HOW CROWDED WILL WE BECOME?

Nathan Keyfitz *University of California, Berkeley*

ALL STATISTICAL facts concern the past. The Census of April 1970 counted 205 million of us, but we did not know this until November, despite the census emphasis on speed, pursued with ingenuity and with much new electronic equipment. Stock-market prices and volumes are hours old before they appear in the evening paper. Statistics of plans or intentions are only an apparent exception. No one can ever gather data directly on the future.

Yet the actions that statistics serve to guide can occur only in the future. The local telephone company wants to know how much this town will grow in population over the next few decades. Its interest is not abstract curiosity, but contemplated construction of new lines out toward a certain suburb. The investment might occur in the next two or three years, and the service given by the investment along with the income derived from it would be spread over 30 years. If the town does not grow as much as expected, the construction would be wasteful. If the growth is in the direction of a different suburb, then lines will be idle on one side of the town and too often busy on the other side. School authorities, the bus company, a textile manufacturer, all similarly need statistics on the future for the conduct of their business, and these are nowhere to be collected until the future has become past and it is too late.

With producers of population statistics all working on the near side of *now* and users all concerned with the far side, it is lucky that even in times of rapid change, some continuities are to be found between past and future. Population projection rests on these continuities.

The continuities are not be found in simple totals. We know that the number of people in the U.S. does not increase evenly from year to year, and still less does the population of one town or one age group increase evenly. The age classes especially have fluctuated erratically in recent decades. Today the U.S. includes an exceptionally large proportion of young people 10 to 25 years of age, the result of the baby boom of the forties and fifties. They have crowded the high schools and colleges, and they are seeking jobs and entry into graduate schools across the country. But during the sixties, births fell sharply, and the number of pupils entering elementary schools leveled off.

Yet we can say something about the future. At the end of the seventies, schools and the labor market will be reached by the wave of what may be called the nonbirths of the sixties. But, though kindergartens and public schools will slow their expansion in the seventies, they may have to accelerate it again in the eighties to accommodate a new generation—children of the children born in the postwar baby boom. How such things can be projected with some confidence is our subject.

The approach, or model, that we shall build for projection serves other purposes than prediction. It is especially valuable for judging the effects on population growth of a possible change or a proposed policy.

PROJECTION WITH CONSTANT BIRTH AND DEATH RATES

The trick in projection is to seek elements that remain nearly constant through time. The increase in total population from year to year plainly does not qualify, but certain *rates* do remain more or less the same, and on these we rest our analysis of the future. For example, the proportion of people aged 30 who die each year is likely to remain much the same in 1960, 1970, and 1980. These death rates are constant enough that some fairly reliable predictions can be hung on them, and we proceed to the exploitation of this constancy.

Our projection of population into the future includes three parts:

(1) The statistical data of a baseline census from which work starts
(2) Effect of death
(3) Effect of birth

Demographers ordinarily recognize five-year age groups, to the end of life, for men and women separately, and they have a computer do the arithmetic. To show the procedure without being swamped in numbers, we consider here girls and women only, and these just up to age 45. Moreover, we need only consider three age groups, each of 15 years' width. For purposes of this illustration, three numbers describe the population at any one time.

We can make a fairly complete analysis for these three groups, and show the whole worksheet. The census of April 1, 1960, counted 27.4 million girls under 15 in the U.S. It showed only 17.7 million between 15 and 29 years. An intermediate number, 18.4 million, were between 30 and 44. (This article follows the census in always counting people at their age last birthday.) Those under 15, born between 1945 and 1960, constitute the baby boom; the next older group, born between 1930 and 1945, are survivors of the meager crop of depression babies; the oldest, aged 30 to 44, were born between 1915 and 1930, when birth rates in the U.S. were higher than in the thirties, but lower than in the fifties.

Now these three numbers can be written one below another in an array known as an age distribution; see Table 1.

So much for the counts made in 1960, our point of takeoff into the future. We now need to know how death and birth will act on this starting distribution. (Migration, which demographers usually take into account in making projections, is probably going to be relatively small and not likely to affect our conclusions seriously, so we shall ignore it.)

