U.S. Census Bureau Conference

The Direction of Fertility in the United States

Hosted by the Council of Professional Associations on Federal Statistics

October 2–3, 2001
Embassy Suites Hotel
Alexandria, Virginia

U.S. Census Bureau Conference

The Direction of Fertility in the United States

Hosted by the Council of Professional Associations on Federal Statistics

October 2–3, 2001
Embassy Suites Hotel
Alexandria, Virginia

Contents

Preface v

Tuesday Morning, Session I , Accuracy of the U.S. Census Bureau National Population Projections and Their Respective Components of Change 3

Tuesday Afternoon, Session II, U.S. Fertility in International Comparison: An Exploration to Aid Projections 51

Tuesday Afternoon, Session III, Should Fertility Intentions Inform Fertility Forecasts? 153

Wednesday Morning, Open Discussion 185

What Did We Learn? 187

Will America Remain an Outlier in Fertility? 187

Where Do We Go From Here? 190

Preface

Gregory Spencer
Chief, Population Projections Branch

The genesis of this conference was the Census Bureau's release of population projections for the United States in January 2000. These showed the possible future population in great detail for each year out to 2100: (a) by single year of age to 100 plus, (b) by sex, (c) by four race groups, (d) by Hispanic and non-Hispanic, and (e) by foreign-born and native. Innovations in this product included: (a) dynamic forecasts of migration, (b) use of the Lee-Carter approach to project life expectancy, (c) incorporation of both a "true" and a "census-level" of undocumented migration, and (d) creation of projections by nativity. The publicity for this document centered around the finding that the U.S. population would double during the new century, reaching almost 600 million in size. There was very little feedback from any group about the results or the methodology.

However, some experts did express concern about the fertility assumptions and methodology. Their primary issue was with the assumption that future fertility would remain somewhat higher than it is now. The overall total fertility rate was projected to rise from 2048 in 1999 to 2207 in 2025 and stay near that level through 2100. The essential reason for this projected increase was our decision to adopt birth expectations data as the primary predictor of future fertility. Moreover, we chose to assume that the "ultimate fertility level" of each race-ethnic group in 2150 would still be at 2100 (the so-called "replacement level" of fertility)—a higher fertility level than the U.S. total population had actually experienced since the early 1970's. This assumption was made based on our application of current forecasting models to the most recent birth expectations data plus the trend oddities, which we found to occur if we attempted to converge a number of race/ethnic groups to any lower fertility level or converge their fertility any more rapidly. Finally, the above-replacement fertility assumption also received considerable attention because the United Nations had just released projections for each of the world's developed nations in which virtually every developed nation was assumed to always have a total fertility rate under 1800.

Because of these general concerns that our fertility assumptions for America were too high, and because we could find no methodology that would yield significantly different results, we thought this was an excellent time to host a conference devoted to this topic. As summarized in the invitation letter, the charge to the conference attendees was the following:

"While there have been dramatic changes in fertility in the rest of the world, American fertility has remained remarkably constant since the early 1970's. Our current projections mirror this stability and are inconsistent with the lower assumptions used in other projections. From this meeting we especially hope to learn about the contribution of compositional aspects such as race, ethnicity, and nativity. We also wish to better understand the nature of this apparently anomalous stability in American fertility and to monitor any nascent signs of pending changes in fertility in the United States."

Three separate papers were commissioned and two discussants were assigned to provide formal comments. Every attendee received the commissioned papers several weeks in advance of the conference. A professional scribe was hired to take down all the comments of other participants.

This report contains all of this information and attempts to accurately convey the opinions of all.

The commissioned papers and formal discussions are shown exactly as they were presented. The comments from the audience were summarized by Gregory Spencer from the scribe's notes. He apologizes for any misstatements or misinterpretations ascribed to the participants.

The Direction of Fertility in the United States

October 2–3, 2001
Embassy Suites Hotel
Alexandria, Virginia

TUESDAY MORNING
OCTOBER 2, 2001

7:45-8:30	**Registration and Continental Breakfast**	
8:30-8:45	Welcome	John Long, Chief, Population Division U.S. Census Bureau
8:45-9:00	Overview	Signe Wetrogan, Assistant Division Chief, Population Division U.S. Census Bureau
9:00-9:40	Session I: The Accuracy of Census Bureau Fertility Projections	
	Chair	Dennis Ahlburg, University of Minnesota
	Author	Tammany Mulder, U.S. Census Bureau
9:40-10:20	Discussant	Hania Zlotnik, United Nations
	Discussant	Jeffrey Passel, Urban Institute
10:20-10:35	**Break**	
10:35-11:15	General Discussion of Session I	
11:15-1:00	**Lunch**	

TUESDAY AFTERNOON
OCTOBER 2, 2001

1:00-1:15	Remarks	Nancy Gordon Associate Director for Demographic Programs, U.S. Census Bureau
1:15-1:55	Session II: Why Is American Fertility So High? (in the Context of the Developed World)	
	Chair	Joseph Chamie, United Nations
	Author	Tomas Frejka, Independent International Consultant
	Author	W. Ward Kingkade, U.S. Census Bureau
1:55-2:35	Discussant	Robert Schoen, Pennsylvania State University
	Discussant	Josh Goldstein, Princeton University
2:35-3:15	General Discussion of Session II	
3:15-3:30	**Break**	

Tuesday Afternoon, continued

3:30-4:10	Session III: The Use/Usefulness of Birth Expectations in Fertility Projections
	Chair. . . . . Martin O'Connell, U.S. Census Bureau
	Author . . . . S. Philip Morgan, Duke University
4:10-4:50	Discussant . . . . Nico Keilman, University of Oslo
	Discussant . . . . George Masnick, Harvard University Joint Center for Housing Studies
4:50-5:15	General Discussion of Session III
5:15	**Adjourn**

WEDNESDAY MORNING OCTOBER 3, 2001

8:00-8:45	**Continental Breakfast**
8:45-9:15	Open Discussion: What Did We Learn?
	. . . . Signe Wetrogan, Chair, U.S. Census Bureau
9:15-11:05	Open Discussion: Will America Remain an Outlier in Fertility?
	. . . . Frank Bean, Chair, University of California, Irvine
11:05-11:15	**Break**
11:15-11:45	Open Discussion: Where Do We Go From Here?
	. . . . Fred Hollmann, Chair, U.S. Census Bureau
11:45	**Adjourn**

Fertility Conference Attendees (Alphabetical)

Dennis Ahlburg	*University of Minnesota*
Adelamar Alcantara	*University of New Mexico*
Frank Bean	*University of California, Irvine*
Suzanne Bianchi	*University of Maryland*
John Bongaarts	*Population Council*
Joseph Chamie	*United Nations*
David Coleman	*Oxford University*
Lincoln Day	*Washington, D.C.*
Jane De Lung	*Population Resource Center*
Tomas Frejka	*Sanibel, Fla.*
Howard Fullerton	*Bureau of Labor Statistics*
M.V. George	*Statistics Canada*
Josh Goldstein	*Princeton University*
Nancy Gordon	*Census Bureau*
Carl Haub	*Population Reference Bureau*
Mary Heim	*California Department of Finance*
Stanley Henshaw	*Alan Guttmacher Institute*
Kenneth Hodges	*Claritas*
Frederick Hollmann	*Census Bureau*
Peter Johnson	*Census Bureau*
Nico Keilman	*University of Oslo*
W. Ward Kingkade	*Census Bureau*
John Long	*Census Bureau*
George Masnick	*Harvard University*
Peter McDonald	*Australian National University*
S. Philip Morgan	*Duke University*
Tammany Mulder	*Census Bureau*
Steve Murdock	*Texas A&M University*
Michael Murphy	*London School of Economics*
Martin O'Connell	*Census Bureau*
Jeffrey Passel	*Urban Institute*
Ann Powell	*Census Bureau*
Karyl Pullen	*Census Bureau*
K.V. Rao	*Bowling Green State University*
Robert Schoen	*Pennsylvania State University*
Mary Jane Slagle	*Census Bureau*
Edward Spar	*COPAFS*
Gregory Spencer	*Census Bureau*
Dirk Van de Kaa	*University of Amsterdam*
Stephanie Ventura	*National Center for Health Statistics*
Ching-li Wang	*Census Bureau*
James Weed	*National Center for Health Statistics*
Signe Wetrogan	*Census Bureau*
Hania Zlotnik	*United Nations*

Proceedings of the Conference

Introductory Remarks

John Long
Chief, Population Division

The Census Bureau hasn't had a conference like this in 30 years. We used to have them because fertility was changing so much. This time we are having it because fertility is not changing at all.

We at Census have been looking at a number of innovations in our projections dealing with immigration, mortality, multiple sets of projections, and stochastic forecasts. But we are concerned that U.S. stability in fertility may not continue, particularly in light of what has happened in other developed countries.

Given all the attention being given to immigration and Census 2000, fertility analysis might have gotten lost if we had not had this conference. I am very happy that we are having it.

A number of you were at the IUSSP conference in Brazil earlier this fall. One of the key issues discussed there was about the future course of European fertility. The mainly European speakers thought fertility would remain low. But their mostly non-European audience thought it would go up some. We are not looking for your vote here, but hope to get your opinions as to what factors are important in possible changes in U.S. fertility, what are the early warning signs of change, etc. I look forward to your participation and the results of this conference.

Overview

Signe Wetrogan
Assistant Division Chief for
Population Estimates and Projections

As part of our mission in the Population Estimates and Projections area, we are responsible for developing and preparing projections of the population of the United States by various demographic characteristics—namely age, sex, race, and Hispanic origin. Our goal is to prepare these projections bi-annually. The last set of national population projections were prepared during 1999 and released in January 2000. At that time, we developed projections out to year 2100.

In preparing these official national projections, we rely upon the basic demographic approach of the cohort component technique. As such, we must make assumptions about the components of population change at the national level—the components of birth, death, and international migration. The projections begin with the latest official census or national population estimates, which are projected forward using the assumptions about fertility, mortality, and international migration. Although, in the past, we have not prepared statistical levels of uncertainty, we do introduce alternative assumptions about the levels of future fertility, mortality, and international migration.

Because the projections are developed by race and Hispanic origin, we must develop these assumptions separately for each of the race and Hispanic groups.

In the most recent set of national projections, we followed the findings of the research community in setting all of the assumptions. For mortality, we used the approach developed by Lee and Carter to project ultimate mortality levels. The overall outcome was to assume that life expectancy for males would increase from 74.0 in 1999 to 77.6 in 2025, and 81.2 in 2050, and finally reach 88.0 in 2100. Female life expectancy would increase from 79.8 in 1999 to 83.6 in 2025, 86.7 in 2050, and 92.3 in 2100.

