COHORT SIZE AND EARNINGS*

Kevin Murphy
University of Chicago

Mark Plant
University of California, Los Angeles

Finis Welch
University of California, Los Angeles

October 1984

UCLA Working Paper No. 352
November 1984

*This research was generously supported by the National Institute of Education. The authors would like to thank members of the UCLA Labor Economics Workshop for helpful suggestions. All errors are our responsibility.

INTRODUCTION

Between 1935 and 1957 the number of babies born in the United States increased from 2.4 million to 4.3 million. By 1975 the number of births had fallen to 3.1 million. Swings in birth rates (numbers of births per thousand people) were even more pronounced, rising from 18.7 in 1935 to 25.3 in 1957 then falling to 14.6 in 1975 -- the lowest level in the twentieth century.[1] These trends are plotted in Figure 1.

The arrival of the baby boom cohorts into the labor force has changed the age structure of workers and the age structure of earnings has responded. There are several papers which link the wage response to the shifting age composition of the workforce. Our purpose is to extend and update these estimates. Most authors agree that increased proportions of young workers reduce average wages of the young. There is, however, disagreement concerning the life cycle wage experience of a large cohort. In one of the early papers, Finis Welch (1979) estimated that the major effect on wages occurred at the point of entry into the labor force, and that the life cycle effects were small. In a recent paper, Mark Berger (1984) reports a contrary finding that confirms earlier work by Richard Freeman (1976, 1979). In this paper we find that the entrance of a large cohort depresses young workers' wages substantially, but that over time relative wages of workers in large cohorts rebound and that in fact the lifetime loss in wages is small.

Our data refer to men in the labor force observed during the period from 1967 to 1982. The first section of the paper describes the data and summarizes the labor force changes observed during this period. Next we describe the basic methods for estimation. In the third section we provide our main results and a number of simulations. The final section contains a summary.

FIGURE 1: SOLID LINE: NUMBER OF BIRTHS
BROKEN LINE: BIRTH RATE
SOURCE: 1982 STATISTICAL ABSTRACT OF THE UNITED STATES

THE DATA AND THE EMPIRICAL SETTING

The data are from the March Annual Demographic Survey of the Current Population Surveys, 1968 to 1983 and refer to white males who are civilians between the ages of 14 and 65, in the labor force and not in school.[2] Self-employed and retired workers are excluded. A detailed description of the procedures used to develop the data base is given by Welch (1979). Our measure of experience in the workforce is not the standard age minus education minus 6, but an imputed measure based on education level and year of birth. This imputation is described by Welch and William Gould (1976). The unit of observation is an experience/education/year cell for which we observe average weekly earnings in the previous year and the number of persons in that cell.[3] We work with 40 experience levels, 4 education levels (8-11 years of schooling, 12 years, 13-15 years and 16 or more years) for the 16 years from 1967 to 1982. Thus there are 2560 observations.

The change in the age distribution of the male labor force in this period was remarkable. As the workers who were born in the baby boom began their working careers the labor force took on a much younger complexion. Table 1 illustrates the changing age distribution of the workforce by giving the percentage of men who are in their first five years of work experience since leaving school. Although we only report four-year averages, the sixteen year period begins with 14.5 percent of all men in the labor force being in their first five years out of school. We have arbitrarily restricted the workforce to the first 40 years of experience, and therefore the percentage in their first five years begins only slightly higher than the 12.5 percent one would expect with the same number of entrants every year and zero mortality. By 1976, 22.6 percent of the labor force was in this low experience group. Thus in only nine years, the share of the youth cohort grew by more than one half.

After this crest of the baby boom, the percentage of young workers began to decline so that by 1982 only 19.7 percent of white male workers are in their first give years in the labor force. Relative to the stationary state, the 1982 labor force is still very young.

Because the baby boom cohorts had higher educational levels than preceding entrants, the pattern for college graduates is even more pronounced. Table 1 shows the four-year averages for college graduates in the third column. In 1967, 19.3 percent of male college graduates were in their first five years in the labor force, and by 1975 this percentage had grown to 25.3 percent. After this peak the percentage began to fall, reaching 17.9 percent by 1982. For high school graduates the increased percentage of young workers reaches a lower peak and falls off less rapidly as the baby boom exits the labor force.

It would be surprising if the labor market could digest such a large increase in the numbers of new entrants without reducing their average wage. Table 2 shows that such a reduction did occur. In Table 2 we report the average weekly wage of young workers (1-5 years of experience) to that of older workers (28-32 years of experience). As in Table 1, we report four year averages of the yearly figures. A typical feature of age-earnings profiles is that earnings increase with time out of school for about 30 years until earnings are roughly 60 to 80 percent higher than for new entrants. After the maximum at 30 years, wages decline gradually until retirement, at which time workers earn 5 to 10 percent less than at the peak. In the early years of our sample, earnings for college graduates were 62 percent of those for peak earners, but fell to 52.5 percent during the 1975-78 period. Thus, compared to prime age earnings, the earnings of new entrants dropped 15 percent.[4] In Figure 2, we plot the cross-section experience profile of weekly wages for

Table 1

PERCENT OF MEN IN THE LABOR FORCE WHO ARE IN THEIR

FIRST FIVE YEARS OUT OF SCHOOL

	All	High School Graduates	College Graduates
1967-70	15.8	16.2	20.7
1971-74	20.6	20.7	23.9
1975-78	22.2	23.1	23.5
1979-82	20.7	22.8	19.0

Table 2

AVERAGE WEEKLY WAGES OF MEN IN THEIR FIRST FIVE YEARS SINCE LEAVING SCHOOL AS A PERCENTAGE OF WAGES FOR MEN IN THEIR 28TH-32ND YEARS

Year	High School Graduates	College Graduates
1967-71	64.2	61.9
1971-74	57.5	53.7
1975-78	56.7	52.5
1979-82	55.9	55.4