Let us start with death, but look at its positive side: the people who do not die, but survive into the next period. The question is, how many of the 27.4 million girls under 15 years of age counted in the 1960 census may be expected to survive to 1975? We have at hand a *life table,* as such collections of survival probabilities are called, that gives the proportion of girls under 15 who survive for 15 years as approximately 0.9924. This life table was calculated from deaths in the U.S. in 1965, and it would not be very different if calculated for any other recent year. Hence the expected number of survivors 15 years later of the 27.4 million counted in 1960 would be

TABLE 1. Age Distribution of American Girls and Women, 1960

AGE	MILLIONS OF GIRLS AND WOMEN
0–14	27.4
15–29	17.7
30–44	18.4

27.4 multiplied by 0.9924, or 27.2 million. These girls would be 15 to 29 years of age in 1975.

No such multiplication can give the *exact* numbers in 1975. Individuals survive or die at random, and even if 0.9924 were the probability for each separate girl 0–14 years of age, a few more or a few less than 27.4 × 0.9924 million could survive in the particular years 1960–75. If the U.S. were subject to serious epidemics, chance events each affecting large numbers of people, then the variation from year to year would be substantial. Because, in fact, death and survivorship act like events affecting each of us more or less independently, the multiplication is permissible, though even then the result could be made wrong by a war or epidemic on the one hand or a medical breakthrough on the other. We shall suppose that the chance of survival does not change very greatly over the period of the projection.

In the same way the proportion surviving 15 years among girls 15 to 29 in 1960 is estimated at 0.9826, and hence the projected number aged 30 to 44 in 1975 would be 17.7 × 0.9826 = 17.4 million. The projections to this point stand as shown in Table 2. Our next task is to fill the upper cell on the right, which requires an estimate of the number under 15 in 1975. (Remember that, to keep things manageable and simple, we are neglecting women 45 or more—of course, only for present simplicity, as the wives of some of us will remind us.)

All of the girls under 15 years of age in 1975 will have been born since 1960, and we need to estimate not how many girl births take place in the 15 years, but how many of these births survive to 1975. We know, also from the 1965 experience, that, on the average a woman 15 to 29 can expect 0.8498 surviving girl babies by the end of a 15-year period. We have counted girl babies only for this purpose because a female model is what we are constructing, and we have deducted deaths among the babies so as to come up with girls under 15 who will be alive in 1975. There were 17.7 million women aged 15 to 29 in 1960, and their contribution to the total girls under 15 in 1975 is expected to be 17.7 × 0.8498 = 15.0 million.

TABLE 2. Projected 1975 Population of American Girls and Women

AGE	MILLIONS OF GIRLS AND WOMEN	
	1960	1975
0–14	27.4	?
15–29	17.7	27.2
30–44	18.4	17.4

Children will be born also to the women 30 to 44 years of age; on the average, these women will have 0.1273 girl babies alive at the end of the 15-year period. The contribution that these make to the total girls under 15 in 1975 is expected to be 18.4 × 0.1273 = 2.4 million. (The actual calculation was made to more decimals than shown here.)

Finally, children will be born before 1975 to girls under 15 in 1960, a large proportion of whom will become of childbearing age during the 15 years. On the average (again at 1965 rates), they will have 0.4271 surviving girls. This average, like the others above, is taken over many different cases; it includes the girls too young to become mothers, those who will be old enough but not yet married, and those who will marry but not have children. The expected contribution here is 27.4 × 0.4271 = 11.7 million.

To find the total number of girl children under 15 surviving in 1975 we must add the numbers reached in the three preceding paragraphs: 11.7 + 15.0 + 2.4 = 29.1 million in all. Figure 1 shows schematically what is happening. (Because so few children are born to women over 44, we can afford to ignore them. Our simple model will give almost the same rate of increase of the population as more elaborate models.)