We projected dynamic rates of international migration. The rates would remain fairly stable and result in levels of international migration of 900,000 to 1 million annually.

In setting the fertility levels we were in a quandary. Analysis of fertility trends revealed a relatively stable pattern for the past decade. However, compared to the other developed countries, the fertility rates of the United States appeared high. Drawing upon the analysis and the results of recent birth expectations data, the set of projections prepared in 1999 assumed that the total fertility rates for the United States would remain fairly stable.

It is now fall 2001, we have the Census 2000 behind us, and we are preparing for the next set of national population projections that we plan to develop and release soon. The approaches and even the levels of mortality and international migration of the most recent projections still appear quite reasonable.

However, what do we do about projecting future levels of U.S. fertility? Why does U.S. fertility appear to be "an outlier" among developed countries? Although recent analysis still reveals relatively stable patterns, will U.S. fertility levels in the future be more in line with those of the other developed countries?

We have invited you to help us explore this topic and, hopefully by Wednesday noon, provide us with some insight and thoughts on the direction of American fertility in the future.

To help guide our discussion, we have commissioned three papers—one addressing the accuracy of recent Census Bureau projections, another examining the trends and differentials in recent U.S. fertility, and a third paper examining the utility of birth expectations data in preparing forecasts of fertility.

We have also invited two discussants for each of the papers. Following the formal discussion, there is ample time for you to provide your ideas and comments about the papers and issues raised in them.

On Wednesday, we plan to begin with a summary of the previous day and work toward hearing your thoughts on the direction of U.S. fertility over the next 50 years.

Session Introduction

Dennis Ahlburg
University of Minnesota

I am pleased to introduce Tammany Mulder, who will present the first paper. Hania Zlotnik and Jeff Passel will then discuss the paper.

Tuesday Morning

Session I

Accuracy of the U.S. Census Bureau National Population Projections and Their Respective Components of Change

Tammany Mulder
(U.S. Census Bureau)

CONTENTS

Abstract

Introduction

Research Design and Methods

Complexities in Assessing the Accuracy of Forecasts

Choosing Among Multiple Forecast Series

Measurement of Forecast Error at Multiple Levels

Forecast Error Patterns

Explanation of Indicators

Comparison of the Census Bureau Forecast Models with a Naive Model

Potential Biases Present Within the Estimate and Forecast Series

Results

Total Population Growth Rate Forecasts

Overall Accuracy and Duration-Specific Forecast Error of the Population Growth Rate Forecasts

Comparison of Growth Rate Forecast Models

Summary of Forecast Error for Growth Rates

Components of Change Forecasts

Fertility Forecasts Error Analysis

Overall Accuracy of Fertility Forecasts

Duration-Specific Forecast Error for Fertility

Comparison of Fertility Forecast Models

Summary of Forecast Error for Fertility

Mortality Forecasts Error Analysis

Overall Accuracy of Mortality Forecasts

Duration-Specific Forecast Error for Mortality

Comparison of Mortality Forecast Models

Summary of Forecast Error for Mortality

Net Immigration Forecasts Error Analysis

Overall Accuracy of Net Immigration Forecasts

Duration-Specific Forecast Error for Net Immigration

Comparison of Net Immigration Forecast Models

Summary of Forecast Error for Net Immigration

Discussion of Results

Conclusion

References

Tables

Figures

Discussion

Accuracy of the U.S. Census Bureau National Population Projections and Their Respective Components of Change

Tammany Mulder
U.S. Census Bureau

Abstract

During the 1900's, knowledge of population trends and their future repercussions for the size and distribution of the population became increasingly important as the U.S. experienced major shifts in fertility and net immigration. Population forecasts produced by the Census Bureau are used widely, informing researchers, planners, legislators, and many others, on the future course of population change. Because forecasts are subject to inherent uncertainty, as they are based on a compilation of reasonable assumptions for the components of population change, it is essential to educate customers as to the amount of uncertainty within the forecasts for the population and the components of population change. To date, the Census Bureau has not published a comprehensive analysis of the accuracy of their forecasts. The aim of this research is to address this gap and systematically evaluate the accuracy of the existing Census Bureau forecasts both in terms of their ability to predict the national population as well as individual components of change.

Overall, the Census Bureau has greatly improved the level of accuracy found within its forecasts. Recent forecasts produced in the 1990's have minimized the inherent uncertainty and provide a reliable product for consumers in the short term. Improvement in the forecast reliability is, in all likelihood, the result of the stabilization of the components of population change. This study reveals that forecasters failed to foresee turning points in population trends, resulting in erroneous forecasts, particularly for fertility and net immigration. The inadequate base data used for certain series severely reduced accuracy upon beginning the forecast. Consequently, the forecasts maintained higher levels of error throughout the forecast period. In addition, the assumptions formulated by the Bureau were often outperformed by simple assumptions of constancy. The research presented here represents a contribution to the discussion of population forecasting accuracy for the United States; however, additional research is needed.

Introduction

Population projections are computations of future population size and characteristics based on separating the total population into its basic components of fertility, mortality, and migration and estimating the probable trends in each component under various assumptions (Srinivasan, 1998). National projections give planners, legislators, policy makers, and researchers, among others, a glimpse of possible future demographic trends for the population and the forces acting to produce population change. The U.S. Census Bureau, in collaboration with Thompson and Whelpton of the Scripps Foundation, began producing population projections and estimates for the national population in the 1940s. Following the first collaborative publication, the Census Bureau independently produced approximately eighteen primary forecasts for the national population (Whelpton, Eldridge, and Siegel, 1947). Because projections are simply a compilation of reasonable assumptions as to what will happen to the current population in future years, the accuracy of forecasts will depend on the validity of the assumptions and the accuracy with which the assumptions are quantified. Correspondingly, it is critical for the consumers of population projections to recognize the level of uncertainty found within population forecasts both in terms of their overall accuracy as well as in terms of the specific components of population change.

To date, the Census Bureau has not published a comprehensive analysis of the accuracy of their forecasts, which means customers depend on the expertise of the demographers producing the product. Long (1987), Stoto (1983), and Ascher (1978), each evaluated the forecast accuracy for the growth rate of the total population, while Ahlburg (1982) evaluated the accuracy of U.S. Census Bureau forecasted total births. However, these analyses have not been updated since their original publication. The aim of this research is to address this gap and systematically evaluate the accuracy of the existing Census Bureau forecasts both in terms of their ability to predict the national population as well as individual components of change.

Projections are used for planning the delivery of various services, such as education, health facilities, employment, water and utilities, communications, transportation, and housing stock among many others, the distribution of federal and state resources, and to assist producers and sellers of various goods and services to predict future markets for their products. Moreover, in addition to understanding the overall size of the national population in the future, planners and policy-makers have an equally important stake in getting an accurate reading of the age and sex composition of the future population (Srinivasan, 1998). An evaluation of the accuracy of the national population forecasts and their components of change, will allow consumers to become more discriminate users of population forecasts. In addition, the research allows forecasters greater insight into how to improve their ability to forecast and where potential problems or biases exist.

The present paper evaluates the accuracy of Census Bureau population forecasts using an ex-post facto approach. That is, the performance of a forecast is evaluated relative to what was observed, which is operationalized here as intercensal estimates from 1947 to 1989, and the post-censal estimates from 1990 to 1999, produced by the Census Bureau (1990, 1993, 1995, 1999, 2000a). In addition, the present study evaluates the assumptions used as input variables in the cohort component method. Specifically, this research will attempt to answer two research questions. First, how accurately did the Census Bureau forecast the total population and its respective components of change? Second, did the forecasts for the population and components produced by the Census Bureau perform more accurately than a naive model assuming constant rates?

Given that this paper represents the first effort to evaluate the accuracy of U.S. population projections on a comprehensive scale, few precedents exist regarding how to properly conduct the assessment. The next section details the complexities involved in assessing the accumulated national projections to date. This is followed by a section on the specific research design used to address these complexities in this paper. Next, the paper provides a discussion of the results of the accuracy assessment. This is broken into two subsections: population growth rate forecasts and components of change forecasts. The paper then presents the results, closing with a discussion and conclusions.

For the purposes of this research, the following terminology, which is consistent with language used among demographers and adapted from Smith and Sincich (1991), will be used to describe forecasts throughout the text:

Base year:	The most recent estimate used to begin the forecast;
Target year:	The designated point[1] (year) the forecast reaches;
Forecast period:	The interval between first forecast year after the base year and target year;
Forecast error:	The difference between the observed and the forecast population at a designated point in forecast period.

When discussing population projections, demographers often specify the difference between a “forecast” and a “projection.” A projection generally represents possible population trends, while forecasts are produced to represent real population trends. In order to analyze the accuracy of the projections, the “preferred” middle series is used (U.S. Census Bureau, 2000b). In other words, this is the series the Bureau feels is most likely to take place, typifying a forecast. Furthermore, the object here is to analyze “forecast error,” meaning the difference between forecast results and estimates.

[1] Throughout the text, “point” refers to a finite time interval within the forecast period.

Research Design and Methods

Complexities in Assessing the Accuracy of Forecasts

Table 1 summarizes the base years, the forecast periods, the authors, and the type of series produced in each Census Bureau forecast product as of 1947. To assess the accuracy of this accumulated body of forecasts is inherently complex and requires a multi-pronged approach. Forecast error:

1) can be assessed for multiple forecast series;
2) can be measured at multiple levels: individual years, periods, multiple series;
3) evaluation can be approached from different perspectives;
4) does not have an indicator recognized as the most reliable and valid among forecasters;
5) can be assessed for the population and the components of change;
6) can be calculated for alternate "naive" models with simplified assumptions providing a benchmark to compare the Census Bureau forecast error;
7) may be skewed by existing biases in the data.

First, for any given national forecasts generated by the Census Bureau, multiple series are produced to represent the potential uncertainty experienced when forecasting the future population. Generally, a middle or "preferred" series forecast is produced with several alternate series based on differing assumptions for the components of change. Second, measurement of error can be calculated at three different levels: (1) forecast error by *individual year* of forecast; (2) averages of forecast error within *intervals* of a forecast period; and (3) averages of forecast errors across multiple series for *specific points in the forecast period*. Consequently, it is possible to examine forecast error resulting within individual series (defined as the error occurring within a specific series forecast period) as well as across multiple forecast series. Third, accuracy evaluations for individual and multiple series are approached from two perspectives: (1) the overall degree of accuracy for the forecasts; and (2) the pattern of error experienced at different points in the forecast. This separation permits analysis of how well the forecaster performed in general, which components of change potentially contributed to the error, and how much error may be attributable to the model upon which the forecasts were built. A fourth complexity inherent in evaluating the accuracy of the national population forecasts is that there is no consensus among forecasters as to the best indicator of forecast error to use. Fifth, because population change is driven by the trends for three components—births, deaths, and migration—forecasts of future population size and growth are built upon assumptions about the annual rate of population growth, as well as trends in the individual components of population change over time. Consequently, the accuracy of any forecast can be assessed according to its ability to forecast the population as well as forecasting the individual components of population change. Sixth, because forecasts are created using various assumptions, the forecasts can be compared to simplified or "naive" models with assumptions of no change in future trends, providing a benchmark to compare Census Bureau forecast error. Lastly, forecast error may be skewed by biases present in the population estimates and forecasts and the individual components of change.