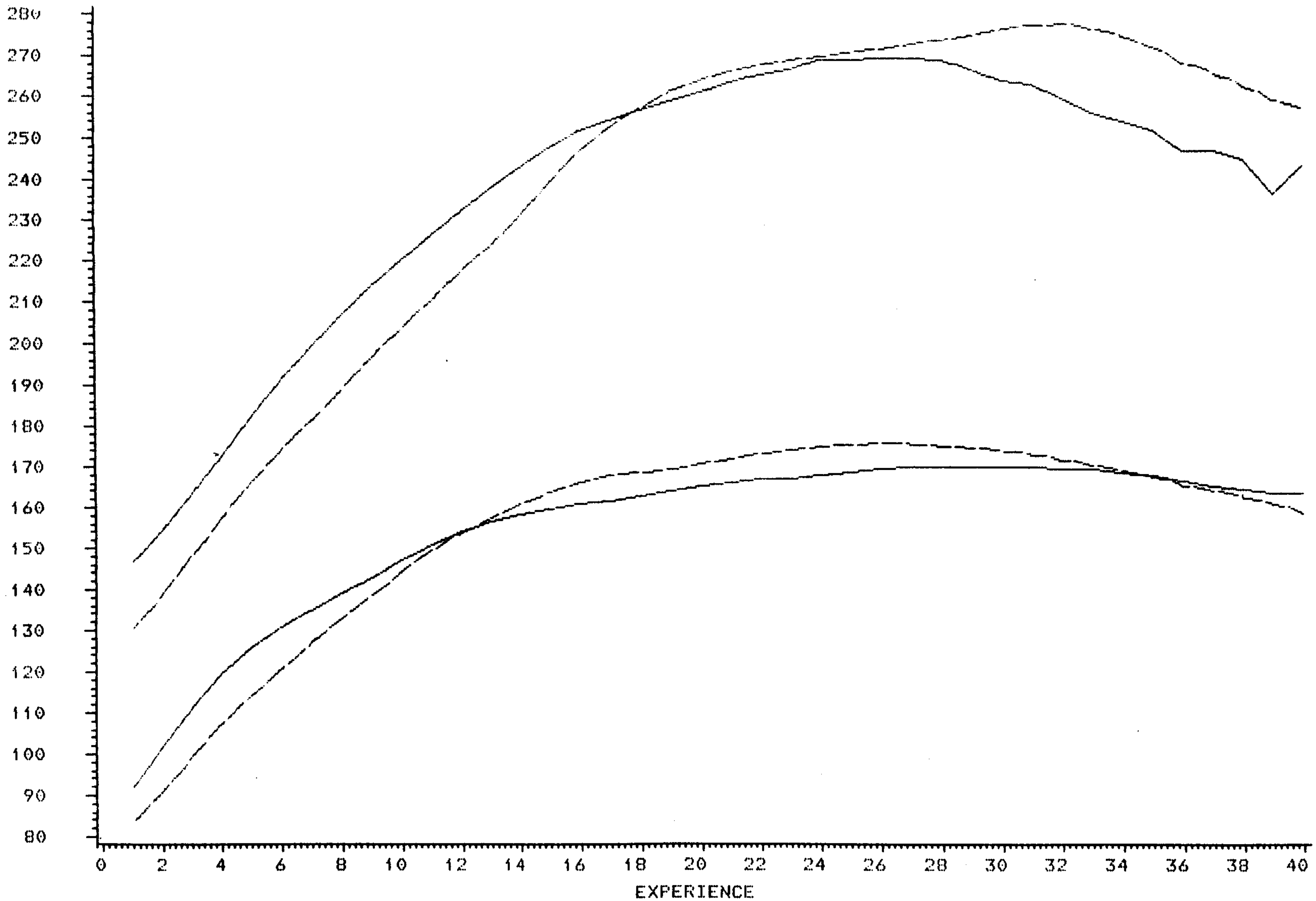

FIGURE 2: WEEKLY WAGES (1967 DOLLARS)

UPPER LINES--COLLEGE GRADUATES   LOWER LINES--HIGH SCHOOL GRADUATES

SOLID LINES: 1967
BROKEN LINES: 1977

high school and college workers in 1967 and 1977, in constant dollars. These figures further illustrate the change in the wage distribution that occurred as the baby boom entered the labor force. For both education levels, the age earnings profiles became steeper, with younger workers taking a 10 to 15 percent decrease in wages, and older workers receiving a higher relative wage. This shift in the age profile of earnings is particularly pronounced for college educated workers.

In working with large samples, one typically observes considerable stability in wage comparisons across age and education groups. Swings such as those summarized in Table 2 and illustrated in Figure 2 are uncommon, and we suggest that there has been a real shift in the age structure of compensation. We, of course, think that this shift was primarily caused by the rapid increase in the fraction of all workers who are young, and it is the relationship between the age structure of earnings and the age structure of the population that we address in this paper. In particular, we wish to develop a theoretical and empirical methodology that will allow us to predict if the decreased relative wages experienced by workers in large cohorts will continue throughout their lives, or whether the life-cycle wage effects are concentrated in the early years of the career. We review this methodology in the next section.

EMPIRICAL METHODOLOGY

The premise on which our empirical work rests is that a worker can be considered as a bundle of attributes or underlying factors of production. Each underlying factor has a well-defined market price determined by ordinary market forces. A worker's wage is the total compensation to these factors. For example, in a two-factor scheme, an individual's wage is the sum of the payments to the two factors where each payment is the quantity of the factor held by the person times its price.[5] As a worker ages, quantities of some skills increase while others may decrease. Workers of differing education may have different initial bundles and different life cycle patterns as well.

The flexibility of this scheme as a descriptor is easy to see. Take, for example, a two-skill model and consider the life cycle path of two groups of workers. Call the skills A and B and the groups high school and college graduates. Suppose that recent high school graduates have only skill A and that as they mature the quantities of A increase. College graduates start with both A and B and as they age quantities of B increase while quantities of A decline. Thus A is the high school specific factor.

One implication of this simple scheme is that young high school and college graduates are better substitutes for each other than mature workers of the two groups. This, of course, assumes that the underlying skills, A and B, are not themselves perfect substitutes.

To see how a scheme like this can produce the kinds of patterns seen in the data, examine Figure 2. In it, wages in 1977 are lower than in 1967 for college graduates with less than 18 years of work experience. The profiles cross at 18 years and 1977 wages exceed 1967 wages for older workers. Presumably the prices of A and B (the underlying factors) in a given year are the same for all workers but change between years.

The general upward drift in wages as workers age (for either year) is presumably caused by the dominance of the growth in B over the loss of A. Yet because young college graduates are relatively A intensive the twist experienced between 1967 and 1977 could have resulted from an increase in the price of B alongside a falling price of A.

Now examine the high school profiles in Figure 2. The model just described assumes that high school graduates have only one factor, A, and the life cycle path refers only to growth in the quantity of that factor. If so, a change in the price of A between 1967 and 1977 would result in a parellel shift in the earnings profile. Thus the observed crossing pattern is inconsistent with the one-factor model.

One alternative to the factor analytic approach is to treat each level of experience and education as a separate factor of production, but then the implicit production function would have 160 labor inputs for our data set so this approach is not feasible. Another alternative is to divide workers into age intervals to decrease the number of factors of production, but such divisions are arbitrary and result in predictions of discontinuous changes in wage profiles.

Returning to the factor structure, suppose that there are M underlying factors, indexed by m, and consider a worker with i years of experience and e years of schooling. Our basic supposition is that this worker's value of marginal product (VMP) in time t, y(i,e,t), is a linear combination of the VMP's of each factor. Specifically:

$$y(i,e,t) = \sum_{m=1}^{M} g(i,e,m)\, w(m,t)$$

where w(m,t) denotes the VMP of each factor of production and g(i,e,m) denotes the amount of factor m imbued in a worker with i years of experience and e years of education. Notice that the life cycle compositional path of a worker with education e is characterized by the pattern of weights g(i,e,m), as i goes from 1 to I (m=1,...,M). The time pattern of wages, holding experience and education constant, is based on the changing VMP's of the basic factors of production. If we array the observed wages of workers with a given level of education by experience level and time we can then write

$$Y(e) = G(e)W$$

where $Y(e)$ is the IxT matrix of observed wages, $G(e)$ is the IxM matrix of factor weights and W is the MxT matrix of the VMP's of the underlying factors, each for educational level e. Assuming that the factors underlying each education level are the same, we can stack the system so that

$$\begin{bmatrix} Y(1) \\ Y(2) \\ \vdots \\ Y(E) \end{bmatrix} = \begin{bmatrix} G(1) \\ G(2) \\ \vdots \\ G(E) \end{bmatrix} W$$

or more succinctly

$$Y = GW$$

where Y is the matrix with (IxE) rows and T columns, G is (IxE) by M and the matrix W is M by T. Clearly such a characterization is not unique since for any MxM non-singular matrix A we can write

$$Y = (GA^{-1})(AW)$$

or

$$Y = G^{*}W^{*}.$$

Given the number of factors, M, we must impose MxM normalizations on the factor analytic representation of the wage data. Clearly if $M=T$, then the data can be described perfectly by the factor scheme. For any $M < T$ the fit in an empirical setting will not be perfect. Our stochastic model, then is

$$Y = GW + U$$

where $M < T$ and U is a matrix of errors. We estimate G and W using weighted least squares.[6] As described elsewhere,[7] we chose to use three factors to characterize our data, since additional factors contributed little to explaining the variance in factor profiles. For our sample T is 16, I

is 40 and E is 4, so we explain Y, the 160x16 matrix of wages, by a matrix G of factor quantities which is 160x3 and a matrix W of factor loadings which is 3x16. We can think of an element of G, say g(i,k), as representing the quantity of factor k in an individual in experience/education cell i, and an element of W, say w(m,t), as the VMP of the m^{th} factor in year t.

