| 40-44 | 600 | 500 | 100 | 0 |
| 45 and over | 300 | 300 | 0 | 0 |

| Females Age Cohort | Total | White | Black | Other |
|---|---|---|---|---|
| 15-19 | 0 | 0 | 0 | 0 |
| 20-24 | 400 | 400 | 0 | 0 |
| 25 and over | 0 | 0 | 0 | 0 |

In general, the behavior of the student population is quite different from that of the resident population. To allow for this, a student submodel was introduced into the population projections model. Before any projections are made, the student population is subtracted out. The nonstudent population is then projected using the cohort-component method. Before normalizing these projected figures, the student population is added back in. Adding the student population back in turns out to be quite difficult. The problem arises in trying to determine whether the schools within the area have attracted many students from outside the area, or whether these schools are mainly being attended by "local" students. For example, a college student aged 20 in 1975 would have been 15 in 1970 and in high school. If he was in high school in the area in 1970, he is already counted in the (20-24) age cohort in 1975 and should not be added in. If however, he came from outside the geoarea, he is not yet counted in the (20-24) age cohort in 1975 and should be added in. This Appendix will elaborate on the procedure that was used to estimate the "nonlocal" student population in each race sex age cohort.

The procedure followed was the same for each race and sex but varied among the age cohorts. For the age cohorts (25-29) and (30-34), whatever was subtracted out in the beginning was added back in. This was due to the facts that the student population was assumed to be constant and that anyone between the ages of 25 and 34 in 1975 would have been between the ages of 20 and 29 in 1970. Students of those ages were subtracted out at the beginning.

The procedure followed for age cohorts (15-19) and (20-24) is more difficult as these age cohorts include some people who would have been in high school at the beginning of the projections. First consider the age cohort (15-19). The possible college students are those aged 18 or 19. To get a measure of how many college student there are in relation to high school students, look at the proportion

$$F = \text{Students (18-19)} \; / \; \text{Students (16-17)}. \qquad (1)$$

If F is greater than 1, it is an indication that the area attracts many college students from outside the area. Then calculate

$$F^* = F - .6, \qquad (2)$$

but do not let it be less than 0. or greater than 1. Then the "nonlocal" students S may be calculated as

$$S = F^* * \text{Students (18-19)}. \qquad (3)$$

This process assures that if F is 1.6 or greater, all students aged (18-19) will be added back in; if F is less than .6 no students will be added back in, and in between a linear percentage will be added back in.

For the age cohort (20-24), a similar type of procedure is followed although different functions are used. First calculate the ratio

$$G = \text{Students (20-24)} \; / \; \text{Students (18-19)}. \qquad (4)$$

G tries to measure the presence of graduate students. If G is very large, there are probably graduate schools attracting people from outside the area. If G is very small, there may be just a junior college in the area and thus few students in the age cohort (20-24). Any students age 23 or 24 would have been 18 or 19 in 1970 and would have been subtracted out. These students should be added back in as well as any 20 to 22 who may have been attracted to the area from outside. Calculate the number of students to add back in, T, as

$$T = \text{minimum}( 1., (.5*G+.2) * \text{Students(20-24)} ). \qquad (5)$$

Thus at least .2*Students(20-24) are added back in and if G is 1.6 or larger, all the Students (20-24) are added back in. For G between .6 and 1.6, a linear percentage of students is added back in.

This report was done with support from the Department of Energy. Any conclusions or opinions expressed in this report represent solely those of the author(s) and not necessarily those of The Regents of the University of California, the Lawrence Berkeley Laboratory or the Department of Energy.

*TECHNICAL INFORMATION DEPARTMENT*
*LAWRENCE BERKELEY LABORATORY*
*UNIVERSITY OF CALIFORNIA*
*BERKELEY, CALIFORNIA 94720*