By repeating exactly the same argument, except that we now start with the 1975 projected population, we obtain the age distribution in 1990; any

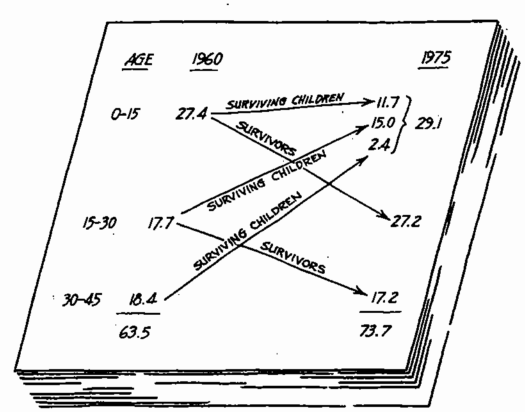

FIGURE 1
Calculation of 1975 population of girls and women under 45 years of age (figures are in millions)

TABLE 3. Millions of Girls and Women Under 45 Years of Age in the U.S. if Birth and Death Rates Remain at the 1965 Level

AGE	1960	1975	1990	2005	2020	2035	2050	2065
0–14	27.4	29.1	37.7	44.1	54.3	65.0	79.0	95.3
15–29	17.7	27.2	28.9	37.5	43.7	53.8	64.5	78.4
30–44	18.4	17.4	26.7	28.4	36.8	43.0	52.9	63.4
Total	63.5	73.7	93.3	110.0	134.8	161.8	196.4	237.1

number of additional 15-year cycles may be calculated similarly. Table 3 shows the resulting numbers up to 2065.

WAVES OF MOTHERHOOD

The first age group, girls under 15, increases less than two million between 1960 and 1975, while the women 15 to 29 increase by almost 10 million. The 15 to 29 group in 1975 are the babies born between 1945 and 1960, the postwar baby boom, and as these succeed the depression babies in any group we expect its number to rise rapidly. Women 30 to 44 actually become fewer during this first 15-year period, even though the 0 to 44 population as a whole is growing.

Because most children are born to mothers 15 to 29 years of age, we can expect a new baby boom, an echo of the first one, at the time when the babies of the fifties themselves pass through childbearing age, and indeed the under-15s grow by 8.6 million from 1975 to 1990 according to Table 3.

In fact, the depression and boom will keep echoing to much later times, supposing, as we do throughout, that childbearing practices remain fixed. But the table also shows that as time goes on the irregularity of the 1960 age distribution steadily lessens. At the end of 105 years all ages are increasing at very nearly the same rate.

That the several ages ultimately increase at the same rate can be seen by dividing each 2065 figure in Table 3 by the corresponding 2050 figure. In Table 4, this ratio is shown to be about 1.2 for the three age groups and the total. By carrying the projection further, we could have had these ratios as close to one another as we wanted; in fact, further calculation shows that they all would converge to 1.2093.

This ratio may be called intrinsic, or the true ratio of natural increase. It can be shown to depend not at all on the 1960 age distribution with which the process started, but only on the rates of birth and death, and it is the most informative single summary measure of that set of rates. It tells us that any population that is subject to our particular birth and death rates

TABLE 4. Increase of Age Groups of Girls and Women in the U.S. from 2050 to 2065

AGE	2050 (MILLIONS)	2065 (MILLIONS)	RATIO, 2065 TO 2050
0–14	79.0	95.3	1.206
15–19	64.5	78.4	1.216
30–44	52.9	63.4	1.198
Total	196.4	237.1	1.207

over a period of time will sooner or later settle down to an increase in the ratio 1.2093, which is to say by about 21% per 15-year period. Under the operation of the projection, applying the assumptions we have made, ***a stable age distribution*** is sooner or later attained in which all the irregularities of 1960 due to boom and depression have been forgotten. Age distributions tend to forget their past when persistently pushed forward by the method developed above.

Let us find numerically the component of population growth that increases in the same ratio in every cycle, a mode of increase spoken of as *geometric*. If we divide each of the numbers shown under the year 2065 in Table 3 by 1.2093, we get back to an estimate for 2050; if we then divide again by 1.2093 we get back to 2035, and so on. To get back to 1960 we would divide by the seventh power of 1.2093, written $(1.2093)^7$ and equal to 3.78. Carrying out the division gives 95.3/3.78 or 25.2 million for age 0 to 14, and similar calculations for the other ages provide what we may call the stable equivalent for 1960; see Table 5.