Let p(i,t) denote the number of men in experience/education cell i in year t. Then P = (p(i,t)) is the 160 by 16 matrix that describes the experience/education distribution of the population. Now define

$$X = G'P$$

where X is the M by T (3x16) matrix that represents the total quantity of each factor over the T (16) years of the sample.

Presumably the wage structure W will depend on the amount of every factor X. If the factors that we extract from our estimation process are the underpinning economic characteristics of workers then the wage of factor m should decrease when the amount of factor m present increases, <u>ceteris paribus</u>. Thus the wages, W, and the factors X represent a 16 observation data set, with wages being the dependent variables. We regress the wage of each factor on the quantities of all factors. With these regressions in hand we have a means by which we can calculate the effect of changes in the distribution of education and experience on the labor force. Specifically, let the estimated regression structure of factor wages W be

$$\hat{W}' = X'\hat{B}$$

where $\hat{B}$ is the matrix (3x3) of regression coefficients. Suppose that in a given hypothetical year, the experience/education distribution is characterized by a 160 by 1 vector $\overline{p}$. Given $\overline{p}$, we want to know what the wage structure will be. The 3 by 1 vector of factor quantities is $\overline{x} = G'\overline{p}$, and the predicted factor VMPs are $\overline{w}' = \overline{x}\hat{B}$. Given these factor predicted wages we can compute the wage structure $\overline{y} = G\overline{w}$. The factor decomposition gives us a means of predicting the wage structure across experience/education cells from any distribution of population among those cells. The factor loadings G and the factor wages W are arbitrary since we have chosen MxM normalizations, but we show in Appendix A that the predictions $\overline{y}$ are not in any way dependent on the the normalization chosen. The factor analytic decomposition has given us a means to characterize the sensitivity of wages to various changes in the composition of the labor force. By choosing appropriate hypothetical distributions, $\overline{p}$, we can isolate the effects of age, education and business cycle changes on the wage distribution.

The factor decomposition we performed used three factors to explain the 160 by 16 matrix of observed weekly wages. Here we describe the factor loadings and factor weights generated. We present the actual results elsewhere.[8] Within an educational level, the first factor is shaped like the mean experience-earnings profile, and as education increases, the endowment of the factor increases. Alone, the first factor explains 99 percent of the variation in the observed earnings profiles. If we analyzed our data using this one factor alone, we could only explain variation in experience-earnings profiles as parallel shifts of this mean profile within education group and could not explain any change in the shape of the experience-earnings profiles. The second factor profile is primarily youth intensive, with a peak between 5 and 10 years of experience and a steady decline until the last few years of the profile at which time it turns upward. Like the first factor, more highly educated groups have higher endowments, except that the men with 8-11 years of schooling have a higher endowment than high school graduates. The third factor profile is hump-shaped, peaking around 20 years of experience. Youth have negative endowments of this factor as do the oldest members of the workforce, and endowments once again increase with increasing education. The second and third factors explain 80 percent of the residual variance after the first factor is added.

The time pattern of the factor wages is critical in understanding how the factors change the wage profiles. The wage of the first factor increases throughout the period essentially accounting for the steady upward drift of wages during the three year period. The second factor has a positive wage in the early years of the period, but by 1975 the wage has become negative. Recall that the second factor is youth intensive (and old age intensive) so that in the early years this factor flattens the wage profile by raising young and very old wages relative to prime age earners. The third factor, which is prime-age intensive exhibits the opposite pattern, its wage being negative in the early years of the sample, positive in the middle years, then turning negative in the late 1970s as the baby boom enters the middle years of its career. In the last two years of the sample, during the early 1980's recession, the youth intensive factor has a large negative wage and the middle age intensive factor has a large positive wage. This indicates that there may indeed be some cyclical sensitivity in these wage patterns.

To explain the factor wage pattern, and to allow prediction of factor wages under various age and education structures of the population, we regressed the factor wages on the factor quantities and on the prime age male unemployment rate. The dependent variables in the three equation system of wage equations were factor wages normalized by a constant quantity price index. The prime age male unemployment rate was included to pick up any fluctuations in factor wages due to the business cycle. We constrained the wage system so that it was homogeneous in wages and the partial effects of quantities of factors on wages were symmetric.

The regression results show that the partial effects of quantities of factors on own wages are large, negative and significant. The cross effects are all positive (an increase in the number of workers of one factor lowers the wage of other factors, holding all else constant). The only insignificant coefficient is on the third factor in the second wage equation (and vice-versa). The prime age male unemployment rate has a negative and significant effect on the wage of the first factor, an insignificant effect positive on the wage of the second factor and a significant positive effect on the wage of the third factor. Thus, in bad economic times, when the unemployment rate is high, the average wage profile is lowered, but also the teenage intensive factor (factor two) has a low wage and the middle age intensive factor has a high wage. This confirms the notion that teenagers suffer relatively more during recessions than do middle age workers. The importance of these regressions is that they allow us to generate a full 160 by 160 substitution matrix. We can compute the partial effect of a change in the quantity of workers of a particular experience and education on the wage of a worker of any other experience and education. We also have implicitly estimated the response of the wages of all 160 types of workers to changes in the prime age male unemployment rates. Clearly it is impossible to present all these partial effects in a succinct manner, so we summarize the implications of these results using simulations.

SIMULATION RESULTS

The factor analytic techniques reported here allow us to simulate wage profiles generated by any distribution of education and experience in the labor force. In particular, we can isolate the effects of changes in the age structure of the workforce from those changes due to changing education or business cycle effects. In the simulations we present here, we hold the education distribution at its mean across the period 1967-82 and the prime age male unemployment rate at its mean of 2.9%.

Figure 3 isolates the effects of a change in the age distribution on the cross-sectional wage profiles and relative wage profiles. In Figure 3 we present the cross sectional wage profiles for high school graduates and college graduates using the mean distribution of education under two age distributions: the 1967 age distribution and the 1976 age distribution. We held population constant at its mean for 1967-82. Note that in times when there are a large number of mature workers (1967), young workers do relatively better, and that mature workers do relatively better when they are more scarce. This complementarity between older and younger workers seems to lessen as workers enter the last 5 years of their careers. The substitutability between very young workers and very old workers is evident from the factor patterns we discussed in Section 2.

The effect illustrated here -- that an increase in the relative amount of any particular age cohort of labor will decrease their wages -- has led authors such as Freeman and Berger to conclude that over their entire career, the baby boom will suffer relatively lower wages. Berger, in fact, concludes that the relative wage effect will worsen as the baby boom reaches middle age. The simple simulations we have considered indicate that the more scarce a particular age group, the higher the wage paid to that group. The size of the effect in terms of lifetime income however cannot be ascertained from such cross-sectional simulations. To focus on the lifetime effects, we ran a simulation that tracked a synthetic baby boom generation throughout its career