The Research Methods section provides the details of how these levels of complexity will be addressed in the present paper.

Choosing Among Multiple Forecast Series

In the recent past, the Census Bureau produced a middle series forecast and several alternate series based on differing assumptions for the components of change. Because the Census Bureau refers to the middle series as the "preferred series," and consumers commonly use this series, it is used hereafter for analytic purposes (U.S. Census Bureau, 2000b). The final column of Table 1 specifies the series used in this paper. For ease of discussion, each series will be identified by its respective *base year* (column 1). To evaluate the accuracy of the forecasts for the total population, seventeen forecasts were analyzed with base years ranging from 1947 to 1994 (U.S. Census Bureau, 1949 to 1996; Whelpton, Eldridge, and Siegel, 1947). Twelve series for the components of change are available from 1964 to 1994 (U.S. Census Bureau, 1964 to 1996).

Error for the total population is measured by its annual percentage rate of change, or annual growth rate, which is calculated using the exponential formula shown in Appendix A. Measurement of error for population projections can be influenced by the size of the projected population and the forecast length (Stoto, 1983). Use of the growth rate for the total population and the rate for the components of change removes any effects of the potential error from population size or the length of the forecast period. Evaluation of forecast accuracy for the growth rate of the total population builds on existing research by Long (1987), Stoto (1983), and Ascher (1978). Comparison of total births follows existing research by Ahlburg (1982).

Ex-post facto evaluation compares the forecast results with the historical population that was actually observed. Therefore, to evaluate the performance of past forecasts, each series is compared with intercensal (1947 to 1989) or postcensal (1990 to 1999) national estimates for the total population from 1947 to 1999. The forecast components of change and the corresponding crude rates are compared with the components produced as a part of the Census Bureau national estimates and vital statistics from the National Center for Health Statistics from 1963 to 1999 (National Center for Health Statistics, 1993; U.S. Census Bureau, 1990, 2000a; Ventura, et. al., 1999, 2000). Both the estimated and the forecast population growth rates are calculated for annual intervals ending on June 30, while the components of change are summed for calendar years. Note that the lengths of forecasts vary, ranging from 7 to 101 years, and that the forecast period of subsequent forecasts always overlap to some extent with that of prior forecasts. Because few forecast series for the components of change and the total population are available in a consistent time series beyond 20 years in length, this analysis does not extend past the 20-year period.

Because forecasts and the input assumptions are created with several characteristics, this provides greater detail for analysis, including variables such as age, sex, race, and Hispanic origin. Additional detail, however, may either not be available in a consistent time series, or is not categorized in a consistent manner across products since 1947. Therefore, this analysis pertains only to the total number and crude rates for the total population and the components of change.

Measurement of Forecast Error at Multiple Levels

A complicating factor in evaluating forecast error is that it can be calculated at different levels. It is possible to analyze an *individual point* in the forecast, the *individual series* to determine the error for specific products, as well as the error for *multiple forecast series* (one series per product) averaged to assess the aggregation of error generally associated with the Census Bureau forecasts. The schematic diagram shown in Model 1 depicts how these types of accuracy assessments are made and how they compare to one another. In each case, *forecast error terms*—the difference between the observed and the forecast population—are used.

First, consider the assessment of the level of error for the forecast error term using a series with the base year 1947 (S_1) (see Model 1). The years analyzed in this forecast period cover 1948 (S_1+1) to 1955 (S_1+8). Notice that for each year in this forecast period, a forecast error term is calculated within each cell of column (2) as the difference between the forecast and the observed values, both in terms of the population and the components of change.[2] Each cell conveys the error that occurred at a specific point in the forecast period. In this particular instance, the forecast period contained 8 years.

The second level of interest, the *individual* series, represents the *average* of the error associated with any specific interval of interest, for example, over the first 5 years of the forecast, the first 10 years, etc. Referring to Model 1 (column 2, final row), using the same 1947 based series, the gray-filled cells of column (2) show how in the case of the 1- to 5-year interval, the five forecast error terms are summed and divided by five. The same logic applies to the other targeted intervals.

The third level of accuracy assessment relates to aggregating the past forecast error to reflect on experience in a cumulative manner. In this case, *multiple series*, the middle series from each product, are used for the input. Specifically, for any given year in a forecast period (e.g., the 1st, 2nd, 3rd,…, 20th), forecast error terms are averaged across each product for the specific time elapsed from the base dates of the series. An example of the formula for assessing the accuracy of the forecasts for their first year (point) is depicted in Model 1 as the bold-framed cells. The forecast error terms for the first year in each series are summed, then divided by the number of series included (final column). Again, this same logic extends to each of the other period target years in the series.

Forecast Error Patterns

Accuracy evaluation can be approached from two perspectives. Until now, the focus has been on evaluating *overall* forecast error. These evaluations relate strictly to the general performance of the forecast(s). The second, and more specific approach in performing a comprehensive assessment of forecast accuracy is that in addition to overall series error, there may also be patterns of error across time. In other words, how well did the forecasts perform throughout the length of forecast period and does a particular pat-

[2] For forecasts produced before 1986, the components of change were published for the mid-year population (July 1). Whereas, rates for products published in 1986 and afterward were calculated for calendar years. Therefore, for purposes of this research the components of change and their respective rates were recalculated to represent the calendar year event.

tern exist? Smith and Sincich (1991: pg. 261) found that "…there is a linear or nearly linear relationship between forecast accuracy and the length of the forecast horizon,…" Uncovering these patterns helps to decipher the relationship between the error attributed to the different components of change, as well as if they demonstrate different patterns of change throughout the forecast period. In order to assess the patterns of error throughout the forecast period, a supplemental analysis is presented for both individual and multiple series. Hereafter, *duration-specific forecast error* references the observation of patterns of error. Indicators used to measure overall error also measure the duration-specific forecast error for both the individual and multiple series.

Explanation of Indicators

Statistics used to measure the accuracy of forecasting methodology and assumptions originated from economic forecasting analysis. Demographers and statisticians apply these statistics to measure the accuracy of population forecasts at the national and subnational level. Researchers have not reached a consensus as to which indicators are most indicative of the accuracy of national population forecasts (Ahlburg, 1992; Armstrong and Collopy, 1992). Consequently, several statistics are often used to afford analysis from different perspectives. Some of the most common, and those used in this report, include the percent error, the mean percent error, the mean absolute percent error, the median absolute percent error, and the root mean squared error. The equations of the aforementioned statistics are presented in Appendix A.

The mean absolute percent error (MAPE) also calculates the difference between actual and forecast values, but is the average of the absolute value (irrespective of whether the error is positive or negative) of the error terms. Positive and negative errors therefore reinforce each other, rather than cancel each other. Each forecast error term is weighted equally. The MAPE is commonly used by forecasters because of the ease of calculation, analysis, and reliability (Tayman and Swanson, 1996). In addition, Swanson, Tayman, and Barr (2000: pg. 193) argue that the MAPE possesses "…highly desirable statistical and mathematical properties." The MAPE, however, is an arithmetic mean with an asymmetrical distribution and is prone to being influenced by outlier values, thereby tending to underestimate accuracy. Consequently, the aforementioned authors argue that in reference to evaluating the accuracy of sub-national estimates, the MAPE may lack validity. Contrary to the arithmetic mean, the median is not influenced by outlier values within the distribution. Consequently, the median absolute percent error (MdAPE) was calculated as a supplementary statistic and is presented in the data.

Another commonly used statistic to measure the accuracy of population forecasts is the root mean squared error (RMSE). Forecast error terms are squared and converted to a square root and averaged, providing a statistic in the same unit of analysis as the original variable. In comparison to the MAPE, the RMSE gives additional weight to larger error terms because of squaring. Therefore, as an arithmetic mean, outliers influence both the RMSE and the MAPE. The RMSE gives even greater weight to those series experiencing large error values. The root mean squared percent error (RMSPE) provides the same properties as the RMSE, but is expressed as a percent.

These evaluative statistics apply to the individual and the multiple series analysis for both the overall forecast error and the duration-specific forecast error. To assess overall error, the PE is used to measure the forecast error that occurred at specified points in the forecast period (1, 5, 10, 15, 20 years). The MPE and the remainder of the statistics present the average within an individual series forecast period at specified intervals (5-, 10-, 15-, and 20-year intervals). These indicators also measure the average across multiple series at designated points of the forecast period (1st, 5th, 15th, and 20th year from the base) as opposed to within series averages. Duration-specific forecast error is measured using the same indicators; however, for multiple series each indicator is analyzed annually (for each point) as opposed to designated points.

Comparison of the Census Bureau Forecast Models with a Naive Model

Each Census Bureau forecast is based on a complex set of assumptions about how patterns of fertility, mortality and migration will behave over time. In order to understand the uncertainty related to these assumptions, each component of population change, as well as the population growth rate, is compared with a "naive" model. Comparing the forecasts with a simplified naive model assuming no change in future trends provides a benchmark to evaluate and compare the error experienced by the forecast model (Keyfitz, 1977: pg. 230). It provides additional insight into the assumptions made both in the long and short term of the forecast period. Lastly, it contributes to the knowledge of the quality of base data used for the forecast.

The naive model is created by assuming the annual growth rate for the total population or the crude rates for the individual components remained constant as of the base year or "jump-off" population for the forecasts. For example, annual growth rates for the forecasts produced from 1967 to 1990 in P25–381 are compared with the constant annual growth rate for 1966, the designated population base of that forecast. The naive model for number of deaths, however, cannot be simply held constant, as this would not be representative of actual trends. The naive numbers of deaths were recalculated for each series based on the associated forecast population and the constant crude death rate. The RMSE is also calculated for the naive model to determine whether the assumptions made within the forecast performed better than simply forecasting a constant. Therefore, if the value of the forecast RMSE is smaller than the naive RMSE, the forecast assumptions or forecast growth rates outperformed the naive model.

Potential Biases Present Within the Estimate and Forecast Series

An accurate assessment of forecast error depends upon the characteristics and the quality of both the estimates and forecast series for the population and the components of change. Therefore, it is important to discuss discrepancies and irregularities found between and within data sources.