Table 5 shows the set of numbers that, increasing in the constant ratio 1.2093, would sooner or later exactly join the track of our projection in each age group. If we multiply the stable equivalent by the fixed number 1.2093 to obtain the geometric track, and subtract this from the projection of Table

TABLE 5. Main Component of Female Population in the U.S., 1960

AGE	STABLE EQUIVALENT (MILLIONS)
0–14	25.2
15–29	20.7
30–44	16.8
Total	62.7

TABLE 6. Departures of Projected Population in Table 3 from Geometric Progression in Millions

AGE	1960	1975	1990	2005
0–14	2.2	−1.4	0.9	−0.6
15–29	−3.0	2.2	−1.4	0.9
30–44	1.6	−2.9	2.1	−1.4

3, we obtain Table 6. For example, for girls 0 to 14 in 1975, we have $29.1 - 25.2 \times 1.2093 = -1.4$. Our analysis has separated the prospective population change into two parts, one a smooth geometric increase, the other a series of waves that are departures from the geometric.

These departures gradually diminish in amplitude. For 1960, we have 2.2 million as a measure of the temporary "excess" of the 1945–60 babies. The −3.0 million are the deficiency of the depression babies, and 1.6 million, again an excess, relate to the twenties. Note that by 1990, each of these has an echo, of the same sign but on the whole of smaller amount.

The tendency of the waves to diminish in amplitude is related to women having their children over a range of ages. If all children were born to mothers of the same age, the waves would steadily *increase* in amplitude. With such concentration any irregularity in the age distribution caused, for example, by a war or depression would not only continue echoing through all later generations, but become magnified. In the U.S. today, women prefer to have their children around age 25, whereas our grandmothers spread theirs from about 20 to 45. The new style, associated with the effective use of birth control, could mean diminished stability.

In this analysis of the U.S. population we have gone from the facts of the 1960 census, through various more or less realistic calculations concerning 1975 and even 1990, into a kind of fantasy as we proceed far into the future. The early part of the projection can within limits be useful for practical purposes; the later part is so dependent on various *if's* that one would be very foolish to count on it. The biggest doubt attaches to the birth rate. It may seem that birth is as individual a matter as death, and therefore births across the country ought to be independent of one another, yet in fact high and low birth rates spread like epidemics across the country.

Why then do we bother with the fantasy of such far-out projections referring to the distant future? The answer is that they help us understand the present. We ascertain the meaning of the present rates of birth and death by calculating what they *would* lead to if they continued for a hundred years or more. Let us see why this is even more important in study of the

birth and death rates of developing countries than of a developed one like the U.S.

GROWTH OF DEVELOPING COUNTRIES

The task is in some ways easier for developing countries because they do not have a history of changing birth rates. It is true that their death rates have been falling, and where this occurs for young children only, it is the equivalent of a rise in the birth rate: as far as the population mathematics is concerned, a fall in infant deaths has the same effect as a rise in births. In fact, however, deaths have been falling at nearly all ages, and births are relatively unchanged. The fact is that age distributions are already more or less in the condition we called stable, and which could be attained by the U.S. only in the course of several generations of fixed rates. Because of past uniformly high birth rates, developing countries tend to grow much faster than the U.S. Moreover, they show a simple geometric increase, with all ages rising uniformly. The sort of waves that we have been studying do not occur for them.

Let us concentrate then on the geometric component, and take Malaysia as an example. In the mid-sixties, Malaysia was growing in the intrinsic ratio defined above of 1.59 per 15 years. This corresponds to an annual rate of increase of the 15th root of 1.59 or 1.031, that is, about 3.1% per year, against about 1% for the U.S. To convince ourselves of this we could multiply 1.031 by itself 15 times, that is calculate $(1.031)^{15}$, and we would find the result to be just under 1.59.

We can see the long-term prospect more clearly by translating into doubling times. How long does it take a country that is increasing at the rate r % per year to double in population? The equation to be solved for the unknown time t is $[1 + (r/100)]^t = 2$. The solution is obtained by taking logarithms of both sides and comes out very near to $t = 70/r$, where r is expressed as percent increase. This rule applies to money lent out at interest, and financiers use it because they are very interested in doubling times. The same sort of rule works for the half-life of a piece of radium or other substance under radioactive decay.