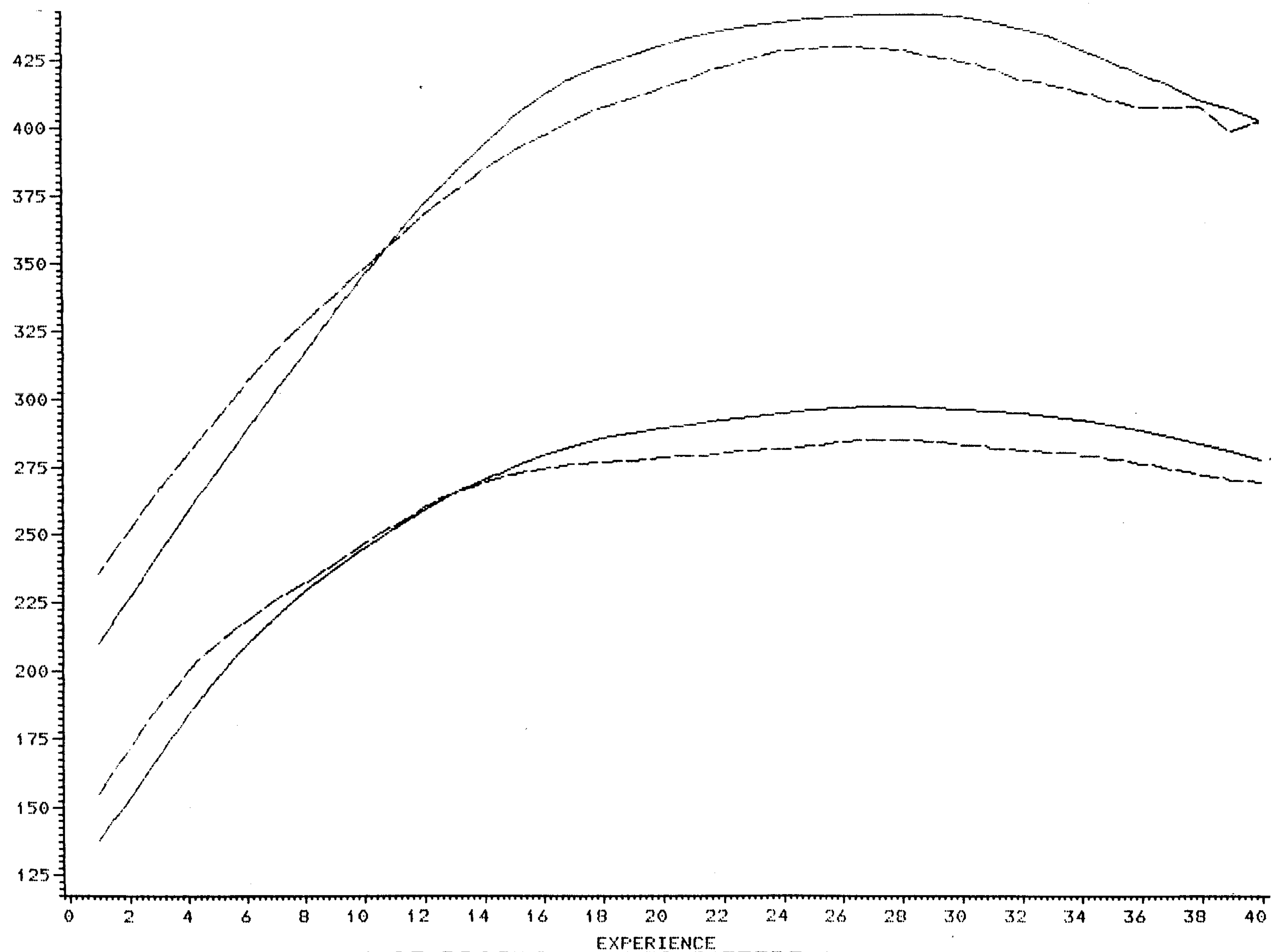

FIGURE 3: WAGE PROFILES UNDER DIFFERENT AGE DISTRIBUTIONS
SOLID LINE: YOUNG (1976) AGE DISTRIBUTION
BROKEN LINE: OLD (1967) AGE DISTRIBUITON

to see if the effects were indeed lasting. We began with the 1967 experience distribution, as a hypothetical long run equilibrium value for the population. We then passed a 31 year "V-shaped" baby boom through the population. Specifically we introduced the first year of the baby boom by assuming a growth in the number of first year workers of 2.6%. The following year we aged the first workers to second year workers and introduced a new first year cohort 5.2% larger than steady state. We continued in this fashion increasing the percentage augmentation by 2.6% per year until it reached a value of 39% in the 15th year, and we then decreased the augmentation by 2.6% per year until it reached zero in the 32nd year. Assuming the baby boom began in 1942, the peak year of birth rates in our simulation was 1957 when the cohort was 39% larger than steady state. Throughout the simulation the education level was kept at its mean level in our sample, and was independent of year of birth. Thus we expect the effects reported to be smaller than that would occur if the education distribution changed simultaneously. These figures roughly describe the actual observed pattern of population increase. We trace the wage distribution of the population from the time the first year of the baby boom enters the labor force until the last year exists the labor force, when the population returns to steady state.

In Figure 4 we plot the lifetime wage profile of college graduates who were born in 1957 and 1981 in our hypothetical simulations. As the results below indicate, these two years represent the extreme values in the present value of total lifetime earnings. Babies born in 1957 had the lowest earnings, and those born in 1981 had the highest. Recall that in our simulation 1957 is the peak of the baby boom. By the time the 1981 cohort reaches the labor market, the 1957 cohort is in their twenty-fourth year of experience so there is a relative shortage of young workers. The wage pattern shows that