The postcensal national population estimates are derived from the most recent national census. This complete enumeration often contains error relating to such issues as under enumeration and data problems in the estimation of population change. Following the census, the postcensal estimates are adjusted for the *error of closure*. The 1980 census results determined that the 1970s population estimates underestimated the total population by approximately 5 million people in 1980. Consequently, the 1970 estimates were adjusted for the *error of closure* by adding approximately 1/2 million people, compounding each year. Therefore, the base populations used for the 1972, 1974, and 1976 series forecasts were off by the respective adjustments in the first forecast period year. For 1972, the forecast erred by 1 million or .54 of a percentage point, for 1974 2.0 million or .95 of a percentage point, for 1976 3.0 million or 1.39 percent. The forecast growth rates were compared with growth rates revised after the forecast production.

Identification of a single middle series permits the comparison of error across products and the error experienced by each individual series. Therefore, in addition to analyzing the forecast error for each series, the error is calculated for the combination of series at specific points in the forecast period. Note that in Table 1 several products produced before 1974 failed to designate a specific middle series. Alternatively, four series were created based on differing assumptions ranging from lowest to highest values, which are not equidistant in value. In order to create a middle series for evaluation, we computed the average of the two series between the lowest and highest valued alternatives. This was done for the total population, births, and deaths, and is specified in Table 1, column 6. Among the products included in this research, eight products with base dates between 1953 and 1972 did not designate a middle series. Five series, produced between 1963 and 1972, are averaged for the components of change.

The universe for net immigration changed throughout the history of Census Bureau forecasts. For most of the products, net immigration referred to net civilian immigration with the Armed Forces Overseas (AFO) population as part of the base population. The Census Bureau changed the definition of net civilian immigrants to net migration to the U.S. and began treating the AFO as a separate universe by not including it within the base population. The national estimates and national forecasts used this methodology beginning in the 1990s (U.S. Census Bureau, 1993). Therefore, to maintain consistency, the AFO population was added to each total population estimate and forecast. For the total population forecasts, the AFO experienced in the base population were simply held constant throughout the forecast period.

Before the 1986 forecast series, the assumed number of immigrants for the national forecasts did not include undocumented immigrants nor the number of emigrants from the U.S. Following the 1980 census, the national estimates included the number of net undocumented immigrants and emigrants (U.S. Census Bureau, 1990). Discussed later, undocumented immigration began to increase in the 1970s. Consequently, the observed number of immigrants net of emigration and the corresponding rates for the observed estimates from 1970 to 1979 were adjusted upward by 76,000 for each elapsed year after 1970, to include the movement of these groups. The forecast series produced before 1986 did not include these flows in its universe. Therefore, for this analysis, the series produced from 1963 to 1983 are compared with the adjusted observed number (and rates) of immigrants net of emigration, hereafter referred to simply as immigrants and the net immigration rate. In addition, the naive model used the adjusted observed estimates to create forecasts. Consequently, Census Bureau net immigration forecasts for 1970, 1972, 1974, 1976, and 1982, are being compared to a naive model based on adjusted observed data mentioned above.

Results

Total Population Growth Rate Forecasts

The U.S. population tripled between 1900 and 1999 as the nation maintained growth rates ranging between approximately a high of 2.0 percent and a low of .6 of a percentage point, with current rates leveling off near .9 percentage points (U.S. Census Bureau, 1999). Graph 1 presents the annual growth rate for the total population from 1947 to 1999, the respective years analyzed for this research. Analyses of how well the Census Bureau forecast the nation's growth trends are first discussed for the multiple series followed by a discussion of each individual series. As mentioned earlier, accuracy assessment is approached from two perspectives: (1) in terms of overall error in the series; and (2) in terms of duration-specific forecast error. Overall error is analyzed for the direction of error (the tendency of the forecast growth rate to generally over- or underestimate the observed growth rate, which is measured using the PE and MPE) and the magnitude of error (which is measured with the MAPE and RMSE). The duration-specific forecast error analyzes the pattern of the error throughout the forecast. Lastly, a comparison of the naïve and the forecast model will be made using the RMSE results.

Because previous authors have examined the historical performance of the forecast population growth rate, the following discussion will remain brief (Ascher, 1978; Stoto, 1983; Long, 1987). This research improves and extends existing research by: (1) evaluating forecasts that are more recent; (2) utilizing more recent national estimates and vital statistics data for the observed series; (3) comparing individual and multiple series results; (4) increasing the sample size for multiple series error statistics; and (5) calculating several statistics to compare results.

Overall Accuracy and Duration-Specific Forecast Error of the Population Growth Rate Forecasts

The multiple series and individual series statistics presented in Table 2 allow for an assessment of whether the total growth rate is generally over- or underestimated by the Census Bureau. As shown in the final column of row (1), the multiple series MPEs for the annual growth rate indicate that the Census Bureau generally underestimated growth rates within the first five years (MPE= –3.8 at the fifth year). In contrast, beyond the five-year period, on average the growth rates were overestimated, as indicated by positive MPEs.

Table 3 presents the percent error occurring at designated points of the forecast period (1st, 5th, 10th, 15th, and 20th years). The wide variations between the MPE, MAPE, and MdAPE (Table 2), and the wide range between individual PEs, within each of the four target forecast periods, indicates that potential outliers influence the multiple error statistics. The PEs range between –26.5 percent (1974) and 6.4 percent (1966) at the first year and from –48.6 (1947) and 29.2 (1963) at the fifth year (n=17). This implies that the multiple error statistics are not representative of the general performance for the growth rates forecast between 1947 and 1999. Within the more recent forecast publications, the Census Bureau includes multiple series RMSE results for the growth rate of the total population as a way of addressing the uncertainty of their forecasts (U.S. Census Bureau, 1996). The RMSE results question the validity of such multiple series growth rate statistics and underscores the need to examine individual series.

An evaluation of the statistics for the individual series reveals a more complex trend of over- and underestimation. Forecasts produced in 1955 and earlier consistently underestimated growth rates. This trend reversed for series produced between 1957 and 1972. Following 1972, the growth rate for each series is again underestimated. Of the seven forecast series produced between 1974 and 1994, three series resulted in small overestimates in the first 5 years (MPE=3.9, 1.9, and 7.6 percent respectively). Otherwise, within and beyond the 5-year period, growth rates for those series were underestimated.

For series with base years between 1947 and 1957, the accuracy improved from series to series within the first 5 years. Series produced in 1947 and 1949 have the largest percent errors at the fifth and tenth year period, with 5 year MAPEs of 31.2 percent and 18.5 percent, respectively (Table 2). Series produced in 1953, 1955, and 1957 improved in overall accuracy within 10 years, averaging 11.5 percent for 1955 and 15.6 percent for 1953. Series 1957 experienced the lowest MAPE of 2.0 percent within the first 5 years for all series. The accuracy decreased for this series throughout the remainder of the forecast period.

Forecasts for 1963, 1966, 1969, and 1970 did not generally improve in accuracy over the 1953, 1955, and 1957 series in the first 5 years. The 1972 series showed an improvement, but then the 1974 and 1976 series showed more error. Series 1974 and 1976 increased in error within the first 5 years with MAPEs of 20.8 and 21.5 percent, respectively, from the improved 1972 MAPE of 4.1 percent. The increase in error and the pattern of underestimation for the 1974 and 1976 series may be the result of the error of closure adjustment made to the intercensal estimates mentioned above. When not allowing for the error of closure, Long (1987) calculated lower RMSEs for 1974 and 1976. Within the first 4 years of the forecast period, Long (1987) calculated a RMSE of .09 percentage points for 1974 and .18 percentage points for 1976 (Table 1). In comparison, when accounting for the error of closure, Long obtained RMSEs similar to the results presented in Table 2 (Table 1A).

The accuracy of forecast growth rates improved after the 1970s within the first 5 years. The MAPEs ranged between a low of 2.5 percent (1991) and a high of 9.9 percent (1986). Forecasts produced in 1982, 1991, and 1994, for the first 5 years improve in accuracy with MAPE values below 4 percent. Although series 1986 and 1992 maintain higher 5-year MAPEs of 9.9 and 7.6 percent than those produced after the 1970s, these series still maintain lower averages than most previous series.

An analysis of the percent error in Table 3 and the statistics in Table 2 reveal that the pattern of error, the duration-specific forecast error, throughout the forecast periods did not increase linearly for each series. To the contrary, certain series both under- and overestimate the growth rate throughout the period. In addition, the magnitude of error fluctuated throughout certain series. For example, the PE changes direction throughout the forecast period of 20 years for 11 of the 17 series. In addition, the error does not generally increase in size throughout the forecast period; i.e., as the growth rate is forecast for longer time intervals, the error does not generally increase. Both the percent error statistics and the average error statistics for the individual series demonstrate this trend. The MAPEs and MdAPEs for series 1953, 1974, and 1976, among others, both increase and decrease beyond the 5-year period.

Comparison of Growth Rate Forecast Models

Table 2 shows the results for the naive and Census Bureau forecast model RMSE. At the fifth year period, on average the naive model outperformed the forecast model. The RMSE of .30 percentage points at the fifth year is larger than a RMSE of .18 percentage points for the naive models, a difference of .12 percentage points (n=17). This trend changed throughout the average forecast period. Beyond 5 years, the disparity between models diminished and the performance of the naive model deteriorated more than that of the forecast model. At 10 years, the difference decreased by –.05 percentage points (n=13). At 20 years, the trend reversed and the RMSE for the naive model increased to .46 percentage points compared with a smaller forecast RMSE of .43 percentage points (n=10).

Individual series analysis indicates that the naive model generally outperformed each forecast model with exception to 1955, 1957, and 1963, throughout most of the 20-year forecast period. Within the first 5-year period, the RMSE for the forecast model was smaller than or equal to the naive model 8 out of the 17 series (47.1 percent). Of the 51 points compared for all series combined, the naive model outperformed the forecast model 32 times (62.8 percent). Nonetheless, approximately half (51.0 percent) of the 51 comparison points maintain differences smaller than .10 of a percentage point.

Recent forecasts indicate an improvement in the Census Bureau forecast model for the short term (5 years) over the naive model. The series 1982, 1991, 1992, and 1994 model outperformed the naive model within the first 5 years with very small RMSEs ranging between .03 percentage points and .08 percentage points. Beyond 5 years, however, the naive model is smaller for series 1982 and 1986.