As an example of a geometric projection of a population, suppose that Malaysia's rate of 3.1% per year were to go on for about $70/3.1 = 23$ years. This would carry it from the present 10 up to 20 million people. Suppose it went on for another 23 years; this would mean another doubling. At the end of 115 years at this rate, the population would have doubled five times, which means multiplying by 2^5 or 32; Malaysia would contain 320 million persons. By the end of 230 years, it would have doubled ten times and would contain 2^{10} times as many as now, or 1024 times ten million.

No one could mistake such calculations for predictions of what will happen. In a sense they are the opposite; we might call them counterpredictions, for they show that in much less than 100 years, the birth rate will go down or the death rate will go up or both. Most demographers are optimistic enough to believe that the adjustment will be through the birth rate.

Other countries are today growing faster than Malaysia. Mexico's present 50 million population is increasing at about 3.5% per year, which, by our rule, would give it a doubling time of 70/3.5 = 20 years. At this rate, it would be 100 million by the year 1990, 200 million by the year 2010, and 400 million by the year 2030. This again is a counterprediction; shortage of food, excess of pollution, and many other reasons would prevent it from coming true. The usefulness of the calculation is in showing that births must be reduced; anyone who makes a *principle* of permanent opposition to birth control in effect favors an increase of the death rate sooner or later. The most that opponents of birth control can argue is that it should be delayed a few years.

In diluted form, the same is true of the U.S. Our calculation showed that the geometric component, neglecting waves, was an increase of 21% for 15 years, or about 1.2% per year, according to births and deaths of 1965. That means a doubling in 60 years, quadrupling in 120 years, and so on. Contract the U.S. time scale by about three, and the future growth of the U.S. is the same as that of Mexico. And even our having three years to Mexico's one is partly offset by the greater damage to the environment caused by our more advanced industry.

RELIABILITY OF PREDICTION

The techniques presented in this article and obvious extensions of them are much used for predicting the future. They are used not because they are perfect, but because nothing better is available. Whatever continuities exist in birth and death rates are exploited by the makers of projections. From about 1870 to 1935 in Western Europe and the U.S., the birth rate and the death rate were both falling; projections could be made by the method outlined here, except that instead of using fixed rates, the past downward trend in birth and death was projected into the future. Such projections were acceptably accurate as long as the downward trend continued.

But these same countries reached a turning point in the forties. People married younger, and births rose rapidly. Moreover, couples varied the timing of their children as well as varying the total number. The fact that in a modern society couples plan their children, both in number and in timing, can be used to strengthen the predictions. Samples of young couples are surveyed to find what their childbearing intentions are, just as we ask intentions on buying houses and automobiles. The official estimates of the U.S. Bureau of the Census take account of these intentions.

The Census Bureau's projections, which use a vastly elaborated form of the method of this article, can be compared with ours. Theirs are more detailed than ours above have been, and they are also cautious enough to make a variety of projections rather than betting on just one. They end up with four numbers for each age, sex, and future year. For example, for 1990, their four numbers for girls 0 to 14 years of age range from a low of 29.5 million to a high of 41.8 million. Our Table 3 shows 37.7 million.

How well would past application of our model have foretold the 1970 population of the U.S.? If we had worked forward from the 1920 census total of 106 million, using exactly the procedure of this article, but applying it in five-year age intervals to all ages and to both sexes, we would have found about 185 million for 1970. If we had allowed for immigration less emigration of 200,000 per year, this would have brought us to 195 million against the 205 million actually counted. Something a little lower would have been found starting from 1950; starting from 1960, we would have slightly overestimated the census figure. An error of about 5% in estimates made up to 50 years ago is not bad, considering that the Bureau of the Census estimates its own actual count to be subject to nearly 2% error.

We would have done much worse starting in 1940, however; the method of this essay, plus about 6 million immigrants, would have produced a total of only about 160 million. Put another way that sounds even worse: the increase from 1940 to 1970 was about 62 million, and of it, we would have estimated less than 30 million. This is not a good score. The baby boom of the fifties was a historic event about as hard to predict in advance as the war that sparked it.