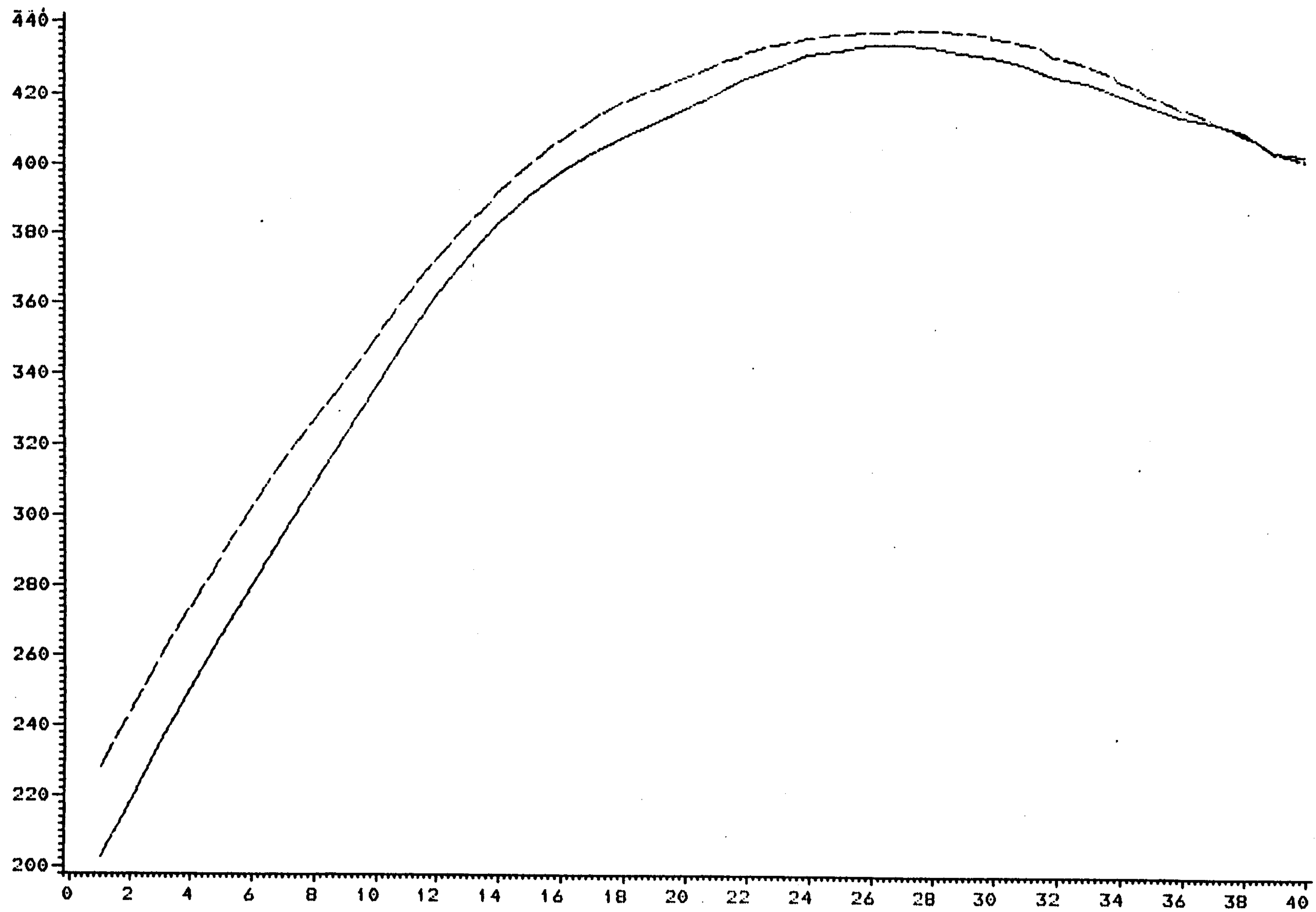

FIGURE 4: WAGE PROFILE OF COLLEGE GRADUATES BORN IN 1957 AND 1981
SOLID LINE: 1957
BROKEN LINE: 1981

workers born in 1957 receive wages approximately 10% lower than the cohort born in 1981 in the initial phases of their career. Until the last few years of the career the wage of babies in the large cohort is consistently less than that of the small cohort, but the percentage difference declines steadily. By the time the cohorts are at the peak earnings level the percentage difference is only 1%. In Figure 5, we graph the ratio of the 1957 cohort earnings to the 1981 cohort earnings by year of experience. This figure reinforces our observation that at the beginning of the career workers in large cohorts suffer large wage losses, but as their career progresses and they are absorbed into the labor force the wage loss they suffer declines, and by the end of their career it vanishes altogether. These simulations belie the claim that the baby boom will suffer from increasing large wage penalties as their careers proceed.

Since the large decrease in wages we see in the early phases of a large cohort's career does not last throughout the career, it is appropriate to ask what the lifetime effect on earnings might be. Assuming a 5 percent real rate of discount, we calculate the present value of earnings of each birth year cohort relative to the present value of earnings in the steady state. This relative present value is plotted by birth year for high school graduates in Figure 6 and for college graduates in Figure 7. Recall that the synthetic baby boom begins in 1942, peaks in 1957 and ends fifteen years later in 1972. High school graduates who were born before 1947 experience an increase in relative wages since they are prime age workers when the baby boom enters the labor force. Their relative scarcity drives their wages up during their peak earnings years in their careers. Lifetime wages go below steady state levels in 1947 and reach the lowest point for cohorts born in the late 1950s. Workers born in the 1970s and early 1980s receive a premium for being young