Summary of Forecast Error for Growth Rates

Except for the 1974 and 1976 series, the pattern of under- and overestimation and level of accuracy for the individual series are closely related to the Census Bureau's assumptions for fertility and will be discussed in detail in the following sections. The first two forecast series, 1947 and 1949, greatly underestimated the overall population growth rate as fertility rates began to rise in 1947, resulting in the Baby Boom. Short-term (5 year) accuracy improved between 1953 and 1957 as growth rates remained at high levels resulting from high fertility rates. Following 1957, the growth rate began to decline, while the Census Bureau continued forecasting high growth rates. The total populations' forecast growth rates became more accurate within the recent past with average error statistics (excluding the MPE) falling below 10 percent within the first 5 years for the past five series as population growth stabilized in the 1980s and 1990s. The average error generally increased after the 5-year forecast period; however, the direction and magnitude of error did not increase or decrease in a consistent manner. Because of large outlier error terms, the multiple forecast error statistics do not represent the actual error experienced overall for the Census Bureau's forecasts. In general, the naive model outperformed the cohort component forecast, particularly in the latter half of the forecast period. Except for the 1957 series, the naive model outperformed the forecast model for a minimum of one point in the measured forecast periods for each series. In contrast, recent cohort component forecasts consistently outperformed the naive model in the first 5 years. The overall error remained high in comparison to a naive model until the 1980s and 1990s.

Components of Change Forecasts

Fertility Forecasts Error Analysis

Throughout the first part of the 1900s, fertility rates in the United States declined until 1946 when rates increased dramatically. Graph 2 depicts the trends of the U.S. general fertility rate (births per 1,000 15- to 44-year old women) between 1943 and 1998. Following World War II, fertility rates among American women increased from 85.9 births per 1,000 women in childbearing age to 101.9 births between 1945 and 1946, representing an increase of 16.0 births (National Center for Health Statistics, 1993). Fertility rates remained unusually high, peaking at 122.7 births per 1,000 women in 1957. After 1957, rates declined until the mid 1960s. Referred to as the Baby Boom, this historic abnormality in U.S. fertility occurred between 1946 and 1964. Subsequent to the Baby Boom, except for small increases in the later part of the 1960s into the early 1990s, fertility remained stable. After 1973, fertility rates ranged between a low of 65.2 births in 1976 and a high of 70.9 births in 1990, which is a difference of 5.7 births.

Of the three components of population change, fertility assumptions are subject to the largest levels of uncertainty. When formulating fertility assumptions as inputs for the cohort component model, demographers must attempt to forecast the trends of American women by age and in the more recent past, by race and Hispanic origin. This encompasses anticipating changes in many variables that directly or indirectly affect fertility, such as contraceptive prevalence, marital status, and female labor force participation rates. Most importantly, demographers try to anticipate potential turning points and/or the stability of the current trends.

For series produced in 1963 to 1972, the Census Bureau formulated fertility assumptions using a cohort fertility methodology as opposed to building from estimates of period fertility. That is, series were formulated based on the completed fertility of cohorts of women in childbearing ages and further adjusted for timing patterns. Timing patterns were generally based on age-specific fertility rates from past years and the average age of childbearing.[3] Assumptions pertaining to the expected level of completed fertility and timing patterns did not remain consistent across products. Estimates for the ultimate completed fertility rates were generally formulated using birth expectation data from different surveys and demographic theory, such as stable population theory and replacement level fertility (U.S. Census Bureau, 1970).[4]

Series produced in 1974, 1976, and 1982, continued the use of the cohort fertility model; however, timing patterns used previously were replaced with assumptions about short- and long-term fertility trends. These trends were also based on survey-generated birth expectations data as well as theory. Estimates used for the fertility assumptions for 1986 and 1991 continued to be based on the cohort fertility method while using Box-Jenkins time series methods to forecast short-term trends. Production of the two latest or most recent series, 1992 and 1994, switched to a period fertility methodology and assumed that the current age and race specific fertility rates remained constant throughout the forecast period.

To calculate the number of live births for a designated forecast period, age-specific birth rates were applied to the average number of women in childbearing ages. Once calculated, the births were survived forward to account for infant mortality. The number of births was summed for each calendar year. The crude birth rate is defined as the number of births per 1,000 people occurring within a calendar year.

Overall Accuracy of Fertility Forecasts

According to the MPE for the multiple series, the Census Bureau consistently tended to overestimate the fertility of American women with the absolute level of error decreasing in the 1990s. Tables 4 and 5 show that multiple series MPEs for the number of births and the crude birth rate never fell below 12 percent. Within the 20-year forecast period, the average error falls to percentages below the average error experienced within the first 10-year period. The MAPE for births increased from 13.9 percent within 5 years, to 28.3 percent and 29.4 percent at the 10th and 15th forecast period years, followed by a decline in the average error to 26.8 percent within 20 years. The average errors for crude birth rates are generally smaller than those experienced for the number of births.

Examination of the individual series forecasts for births and the crude rate display a consistent trend of overestimation until series 1982. Graph 3 displays the estimated or actual crude birth rates and the forecast crude rates for each series. According to the average statistics for the number of births (Table 4); the series produced from 1963 to 1972 greatly overestimated the

[3] Please refer to the original publication for further discussion of assumptions and methodology.

[4] To collect birth expectation data, the Census Bureau used national survey data from the Growth of American Families Studies, the University of Michigan national sample surveys, and the Current Population Survey (U.S. Census Bureau, 1967, 1984).

number of births in comparison to later series. The series with the largest error during the first 5 years, 1970, experienced a MAPE of 29.0 percent. This error increased to 37.1 percent during the 10-year period and 39.4 percent within 15 years. MAPEs for the remaining series (1963, 1966, 1969, and 1972) ranged between 12.5 and 17.6 percent during the first 5 years and 20.3 and 37.1 percent within 10 years. The series for 1972, however, did not increase as rapidly with an average error remaining between 18 and 21 percent throughout the period. Series 1963 and 1966 experienced the largest MAPE statistics, 42.9 percent and 46.2 percent, respectively, for long-term forecast periods (15 and 20 years).

Table 6 shows that PEs for the first year of forecast births and rates for 1966, 1970 and 1972, were larger than other series. The PE in the first year for 1972 of 10.7 percent (CBR=11.3) and for 1970 of 8.6 percent (CBR=9.1) indicate that these series began with inadequate base data. In addition, 1970 represents a turning point in fertility trends as the number of births declined from 1970 to 1973. Each forecast with base dates before 1974 failed to incorporate the decline and subsequent stability in fertility patterns seen throughout the early and mid-1970s.

After 1972, forecast error for the number of births decreased substantially from previous series, with continued improvement in the recent past. During the first 5 years, the MAPE for series produced after 1972 ranged between a low of .5 percent (1991) and a high of 8.3 percent (1986), and within 10 years 4.0 (1982) and 9.3 percent (1986). The lowest error was experienced throughout all periods by the 1991 and 1994 series. Within 5 years, series 1991 had a MAPE of .5 percentage points and 1994 a MAPE of 0.9 percentage points.

Duration-Specific Forecast Error for Fertility

Graph 4 shows the multiple MAPEs for each component of population change for the 20-year forecast period for each single year. This MAPE represents the average absolute error occurring on the specific year of the forecast period. Error for the number of births increased throughout the first 9 years and began to stabilize past 10 years. The average error for the crude birth rate stabilized and actually declined after 10 years. This trend is attributable to specific series included with the later forecast periods and the actual trend of fertility. Specifically, series 1972, 1974, 1976, and 1982 first overestimated fertility. Later in their respective forecast periods, these series then underestimated fertility. The series underestimated fertility as the observed number of births increased in the 1980s. Therefore, because observed fertility trends increased during the 1980s and particular series forecast an eventual decline in the long term (with forecast periods falling within this time interval), the referenced series average error statistics decreased later in the forecast period. In contrast, the early series, 1963 to 1969, consistently overestimated fertility during a period of decline following the Baby Boom.

Comparison of Fertility Forecast Models

Analysis of the RMSE for the multiple series statistics indicates the naive model forecast the number of births and the crude rate more accurately (Tables 4 and 5). In addition, the values for the naive model RMSE remained at least 40 percent smaller for the number of births than the Census Bureau forecasts throughout the forecast period. During the first 10 years, the multiple series RMSE for the forecasts was 1.2 million births (CBR RMSE=5.0), in comparison to 495.1 thousand births (CBR RMSE=3.0) for the naive model. The large disparity continues throughout the 20-year period, with the naive RMSE remaining smaller than the average error experienced in the first 5 years of the Census Bureau forecast series.

Before the 1974 series, the naive model outperformed each forecast series for births and the crude birth rate. The RMSEs for the naive model never fell below 84.8 thousand for the number of births, maintaining high levels of error for each series. Within 10 years, the naive RMSE ranged between a low of 235 thousand births per year and a high of 604 thousand births. In reference to recent forecasts beginning in 1974, the forecast model outperformed the naive model for the number of births. Of the 16 points measured throughout the periods of the remaining seven series following 1972, the forecast RMSE was smaller than the naive RMSE at 11 points (68.8 percent) of the targeted forecast periods. The assumptions made for the 1976 series consistently outperformed the naive model throughout the entire 20-year period. A constant forecast of births or birth rates for the 1986 series, however, would have performed better. In contrast, the naive model for the crude birth rate outperformed the Census Bureau forecast in general. Of the 16 points observed as of 1972, the RMSEs for the crude rate naive model were greater than forecasts for only 6 points compared with 11.

Summary of Forecast Error for Fertility

The Census Bureau remained extremely optimistic about fertility trends remaining at levels experienced during the Baby Boom from 1963 to 1972, despite the continued decline experienced following the peak in 1957. Error decreased for series 1974 and 1976 because of two main factors. The 1974 series reduced the number of alternate series from four to three, resulting in one middle series with a lower completed fertility of 2.1, compared with an average of 2.5 and 2.1 for 1972. In addition, the number of births that actually occurred began to increase in the long-term forecast period. The 1976 series improved over the 1974 series by further reducing the short-term assumptions. In addition to a general improvement in the level of accuracy, the 1974 forecast began a trend of outperforming the naive model of constant rates, with exception to the 1986 model.

In contrast, the 1982 and 1986 series were conservative and resulted in underestimating births. Series 1982 continued the use of the cohort fertility approach, while the 1986 series used a Box-Jenkins time series model for short-term forecasts. The completed fertility level was further reduced to 1.9 for 1982 and 1.8 for 1986. Following the 1990 turning point, the number of births remained stable. Accuracy improved for series 1991, which continued the use of the time series model, increased the completed fertility to 2.1, and abandoned the racial convergence assumption, among other changes. This stability, combined with improved assumptions, permitted a more accurate forecast for those series produced within that decade. High levels of accuracy for short-term forecasts were duplicated for the 1994 series, which abandoned the cohort fertility method and assumed constant trends among the largest racial groups.[5]

The results of the comparison between forecast models differed for the number of births and the crude rate. The Census Bureau forecasts for the number of births were more accurate in the recent past. This is not necessarily true for the crude rate forecasts.