MODELS PERMIT EXPERIMENTS

We have discussed the population projection as a way of making predictions, and also as a way of making counterpredictions—calculating what would happen if present rates continued as a way of showing that they cannot continue. This last suggests what may be the most important use of the model of this essay, originally developed and still often applied to making predictions. This use is experimentation. Not only does our model answer the question "What would happen if the birth and death rates of the present time continue into the future?" but it answers a great variety of other important questions. What would be the effect on total population numbers if intensive research on heart disease was undertaken, and it reduced deaths from that cause by 90%? This could conceivably result from a research effort comparable to present investigations of outer space. But the effort could equally be put into reduction of infant mortality. Our model could compare the effects of these alternatives, taking account of the fact that the person dying of heart disease is of such an age that he will soon die of something else; the child saved from some lethal ailment, on the average, will have a long life ahead of him. A given fall in infant mortality increases the population by much more than the same fall in heart disease.

An example of a question that has been frequently asked, and to which our model provides a clear answer, is: what would happen if, starting now, each member of the population averaged one descendant? This means that each fertile couple would need to have somewhat more than two children, to allow for those who do not marry, for those who marry but are infertile, and for deaths in childhood. An average of about 2.3 children per married couple would constitute bare replacement, that is to say, would keep the population stationary at modern death rates.

But the stationary *level* at which it would keep it would be above that of the starting point. Any population that has been subject to birth rates higher than bare replacement in the past has a large proportion of girls and women of childbearing age. These will produce increasing numbers of children for about 50 years after the date at which the birth rate falls. The projection model developed in this article tells how high the population would rise if we drop to bare replacement numbers of children.

Application of the model shows that the U.S. would rise to about 270 million persons if bare replacement were adopted today, and Mexico would rise from its present 50 million to over 80 million. Most underdeveloped countries, such as Mexico, would increase by about 65% from the point at which they drop to replacement, and they would do this over 50 or so years. No country ought to fear that immediate adoption of contraception would freeze total population where it stands; a kind of momentum operates simply because of the favorable age distribution that results from past high fertility.

Former President Sukarno of Indonesia was against birth control because he thought Indonesia should have 250 million people. The present government applied the model described in this essay and found it will probably exceed 250 million even if the brakes are put on immediately; consequently it has formally sponsored a program of birth control.

CONCLUSION

We started this essay by developing a model to forecast the future. The model works for forecasts over short periods and over longer periods in which either the trend is steady or in which ups and downs offset one another. For forecasting major turning points it is of little use, but so is any other model so far developed.

While the model is moderately, but only moderately, successful for the purpose for which it is designed, it has the power to analyse hypothetical futures whose consideration is urgent. If the marriage age in India is raised to 20, what effect will this have on the birth rate? If 20% of couples aged

30 accept sterilization, how far will this take a given country towards zero population growth? How much long-run effect does the emigration from Jamaica have on its population increase? What kind of population waves would follow a sudden drop in the birth rate of an underdeveloped country? It is in answering questions such as these, of which examples have been given in the course of this essay quite as much as in making predictions, that the projection model presented here finds its use.

PROBLEMS

1. What is a "life-table"?

2. What assumptions must one make to be able to predict, with some accuracy, the future population size using a life table?

3. In Table 2, why cannot one use the same method to calculate the three entries in the last column?

4. What is meant by "diminished stability" as a result of the new style of childbirth associated with the effective use of birth control?

5. Give some reasons why the method of prediction used in the article might give biased results.

6. What is the usefulness of the proposed model of population growth? Describe a few questions for which the model can provide answers.

7. Using the stable equivalent for 1960 females in the age group 0–14 given in Table 5, we can get the stable equivalent for the year 2065 of the same age group by multiplying 25.2 by $(1.2093)^7$.
   a. What power of 1.2093 should 25.2 be multiplied by to get the stable equivalent for the year 1990? Calculate this number.
   b. Is the stable equivalent obtained for 1990 in part a. equal to 37.7 (the number given in Table 3)? If not, how big is the difference? Compare with the difference 0.9 shown in Table 6.

8. Verify the numbers in the column for the year 2065 of Table 3 using the method given in the text in connection with Table 2.

9. Verify the solutions for the doubling time $t=70/r$ of the equation $(1+(r/100))^t=2$. If the growth rate of a country is 2%, what is the doubling time?