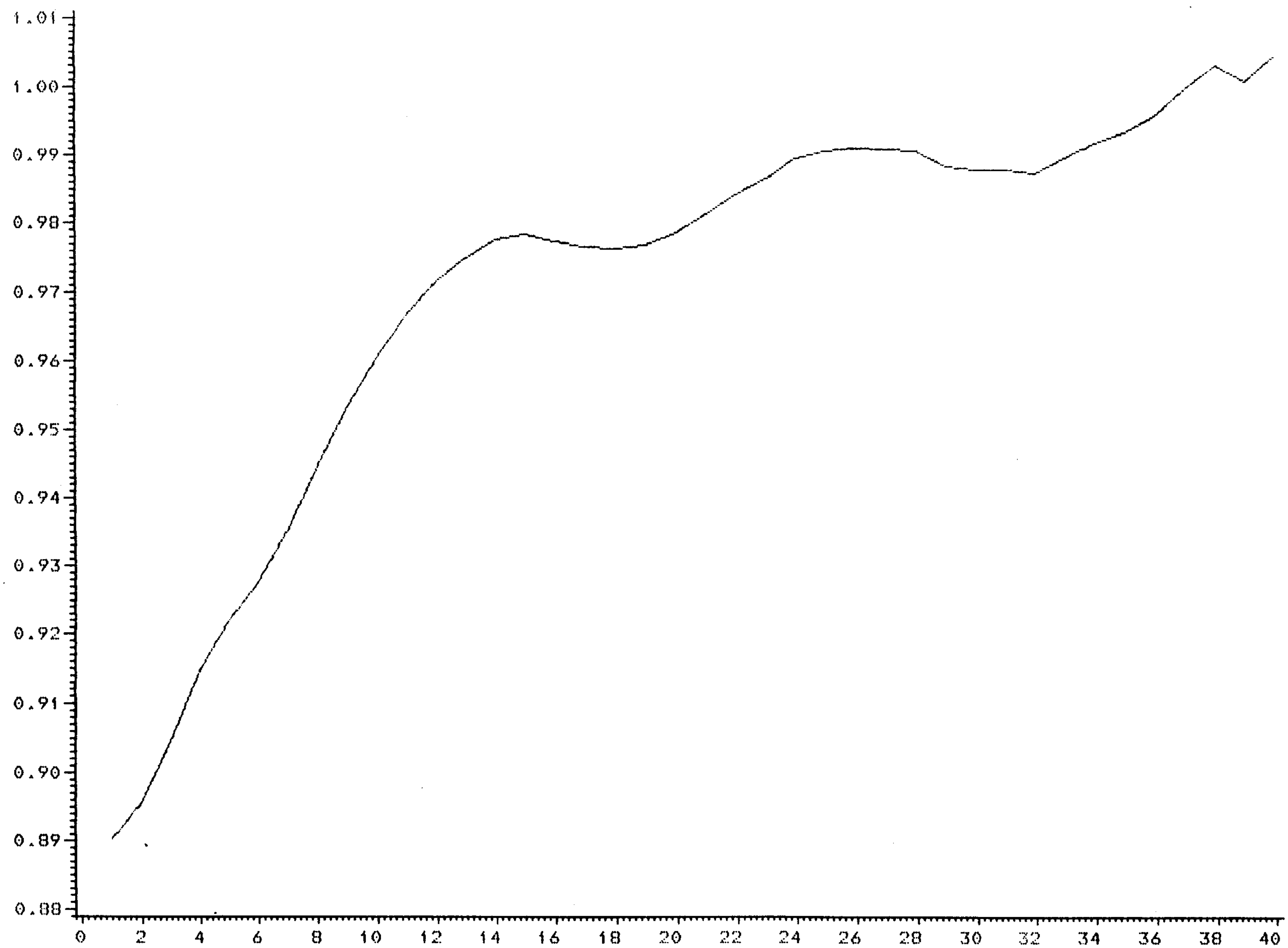

FIGURE 5: RATIO OF 1957 WAGE TO 1981 WAGE--COLLEGE GRADUATES

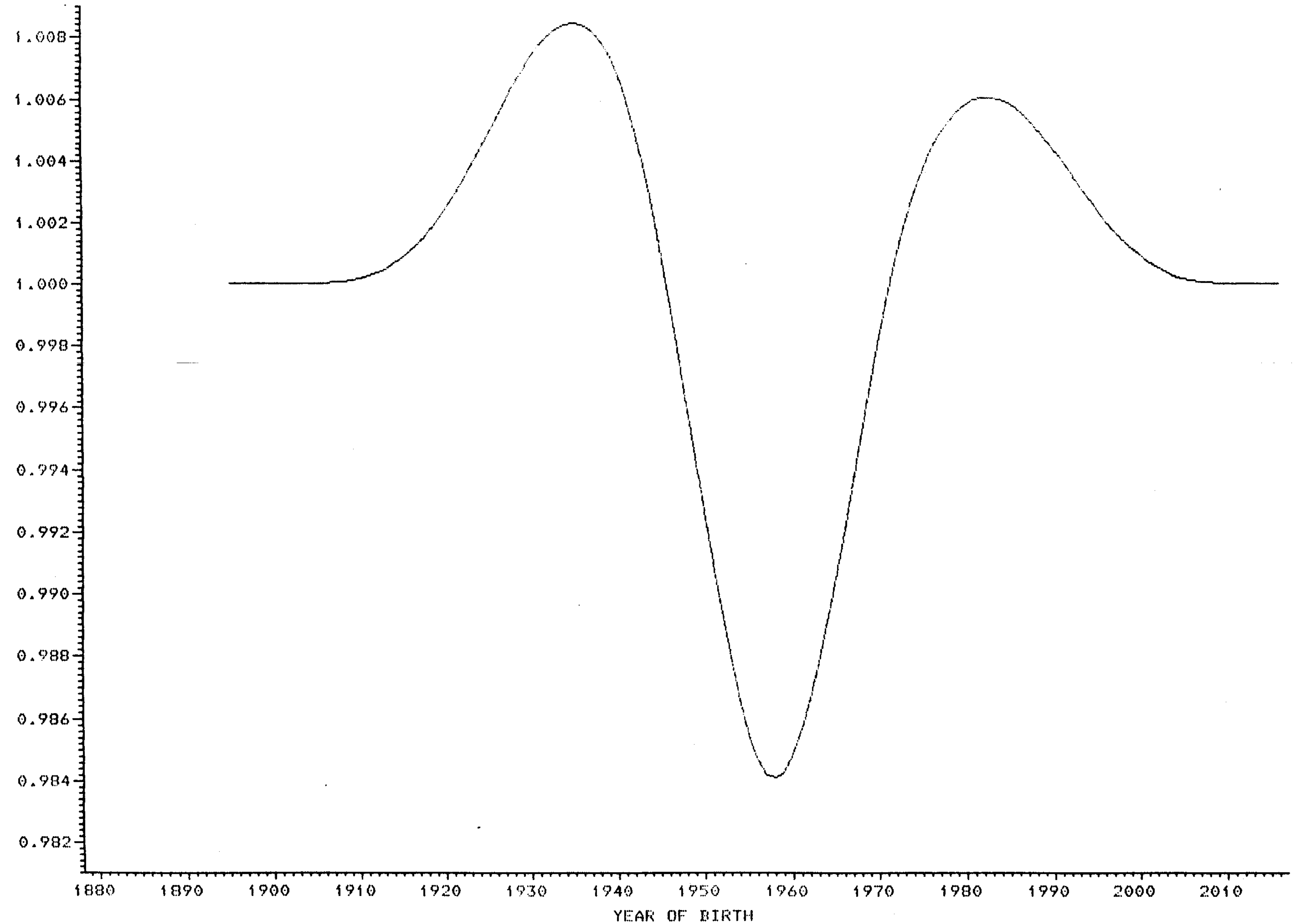

FIGURE 6:RELATIVE PRESENT VALUE OF WAGES UNDER SIMULATED BABY BOOM
HIGH SCHOOL GRADUATES

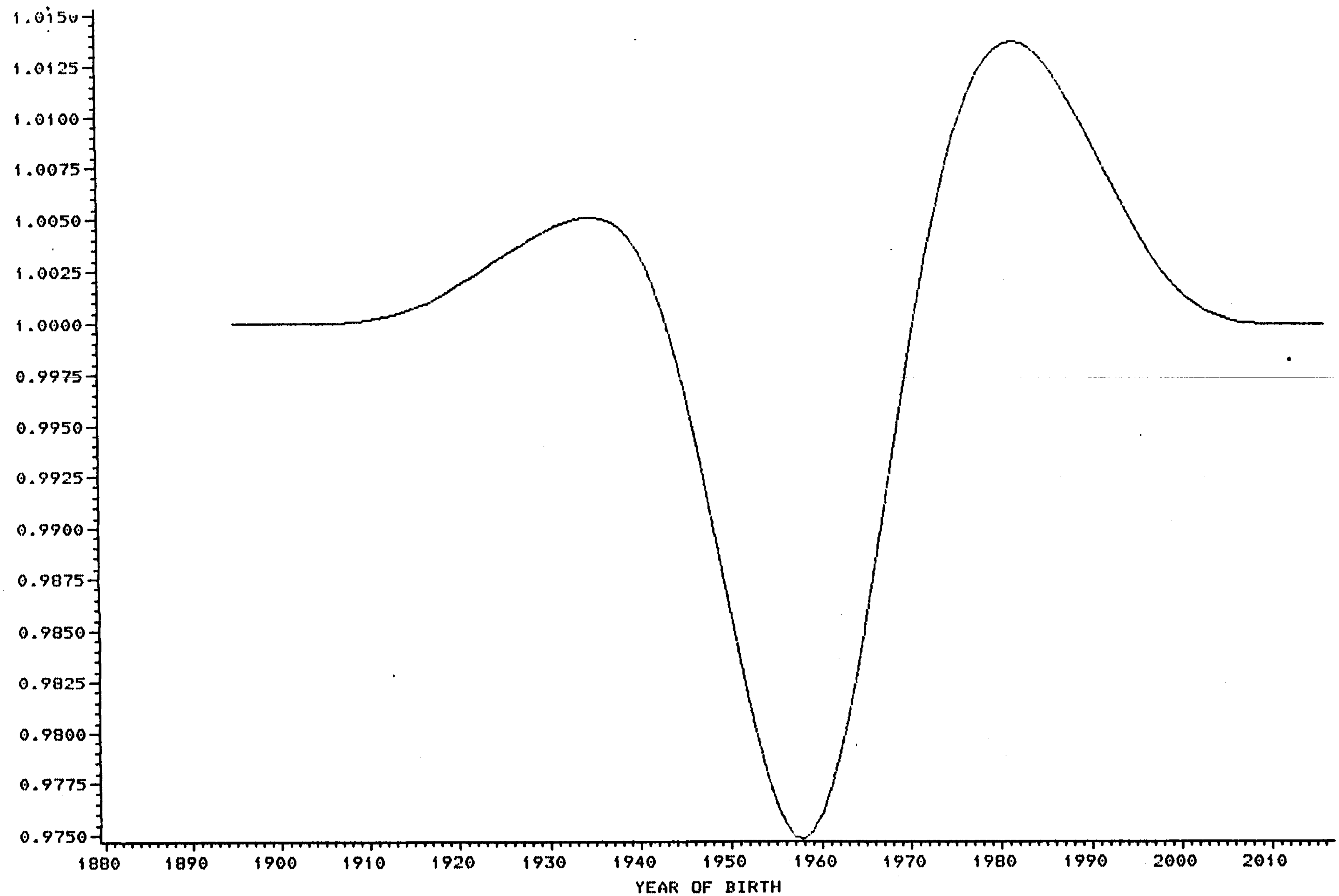

FIGURE 7:RELATIVE PRESENT VALUE OF WAGES UNDER SIMULATED BABY BOOM COLLEGE GRADUATES

workers during the time in which the baby boom are peak earners. The level of lifetime wages returns to its steady state when the last of the baby boom cohorts leaves the labor force -- the year the babies born in 2012 enter the labor force. The story for college graduates is roughly the same as for high school graduates, except the initial peak is much smaller, and the premium to those born in the 1970s and 1980s is much larger.

The most striking observation to be made about Figures 6 and 7 is that the difference in lifetime wages is very small. From peak to trough the difference is only about 3% for high school graduates and 4% for college graduates. High schools graduates born in the least favorable year experience lifetime wages only 1.6% less than steady state levels, and for college the discount is 2.5%. This contrasts markedly with the initial impact of increased cohort size on young workers. Recall from Figure 4 we concluded that the impact of cohort size in the early stages of a large cohort's career is large -- around 10% in our simulation. However, the entire lifetime the net effect is minimal for most workers.