In summary, accuracy for the number of births improved in the recent past. Improved accuracy, however, does not seem to be explicitly determined by the different approaches toward deriving forecast assumptions (cohort vs. period) used to forecast short-term trends.

Mortality Forecasts Error Analysis

Mortality rates decreased consistently throughout the 20th century as life expectancy at birth increased from 47.3 years in 1900 to 77.0 in 1999, an increase of 29.7 years in approximately 100 years (Anderson, 1999; U.S. Census Bureau, 2000b). Graph 5 displays the observed and forecast crude death rates from 1964 to the present. Crude death rates generally decreased throughout the 1960s and 1970s, falling from 9.4 deaths per 1,000 people in 1964 to 8.6 deaths by 1977, a time span of 13 years. Following 1977, the rate remained stable, ranging between 8.5 and 8.8 deaths for 21 years. As rates stabilized or decreased, the base population continued to grow in size, resulting in an increase in the number of deaths. The number of deaths steadily increased from approximately 1.8 million in 1964 to 2.4 million in 1999. Graph 6 displays the observed number of deaths from 1964 to 1999. Between 1964 and 1983, the number of deaths increased from 1.8 to 2.0 million. Beyond 1983, the number of deaths increased to 2.4 million. These trends differ by age, sex, race, and Hispanic origin at the national level (Anderson, 1999). For the purposes of this research, only the forecast number of deaths and the crude death rate for the total population will be examined.

To forecast trends in mortality, age-specific death rates and survival rates are used as inputs to the cohort component model to survive the population forward. Rates are generally calculated by single year of age, sex, and more recently race and Hispanic origin. Mortality forecast assumptions formulated between 1963 and 1986 depended on life tables created by the Social Security Administration and were adapted to the needs of the Census Bureau. Before 1982, one set of rates was used as inputs for the model. Forecasts following 1976 produced a low, middle, and high mortality series. For series produced in 1991 forward, the Census Bureau used its own forecast life tables based primarily on the rate of mortality change experienced in previous decades.

Overall Accuracy of Mortality Forecasts

Compared to births, deaths are not as numerous and exhibit less fluctuation over time. Therefore, mortality forecasts are subject to smaller numeric magnitudes than fertility and exhibit smaller summary error statistics. Tables 7 and 8 present the error statistics for the forecast number of deaths and the crude death rates. Multiple series error statistics for the number of deaths begin with a MAPE of 5.1 percent (CDR=5.6 percent) at the fifth year of the forecast period. At the twentieth year, the MAPE reaches its highest value of 12.2 percent (CDR=9.7 percent). On average, the error terms for the number of deaths and the crude rates increased throughout the forecast periods. Correspondingly, mortality trends forecast by the

[5] Fertility among non-Hispanic White, non-Hispanic Black, and non-Hispanic American Indian women remained at constant levels, while rates for Hispanic and Asian women were assumed to decline.

Census Bureau were generally too conservative and failed to adequately forecast improvements in life expectancy.

Similar to the results for the individual fertility series, the overall accuracy of the individual mortality series for the number of deaths and the crude rates improve dramatically in the recent past. Graph 6 displays the individual series forecast for deaths and the actual number of deaths. Forecasts produced in 1976 and earlier consistently overestimated deaths. Beginning in 1963, error terms generally increased within the first 5 years for each series, peaking at 1974 (with exception of series 1972 and 1974 beyond the 15 year forecast period). Series 1974 was inaccurate by 9.9 percent (for both the MPE and MAPE), increasing from 1.8 percent in 1963, within the first 5 years. Table 9 displays the PEs for the number of deaths and the crude death rates. Again, series 1974 experienced the largest error term, with a PE of 8.2 percent at the first year for deaths and 9.1 percent for the crude rate.

Following series 1974, the level of accuracy improved. In 1976, the MAPE for the number of deaths fell to 4.6 percent during the first 5 years and again to .91 percentage points by 1982. Forecast deaths and crude rates produced after the 1976 series were consistently more accurate than previous series, except for 1992, which had a MAPE of 3.8 percent within the 5 years. The MAPE within the first 5 years for series produced after 1982, excluding 1992, ranged between .9 percentage points and 1.3 percent. For series 1982 and 1986 with forecast periods beyond 5 years, the MAPE remained near 1.0 percent and 1.1 percent.

Duration-Specific Forecast Error for Mortality

Multiple series error statistics increased throughout the forecast period for both the numbers of deaths and the crude death rates. The crude rate, however, accumulated less error throughout the forecast period. (This can also be witnessed for individual series.) Graph 4 shows the multiple MAPEs for each component of population change for the 20-year forecast periods by single year. The MAPE remains stable after 10 years for both deaths and the crude rate. Within 10 years, the crude rates demonstrated lower average error statistics, increasing the gap between the MAPEs for the number and the rate of deaths as the forecast periods lengthened.

The duration-specific forecast error for individual series deaths generally increased throughout the forecast period, with exception to 1974 and 1986. In contrast, crude rate forecasts with periods 15 years and longer, the average error declined at 20 years for series 1966 and 1969. Series 1974 and 1982 experienced smaller averages within 15 years than 10 years, followed by an increase within 20 years for 1974.

Comparison of Mortality Forecast Models

A comparison of the multiple series forecast and naive models RMSE indicates that the naive model outperformed the forecast series throughout the entire forecast period for both the number of deaths and the crude rates. The difference between the two models' RMSEs diminished within the 20-year period for the forecast number of deaths and the crude rate, with the Census Bureau forecast outperforming the naive model within 20 years for deaths. The multiple series forecast number of deaths RMSE of 265.5 thousand is smaller then the naive RMSE of 278.9 thousand. In contrast, the naive model multiple series RMSE for the crude rate outperformed the forecast series by .19 deaths per 1,000 people at 20 years.

For the individual series forecasts, the naive model of a constant number of deaths and crude rates outperformed the forecast series for every series with exception to 1982, 1986, and 1991, and long-term forecasts for 1963 and 1966. Naive models for series 1974, 1976, and 1986 produced RMSEs below 50 thousand deaths throughout the entire forecast period and were superior to the performance of Census Bureau forecasts. Within 5 years, the naive model RMSE for 1976 averaged 19.6 thousand deaths, the lowest RMSE reported for deaths.

Summary of Forecast Error for Mortality

Beginning in 1963, the Census Bureau generally underestimated improvements in life expectancy. Particular forecasts produced after 1976, in contrast, slightly overestimated improvement. Forecasts produced between 1963 and 1974 gradually increased in error, highlighting a trend of the Census Bureau's historically conservative approach toward forecasting improvements in life expectancy. Recent forecasts experienced superior performance. This improvement in accuracy may be indicative of the stabilization of mortality trends beginning in the late 1970s. In addition, the Census Bureau began producing a middle series mortality assumption for the 1982 series; potentially further contributing to the overall level of mortality forecast accuracy. Similar to fertility, the error terms for the number of deaths are slightly larger throughout the forecast period than those for the crude rate as they are more dependent on the size of the forecast population. Multiple series forecast error generally increased throughout the forecast horizon, stabilizing after the tenth year of the forecast period. Lastly, except three series, the naive mortality models outperformed the Census Bureau forecasts. In comparison to fertility, the most recent forecasts, series 1992 and 1994, did not exhibit superior performance relative to the naive model.

Net Immigration Forecasts Error Analysis

Net immigration for the United States is largely determined by domestic policy and the type of immigration occurring at any given point in history. For example, over 80 percent of the current number of immigrants entering the United States in 1999 were attributable to family reunification policy and of immigrants with refugee status (Kramer, 1999: pg. 2). In addition, the types of immigrants are controlled through bureaucratic and/or political means. During the 1970s, however, research found that the number of undocumented immigrants increased dramatically (Passel and Woodrow, 1987). This increase remains at levels researchers are unable to directly determine. The Census Bureau's current knowledge of net immigration is dependent on *legal* immigration data from the Internal Naturalization Service (INS). Given the limitation of data on the current level of net migration and the inability to predict domestic and international policy, forecasts of this component are especially problematic.

Consequently, the historical forecasts for net immigration have remained conservative. Except the most recent release in 2000 and the 1986 series, net immigration was assumed to remain constant throughout the forecast period for each series. Graph 7 depicts the observed and forecast crude rate of net immigration for each series produced as of 1963. The forecast number of immigrants was applied each year as a constant number with a constant age and sex distribution. Recent products assumed separate distributions by age, sex, race, and Hispanic origin. Characteristics of the net immigrant populations experienced around the time of the base year generally represented the forecast distributions.

As a result of these complicating factors and those mentioned above in relation to emigration, undocumented immigration, the change in the universe, serious limitations to the evaluation of the accuracy of net immigration forecasts exist. Nevertheless, it may still be profitable to examine these data at some level to further understand how they affect results of the forecast and inform us about trends. Analysis of the immigration component for this report is conducted at a general level.

Overall Accuracy of Net Immigration Forecasts

The forecast number of immigrants and the net immigration rate are consistently underestimated in each forecast and the magnitude of error for both variables is larger than either components of population change discussed previously (Table 10). For multiple series error, the MPE for the number of immigrants is underestimated by –21.0 percent at the fifth period year (Table 11). The RMSE at 5 years is 189.2 thousand immigrants. At the tenth year, the MPE increased to –36.5 percent and –50.2 percent at 20 years. The number of immigrants and the rates' MAPE statistics correspond with the MPE statistics.

Among individual series forecasts, the overall accuracy of series 1976 demonstrated the worst performance and series 1966 performed the best. The recent series for 1991, 1992, and 1994, are more accurate within the first 5 years than past forecast series. The average error within the first 5 years for series 1992 had the smallest MAPE of 5.5 percent. The PEs for the first year of the forecast indicates that the base number of immigrants used to create the forecasts is often of poor quality. Table 10 displays the PE for both the crude rates and the number of immigrants. PEs for the number of immigrants range between –0.3 for the 1992 series and –24.0 for 1982. Of the twelve series in the first year, only five series experience PEs below 10 percent.

Duration-Specific Forecast Error for Net Immigration

As the number of immigrants increased throughout 1963 to 1999, the forecast individual series for constant numbers and rates of immigrants resulted in increasing error throughout the forecast period. As previously stated, the multiple series MAPE began at over 20 percent at the fifth year (n=13) and increased to over 50 percent at the twentieth year (n=6) for multiple series error. Graph 4 displays the MAPE by single year for each component. The MAPE for both the number and rate are larger throughout the entire forecast period than the error for fertility and mortality. A large proportion of the error occurred between the first and ninth year, increasing from approximately 10 percent to over 35 percent, a *25* percentage point increase. For individual series, the MAPE within 20 years ranged between a low of 21.9 percent for 1966 and a high of 41.8 percent for series 1976 (n=6).