Some authors have argued that college educated workers are particularly hard hit by the recent changes in labor force composition since there is a surplus of educated labor as well as young labor. In this simulation we do not change the education distribution, but the changing age distribution does have an impact on the return to education. In Figure 8, we plot the ratio of the present value of college earnings to the present value of high school earnings by birth year given the simulated baby boom. In terms of lifetime earnings the decline in the return to education begins for persons born well before the baby boom. This demonstrates the notion that the effects of large cohorts increase with the level of education, which was evident from the increase in all factor endowments as education increased. Once again these

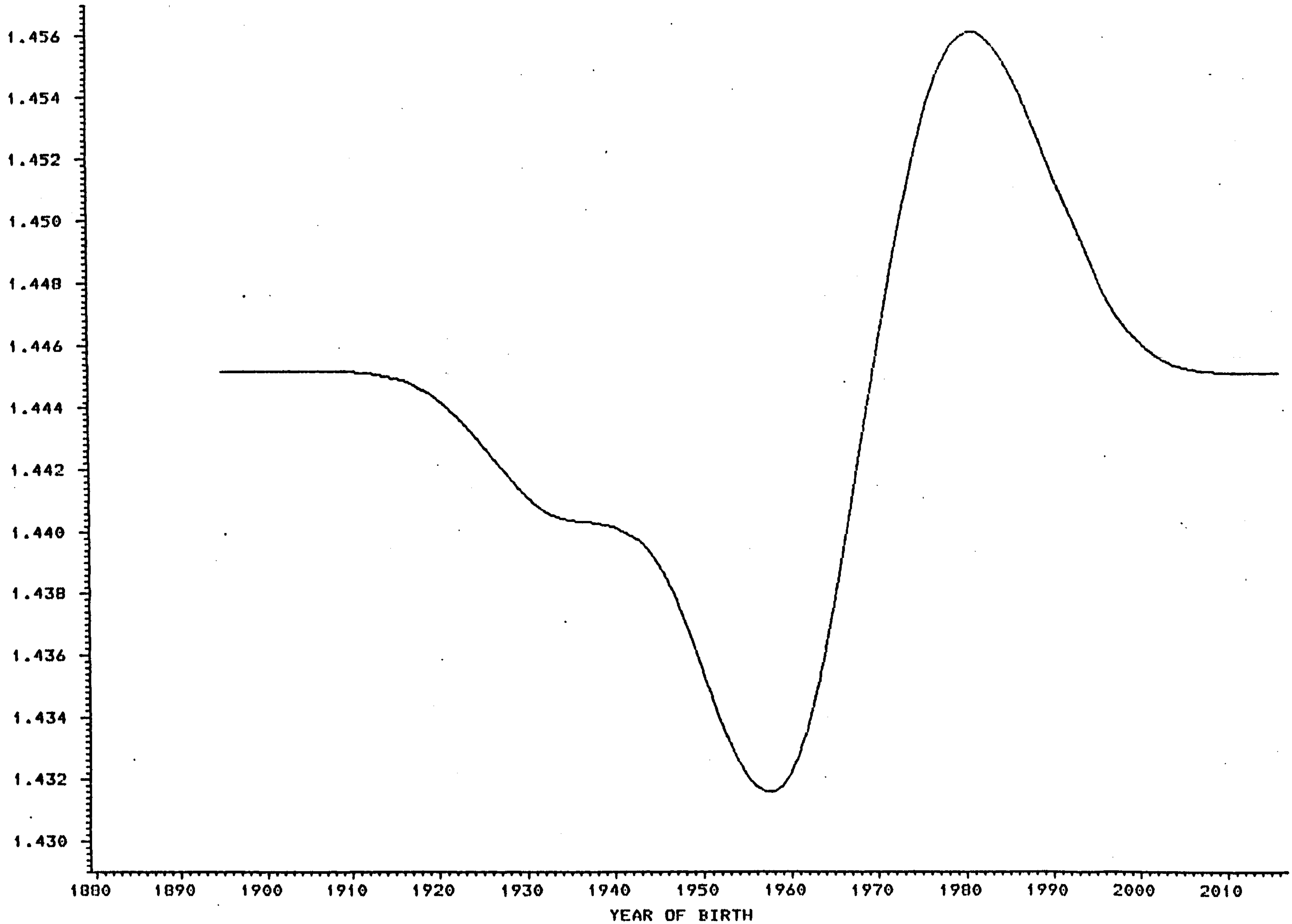

FIGURE 8:RATIO OF PRESENT VALUE OF WAGES: COLLEGE TO HIGH SCHOOL UNDER SIMULATED BABY BOOM

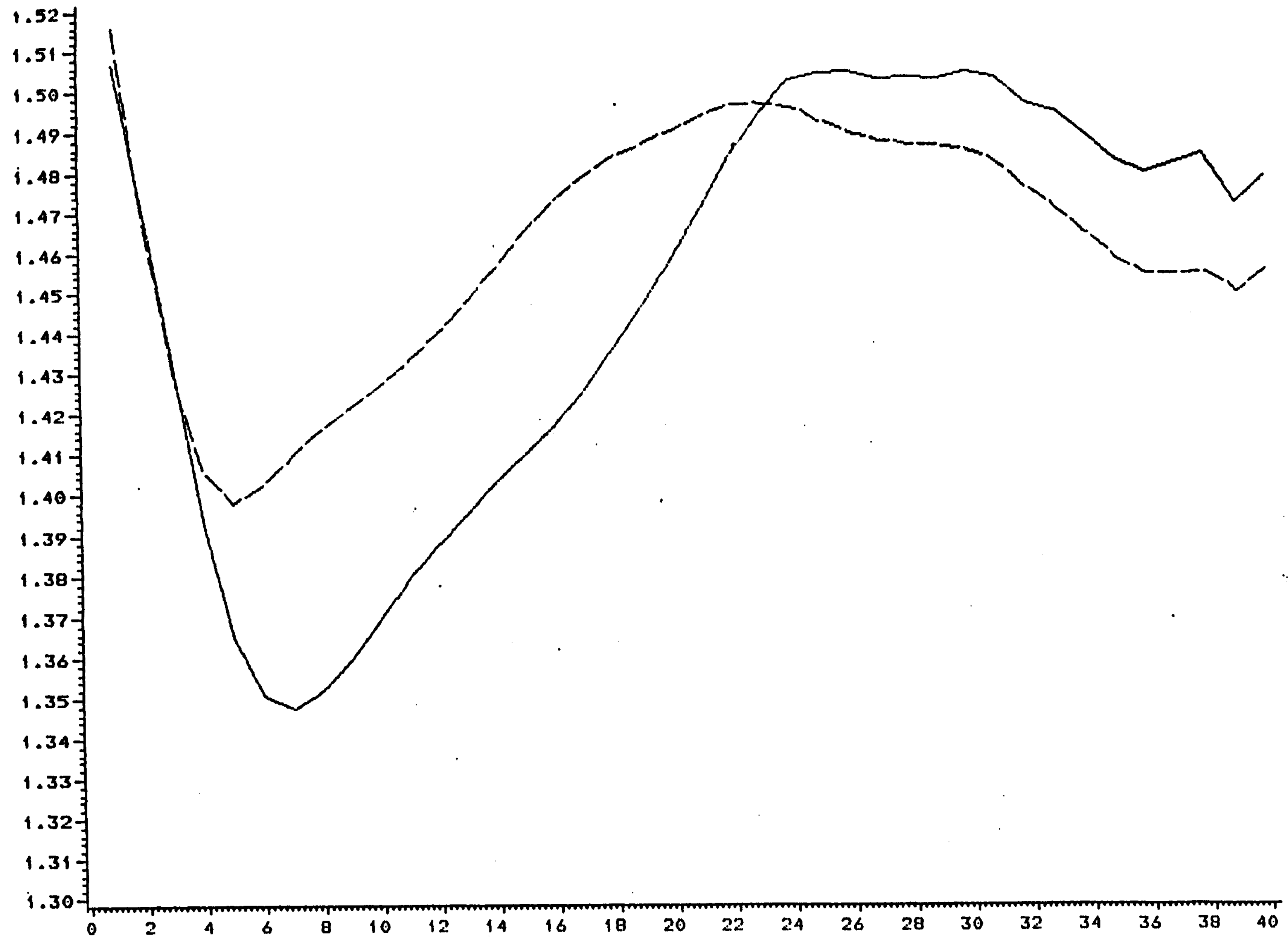

FIGURE 9: RATIO OF COLLEGE TO HIGH SCHOOL WEEKLY WAGES FOR COHORTS BORN IN 1957 AND 1981
SOLID LINE: 1957
BROKEN LINE: 1981

lifetime effects are relatively small, with the lifetime premium to a college education varying from a low of 43% to a high of 45%, the steady state value being 44.5%.

Although these estimated effects on the lifetime premium to a college education are relatively small, specific year effects can be quite large. In Figure 9, we plot the relative wages of college to high school graduates over their lifetimes for the 1957 birth year and the 1981 birth year. College graduates born during this simulated baby boom suffer from a lower educational premium than do their small cohort counterparts, but by the end of their life, the pattern is actually reversed -- the rewards to education increase faster for the large cohort babies and remain at higher levels later in their lives. The decreased premiums observed currently for the baby boom babies have led some authors to assert that these educated workers will suffer throughout their lives from "overeducation." Our analysis leads us to expect that the decreased relative wages of educated baby boom babies is a temporary phenomenon and that their relative wages will increase in the course of their lifetime.

CONCLUSION

The labor force has undergone large changes in demographic and educational composition in the twentieth century. It is clear that demographic fluctuations have had an economic impact on workers, but the nature of that impact has been unclear. In order to attribute fluctuations in wages to fluctuations in labor force composition, it is necessary to develop an empirical framework in which the effects of composition on productivity is well-defined. In this paper we develop such a framework by viewing workers as a composite of several productive factors, a worker's wage will change if the composition of factors changes or if the value of any of those factors change. This leads us to analyze the sixteen years of observed wage patterns using a factor analytic technique. The advantage of such a technique is that it allows us to summarize the changes in the composition of the labor force succinctly in terms of a small number of underlying factors of production rather than having to resort to arbitrary divisions of workers into factor groups or to a complicated production technology with numerous factors and numerous restrictions to make the technology estimable. In our estimation we saw that we could characterize most of the variation in wage patterns using three underlying factors of production. The first factor reflected the increase in skills that comes with experience. The second factor was a youth and old-age factor, and the third factor was a middle age factor. In regressing the quantities of these factors on their wages, we were able to build a structure that allowed us simulations of wage patterns given any educational or age distribution, holding constant business cycle effects. Although this structure enabled us to build a complete substitution matrix, that way of summarizing the interactions of 160 different productive types of workers was hardly succinct, so we resorted to these simulations to illustrate our results. The first simulation simply contrasted the wage patterns in a youthful population with that in an older population. These cross-section results showed that the rewards to scarcity of factors exist but are relatively small. We then went on to consider the effect of the entrance of a baby boom on lifetime earnings and found that although the depression of wages caused by large cohorts could be large during the initial stages of the large cohort's career, the wage differential diminished over the course of the career. The individual year effects in our simulations were as much as 10% but the effects on lifetime earnings were at most 3%. Our evidence also indicates that a youthening of the labor force does not lead to a large lifetime decrease in the return to college education nor to a permanent decrease in that return.

This list of potential avenues for future research is long. Of immediate concern is running simulations of the baby boom that accurately reflect the changing educational composition of the labor force as well as the changing age composition. We also recognize that the emergence of a large number of female workers may have had an impact on the wage structure during this period. Other outside influences such as affirmative action, the end of the Vietnam conflict and the draft have been omitted here. Our contribution has been both methodological and substantial, and we believe the methodology is sufficiently flexible to permit much more fruitful resarch.

FOOTNOTES

[1]These data are taken from The 1982 Statistical Abstract of the United States.

[2]We do not include black men or women in our analysis because there are various issues, such as affirmative action, that analysis of the group would necessarily involve. Such issues are beyond the scope of this paper.

[3]The cell population is created by applying the Welch and Gould experience weights to the CPS person weight.

[4]The detailed year-by-year data show a drop of 25 percent from the peak in 1967 to the trough in 1976.

[5]See Welch (1969) for the basic reference of this topic.

[6]Clearly this estimation could be done using ordinary least squares, but recall that the observations on the wages are cell means. Thus a weighted scheme seems more advisable to account for the heteroschedasticity resulting form varying cell size. The cell sizes are computed by applying the experience weighting scheme described in Gould and Welch to the CPS weights given in the original sample. The weights reflect the expected number of males in the given experience cell in any given year.

[7]See Murphy, Plant and Welch (1984).

[8]See Murphy, Plant and Welch (1984).

Appendix A: Insensitivity of Predicted Wage Profiles to Normalization

Suppose we decompose the wage matrix Y into

$$Y = (GA^{-1})(AW)$$

where A is a nonsingular matrix chosen to normalize the factors and factor loadings. Define

$$G^* = (GA^{-1})$$

and $$W^* = (AW)$$

so $$Y = G^*W^*.$$

We regress $W^{*\prime}$ on the matrix $X = P'G^*$ where P is the population data. Thus we fit

$$W^{*\prime} = XB$$

so

$$\hat{B} = (X'X)^{-1}X'W^{*\prime}$$

$$= (G^{*\prime}PP'G^*)G^{*\prime}\ PW^{*\prime}.$$

The predicted factor wages for a new population vector $\bar{p}$ are

$$\bar{w} = \bar{x}B$$

where $$\bar{x} = \bar{p}'G^*$$

and the predicted wage profiles are:

$$\bar{y} = G^*\bar{w}.$$

Substituting we get:

$$\bar{y} = G^*\hat{B}'\bar{x}'$$

$$= G^*W^*P'G^*(G^{*\prime}P\ P'G^*)^{-1}\ G^{*\prime}\ \bar{p}$$

$$= GA^{-1} AWP'GA^{-1} (A^{-1}G'P P'GA^{-1})^{-1} A^{-1'}G'\bar{p}$$

$$= GWP'G(G'PP'G)G'\bar{p}$$

which is independent of A, the normalization matrix.

References

Anderson, T.W. and Herman Rubin, "Statistical Inference in Factor Analysis," in Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, (Berkeley: UC Press, 1956).

Berger, Mark C., "The Effect of Cohort Size and the Earnings Growth of Young Workers," forthcoming in _Industrial and Labor Relations Review_, 1984.

Easterlin, R., M. Wachter, S. Wachter, "Demographic Influences on Economic Stability: The United States Experience," _Population and Development Review_, vol. 4, no. 1 (March 1978), 1-23.

__________, "What Will 1984 Be Like: Socioeconomic Implications of Recent Twists in Age Structure," _Demography_, vol. 5, no. 4 (November 1977), 397.

Freeman, Richard B., "The Effect of Demographic Factors on the Age-Earnings Profile in the U.S.," _Journal of Human Resources_, vol. 14, no. 3 (Summer 1979), 289-313.

__________, _The Overeducated American_, (New York: Academic Press, 1976).

Harmon, Harry H., _Modern Factor Analysis_, (Chicago: University of Chicago Press, 1967).

Joresky, K.G. and Sorbom, D., _Advances in Factor Analysis and Structural Equation Models_, (Cambridge, MA: Abt Associates, 1979).

Mincer, J., _Schooling Experience and Earnings_, (NBER, 1974).

Murphy, K., M. Plant and F. Welch, "Wage Dynamics: A Technical Analytic Framework," forthcoming, 1984.

Smith, J.P. and F. Welch, "The Overeducated American? A Review Article," in _Proceedings of National Academy of Education_, vol. 5, 1978.

Welch, F., "Effects of Cohort Size on Earnings: The Baby Boom Babies' Financial Bust," Journal of Political Economy, vol. 87, no. 5, pt. 2 (October 1979).

__________, "Linear Synthesis of Skill Distributions," Journal of Human Resources, Summer 1979.

__________, and Wm. Gould, "An Experience Imputation or an Imputation Experience," mimeograph, RAND Corp., 1976.

Whittle, P., "On Principal Components and Least Squares Methods of Factor Analysis," Skandinavisk Aktuarietidskript, (1952), 223-239.