Comparison of Net Immigration Forecast Models

For multiple series error statistics, the naive model outperformed the Census Bureau forecast model. At the tenth year of the forecast, the RMSE for the naive model of 244.0 thousand was smaller than the Census Bureau RMSE of 321.8 thousand immigrants. Series 1974, 1991, 1992, and 1994 are the only forecasts that outperformed the naive model (with exception to 1970 within the first 5 years). For crude rates, only three series (1970, 1991, and 1992) outperformed the naive model and only within the first 5 years. The naive model is based on adjusted numbers for net undocumented immigrants and emigrants in the 1970s and afterward. Graph 8 displays the multiple series RMSE for both models for the forecast crude rate of net immigration. This offers a hypothetical or possible representation of the RMSE for the Census Bureau forecasts if the base error was improved and the adjustment for undocumented immigrants and emigrants were included. With exception to the first three periods, the RMSE could be smaller for the net immigration rate as indicated by the naive model.

Summary of Forecast Error for Net Immigration

Given that actual net immigration increased throughout the period between 1963 and 1999, the forecast assumptions of constant trends resulted in consistent underestimation. Error terms throughout the forecast period increased, and maintained the highest error statistics compared to the fertility and mortality forecasts throughout. Because most of the series begin with large forecast error terms within the first year, the base data used may be contributing to a large proportion of the error throughout the forecast period. Nonetheless, net immigration forecasts have improved in the recent past. This improvement is also evident when comparing the naive and Census Bureau forecast models of net immigration. The naive model consistently outperformed the Census Bureau forecast model, with exception to the fifth year average for 1991, 1992, and 1994, for both the number of immigrants and the crude rate. In spite of this, the naive results are not a dramatic improvement over the Census Bureau forecasts.

Discussion of Results

This paper has evaluated the accuracy of population growth forecasts produced by the Census Bureau beginning with the 1947 series publication. To summarize the findings, the research questions asked previously are restated. First, how accurately did the Census Bureau forecast the total population and their respective components of change? In general, the forecasts produced by the Census Bureau overestimated total population growth. A detailed analysis of the components of population change, however, revealed a more complex pattern of over- and underestimation.

Erroneous assumptions about fertility following the Baby Boom era were largely responsible for a pattern of overestimation of the total population. Specifically, the growth rate forecast performance worsened for the series produced between 1957 and 1972. The number of births and the crude rate were severely overestimated between series 1963 and 1972, influencing the forecast growth rate. Before the 1957 series and following the 1972 series, annual growth rates were underestimated. Therefore, if the fertility component was not as grievously overestimated, the forecast results may be much more conservative and possibly underestimate the series as witnessed before the 1957 and after the 1972 series.

The mortality component of change generally presents the least amount of contributing error to the forecast model in comparison to fertility and possibly net immigration. The MAPE for both the number of deaths and the crude rates begin below 5 percent at the first year and never rise above 15 percent within the 20-year period.

The assumptions for constant levels of net immigration consistently produced underestimated series as the observed number of immigrants continually increased for over 30 years. Forecasts were further troubled by the poor quality base data.

Recent forecasts for series 1991, 1992, and 1994, show improvement in accuracy over previous series within the first 5 years. Series 1991 and 1994 forecasts for fertility and mortality maintain smaller average errors than previous forecasts, while the net immigration forecasts are smaller for the 1991 and 1992 series. This improvement in accuracy may be indicative of the stabilization of the components of change of the total population. In addition, the level of detail has expanded as more race and Hispanic origin groups were added to

[6] Beginning with the 1991 series, the Census Bureau began producing forecasts with greater detail for race and Hispanic origin groups. The vital statistics data and the estimates were used to forecast four race groups by Hispanic and non-Hispanic origin. In 1982, the age distribution of the forecast population was extended from 85 years and over to 100 years and over. Lastly, for the 1991 series, the detail for immigrants was expanded to five types of immigration to the United States (U.S. Census Bureau, 1984, 1992).

the product, the terminal age of the population data rose, and the quality of input data improved.[6]

The duration-specific forecast error generally increased throughout the forecast period for both multiple series and individual series for the growth rate and the components of change. The magnitude by which the error increased differs for each component of population change. Net immigration consistently maintains the highest level of error throughout the multiple series statistics, followed by fertility and mortality. Fertility error increased rapidly within the first half of the average forecast period, but is followed by the stabilization of error terms in the latter half. This stabilization of error is most likely the result of an eventual increase in the actual fertility of American women, following a major decline in conjunction with Census Bureau assumptions for long-term fertility trends. Mortality maintains the smallest error and remains stable throughout the forecast period past the tenth forecast year, as compared to the net immigration and fertility forecasts.

Secondly, did the forecasts for the population and the components of change produced by the Census Bureau perform more accurately than a naive model assuming constant change? With exception to the recent forecasts of 1991, 1992, and 1994, and earlier series 1955, 1957, and 1963, the naive models outperformed the Census Bureau forecasts for the growth rate and each component of population change. It is evident that the Census Bureau's inability to forecast turning points in trends greatly diminishes the accuracy of each forecast series.

The assumption of constancy for the naive model outperformed the Census Bureau forecasts for series experiencing a change in trends. In contrast, once the population stabilized in the recent past or experienced minimal to moderate change before the Baby Boom, the Census Bureau forecasts generally outperformed the naive model.

Conclusion

During the 1900s, knowledge of population trends and their future repercussions for the size and distribution of the population became increasingly important as the U.S. experienced major shifts in fertility and net immigration. Population forecasts produced by the Census Bureau are used widely, informing researchers, planners, legislators, and many others on the future course of population change. Because forecasts are subject to inherent uncertainty, as they are based on a compilation of reasonable assumptions for the components of population change, it is essential to educate customers as to the amount of uncertainty within the forecasts for the population and the components of population change. Throughout the second half of the century, the forecasts produced by the Census Bureau improved in accuracy as a result of several factors including improvements in data quality and methodology. Nonetheless, this study reveals that forecasters failed to foresee turning points in population trends, resulting in erroneous forecasts, particularly for fertility and net immigration. In addition, with exception of net immigration, the assumptions formulated by the Bureau were often outperformed by simple assumptions of constancy.

The forecast reliability is, in all likelihood, the result of the stabilization of the components of population change.

Recent forecasts produced in the 1990s minimize the inherent uncertainty and provide a reliable product for consumers. This research addresses the error experienced for general characteristics of the forecasts. Previous studies by Long (1987), Stoto (1983), and Ascher (1978) examined the accuracy of the Census Bureau population growth rate for individual series. The present study makes a contribution to this body of research by using a multi-pronged approach, combining the analysis of the individual error terms, individual series error, and multiple series error. In contrast to previous studies, this research evaluates and compares the accuracy results with multiple statistical tools, strengthening its validity. This study is unique in that it is the only one to systematically analyze the forecast error for the population growth rates in combination with the respective components of change for the U.S. Census Bureau. In addition, this research represents the only detailed accuracy analysis of the net immigration and mortality forecasts.

In order to reduce uncertainty for future products, further analysis is necessary to understand the uncertainty in forecasting specific characteristics of the population, such as the forecasts of the race and Hispanic origin distribution and the age-specific assumptions for the components of change. Correspondingly, a detailed analysis comparing the specific assumptions made between products and analysis of additional characteristics such as age and race or Hispanic origin-specific assumptions may strengthen the understanding of the weakness in the chosen assumptions.

References

Ahlburg, Dennis A. 1982. “How Accurate are the U.S. Bureau of the Census Projections of Total Live Births?” *Journal of Forecasting* 1: 365-374.

_____. 1992. “A Commentary on Error Measures,” *International Journal of Forecasting* 8: 99-100.

Anderson, Robert. 1999. “United States Life Tables 1997,” *National Vital Statistics Reports* Vol. 47(28), Hyattsville, MD.

Armstrong, J. Scott and Fred Collopy. 1992. “Error Measures for Generalizing about Forecasting Methods: Empirical Comparisons,” *International Journal of Forecasting* 8: 69-80.

Ascher, William. 1978. *Forecasting: An Appraisal for Policy-Makers and Planner*, Baltimore, MD: Johns Hopkins.

Keyfitz, Nathan. 1977. *Applied Mathematical Demography,* New York, NY: Wiley and Sons.

Kramer, Roger G. 1999. “Developments in International Migration to the United States: 1999,” *Immigration Policy and Research Working Paper 34* U.S. Department of Labor, Bureau of Internal Labor Affairs.

Long, John F. 1987. “The Accuracy of Population Projection Methods at the U.S. Census Bureau,” Presented at the Annual Meeting of the Population Association of America, Chicago, Illinois.

National Center for Health Statistics. 1993. *Vital Statistics of the United States,* 1992, Vol. I, Natality. Hyattsville, MD.

Passel, J. and K. Woodrow. 1987. “Change in the Undocumented Alien Population in the United States, 1979-1983,” *International Migration Review* 21(3): 1304-1334.

U.S. Census Bureau. 1949. *Forecasts of Population and School Enrollment in the United States: 1948 to 1960*. Current Population Reports, P25-18. Washington, DC: U.S. Government Printing Office.

_____. 1950. *Illustrative Projections of the Population of the United States: 1950 to 1960*. Current Population Reports, P25-43. Washington, DC: U.S. Government Printing Office.

_____. 1953. *Illustrative Projections of the Population of the United States, by Age and Sex: 1955 to 1975*. Current Population Reports, P25-78. Washington, DC: U.S. Government Printing Office.

_____. 1955. *Revised Projections of the Population of the United States, by Age and Sex: 1960 to 1975*. Current Population Reports, P25-123. Washington, DC: U.S. Government Printing Office.

_____. 1958. *Illustrative Projections of the Population of the United States, by Age and Sex: 1960 to 1980*. Current Population Reports, P25-187. Washington, DC: U.S. Government Printing Office.

_____. 1964. *Projections of the Population of the United States, by Age and Sex: 1964 to 1985*. Current Population Reports, P25-286. Washington, DC: U.S. Government Printing Office.

_____. 1967. *Projections of the Population of the United States, by Age, Sex, and Color to 1990, with Extensions of Population by Age and Sex to 2015*. Current Population Reports, P25-381. Washington, DC: U.S. Government Printing Office.

_____. 1970. *Projections of the Population of the United States, by Age and Sex (Interim Revisions): 1970 to 2020*. Current Population Reports, P25-448. Washington, DC: U.S. Government Printing Office.

_____. 1971. *Projections of the Population of the United States, by Age and Sex: 1970 to 2020*. Current Population Reports, P25-470. Washington, DC: U.S. Government Printing Office.

_____. 1972. *Projections of the Population of the United States, by Age and Sex: 1972 to 2020*. Current Population Reports, P25-493. Washington, DC: U.S. Government Printing Office.

_____. 1975. *Projections of the Population of the United States: 1975 to 2050*. Current Population Reports, P25-601. Washington, DC: U.S. Government Printing Office.

_____. 1977. *Projections of the Population of the United States: 1977 to 2050*. Current Population Reports, P25-704. Washington, DC: U.S. Government Printing Office.

_____. 1984. *Projections of the Population of the United States, by Age, Sex, and Race: 1983 to 2080*. Current Population Reports, P25-952. Washington, DC: U.S. Government Printing Office.

_____. 1989. *Projections of the Population of the United States, by Age, Sex, and Race: 1988 to 2080*. Current Population Reports, P25-1018. Washington, DC: U.S. Government Printing Office.

_____. 1990. *Population Estimates for the United States*. U.S. Census Bureau, Current Population Reports, P25-1045. Washington, DC: U.S. Government Printing Office.

_____. 1992. *Population Projections of the United States, by Age, Sex, Race, and Hispanic Origin: 1992 to 2050*. Current Population Reports, P25-1092. Washington, DC: U.S. Government Printing Office.

_____. 1993. *Population Projections of the United States, by Age, Sex, Race, and Hispanic Origin: 1993 to 2050*. Current Population Reports, P25-1104. Washington, DC: U.S. Government Printing Office.

_____. 1993. *Population Estimates for the United States*. Current Population Reports, P25-1095. Washington, DC: U.S. Government Printing Office.

_____. 1995. *National and State Population Estimates: 1990 to 1994*. Current Population Reports, P25-1127. Washington, DC: U.S. Government Printing Office.

______. 1996. *Population Projections for the United States by Age, Sex, Race, and Hispanic Origin: 1995 to 2020*. Current Population Reports, P25-1130. Washington, DC: U.S. Government Printing Office.

_____. 1999. "Population Estimates for the United States by Age, Sex, Race, Hispanic Origin, and Nativity: 1999 to 2100," Internet http://www.census.gov/population/www/estimates/uspop.html.

_____. 2000a. *Unpublished Data from the National Estimates Product*.

_____. 2000b. "Population Projections for the United States by Age, Sex, Race, Hispanic Origin, and Nativity: 1999 to 2100," http://www.census.gov/population/www/projections/natproj.html.

Smith, Stanley and Terry Sincich. 1991. "An Empirical Analysis of the Effect of Length of Forecast Horizon on Population Forecast Errors," *Demography* 28(2): 261-274.

Srinivasan, K. 1998. *Basic Demographic Techniques and Applications*, Thousand Oaks, CA: Sage Publications, Inc.

Stoto, Michael. 1983. "The Accuracy of Population Projections," *Journal of the American Statistical Association* 78(381): 13-20.

Swanson, D., J. Tayman, and C. Barr. 2000. "A Note on the Measurement of Accuracy for Subnational Demographic Estimates," *Demography* 37(2): 193-201.

Tayman, J. and D. Swanson. 1996. "On the Utility of Population Forecasts," *Demography* 33: 523-528.

Ventura, S.J., J. Martin, S.C. Curtin, and T.J. Mathews. 1999. "Births: Final Data for 1977," *National Vital Statistics Report* Hyattsville, MD Vol. 47(18).

Ventura, S.J., J. Martin, S.C. Curtin, T.J. Mathews, and M. Park. 2000. "Births: Final Data for 1998," National *Vital Statistics Report* Hyattsville, MD Vol. 48(3).

Whelpton, P.K., H.T. Eldridge, and J.S. Siegel. 1947. *Forecasts of the Population of the United States: 1945-1975*. U.S. Census Bureau and Scripps Foundation for Research in Population Problems. Washington, DC: U.S. Government Printing Office.

Table 1. Summary of the U.S. Census Bureau National Population Projections Products: 1947 to 1999

Base Year	Code	Projection Period [1]	Length of Forecast Period	Series Used for Analysis [2]
1947	P25-18	1947-1960	13	Middle
1949	P25-43	1949-1960	11	Middle
1953	P25-78	1953-1975	22	B & C
1955	P25-123	1955-1975	20	A & B
1957	P25-187	1957-1980	23	II & III
1963	P25-286	1963-1985	22	B & C
1966	P25-381	1966-1990	24	B & C
1969	P25-448	1969-1990	21	C & D
1970	P25-470	1970-1990	20	C & D
1972	P25-493	1972-2000	28	D & E
1974	P25-601	1974-2050	76	II
1976	P25-704	1976-2050	74	II
1982	P25-952	1982-2080	98	Middle
1986	P25-1018	1986-2080	94	Middle
1991	P25-1092	1991-2050	59	Middle
1992	P25-1104	1992-2050	58	Middle
1994	P25-1130	1995-2050	55	Middle
1999	WP #38	1999-2100	101	Middle

[1] The projection period reflects single year data published and available for analysis.
[2] The series are identified by the name given to each at the time of production.
Source: Population Projections Program, Population Division, U.S. Census Bureau

Table 2. Error Statistics for the Forecasted Annual Growth Rate for the Total U.S. Resident Population: 1947 to 1999

[In percents. Resident population]

Forecast Periods	Individual Series (by Base Year)																	Multiple Series
	1947	1949	1953	1955	1957	1963	1966	1969	1970	1972	1974	1976	1982	1986	1991	1992	1994	
Five years																		
MPE (%)	(31.17)	(16.51)	(14.09)	(13.58)	0.52	12.93	12.91	(1.07)	13.37	(0.89)	(20.76)	(21.49)	3.88	(8.60)	1.93	7.62	(2.54)	(3.76)
MAPE (%)	31.17	18.52	14.09	13.58	1.98	14.06	12.91	14.16	20.10	4.08	20.76	21.49	3.88	9.88	2.51	7.62	3.30	15.04
MdAPE (%)	30.39	18.05	15.41	13.21	2.65	13.19	6.62	16.66	21.01	3.34	19.86	25.06	3.39	5.43	3.11	9.66	4.16	9.20
RMSE	0.57	0.37	0.28	0.24	0.03	0.18	0.16	0.17	0.22	0.05	0.21	0.25	0.04	0.14	0.03	0.08	0.04	0.30
RMSE Naïve	0.22	0.05	0.11	0.10	0.19	0.30	0.12	0.17	0.21	0.11	0.12	0.12	0.07	0.10	0.10	0.12	0.05	0.18
Ten years																		
MPE (%)		(30.81)	(15.62)	(9.83)	13.53	20.11	29.83	7.14	21.80	3.41	(9.33)	(8.66)	(5.05)	(17.53)				9.36
MAPE (%)		31.81	15.62	11.49	14.26	20.68	29.83	14.76	25.17	6.15	14.87	13.17	8.92	18.17				26.89
MdAPE (%)		39.33	15.41	12.18	5.61	16.26	26.89	15.38	27.24	3.22	17.61	7.84	4.98	23.69				23.66
RMSE		0.61	0.28	0.21	0.25	0.25	0.34	0.17	0.27	0.09	0.17	0.18	0.13	0.21				0.37
RMSE Naïve		0.05	0.11	0.25	0.42	0.33	0.16	0.13	0.19	0.10	0.12	0.10	0.08	0.10				0.30
Fifteen years																		
MPE (%)			(9.59)	2.54	26.56	31.40	35.61	15.33	31.93	10.93	(4.93)	(9.10)	(12.01)					23.94
MAPE (%)			12.88	16.76	27.05	31.78	35.61	20.41	34.17	12.76	11.88	12.50	14.59					34.91
MdAPE (%)			14.25	14.08	25.17	31.58	41.27	17.07	32.18	8.51	10.30	7.37	17.94					31.25
RMSE			0.24	0.24	0.39	0.37	0.40	0.23	0.36	0.15	0.14	0.17	0.18					0.39
RMSE Naïve			0.30	0.45	0.54	0.38	0.15	0.12	0.22	0.13	0.10	0.09	0.07					0.38
Twenty years																		
MPE (%)			(6.77)	9.91	41.60	37.00	44.53	20.62	32.23	7.54	(9.85)	(13.91)						23.44
MAPE (%)			12.21	20.57	41.96	37.28	44.53	24.43	33.92	12.46	15.06	16.46						37.78
MdAPE (%)			12.68	16.54	39.24	38.82	46.98	20.75	32.85	10.24	17.61	22.06						28.66
RMSE			0.22	0.27	0.54	0.42	0.47	0.26	0.35	0.15	0.18	0.20						0.43
RMSE Naïve			0.39	0.54	0.63	0.39	0.18	0.11	0.24	0.12	0.11	0.09						0.46

Source: Population Projections Program, Population Division, U.S. Census Bureau: May 2000

Table 3. Percent Error for the Total U.S. National Population Forecasted Annual Growth Rates: 1947 to 1999

[In percents. Resident population]

Base Year	Percent Error (%) of Forecast Period				
	1st	5th	10th	15th	20th
1947	(12.69)	(48.62)			
1949	(5.02)	(35.80)	(47.27)		
1953	(6.23)	(15.41)	(14.25)	16.77	10.42
1955	(15.05)	(9.20)	8.30	14.08	37.13
1957	0.82	2.79	47.76	64.34	83.74
1963	(2.83)	29.16	50.66	46.20	69.88
1966	6.41	4.66	56.71	61.34	66.69
1969	(16.66)	20.30	10.99	47.59	27.44
1970	(16.83)	27.47	23.66	52.31	10.83
1972	(8.51)	3.09	20.72	21.04	(15.75)
1974	(26.49)	(18.09)	14.58	(5.08)	(26.08)
1976	(25.06)	(6.01)	2.23	(24.00)	(29.87)
1982	2.25	3.39	(24.95)	(31.25)	
1986	3.21	(22.31)	(27.44)		
1991	0.01	3.11			
1992	4.94	1.94			
1994	1.37	(4.35)			

Source: Population Projections Program, Population Division, U.S. Census Bureau: May 2000

Table 4. Error Statistics for the Forecasted Number of Births for the Total U.S. Resident Population: 1963 to 1999

[Resident population]

Forecast Period	Individual Series (By Base Year)												Multiple Series
	1963	1966	1969	1970	1972	1974	1976	1982	1986	1991	1992	1994	
Five years													
MPE (%)	12.47	15.11	14.61	29.04	17.60	6.37	0.56	2.42	(8.34)	0.19	2.58	0.08	11.97
MAPE (%)	12.47	15.11	15.80	29.04	17.60	6.37	1.76	2.42	8.34	0.49	2.58	0.92	13.85
MdAPE (%)	14.18	14.18	17.60	33.09	19.28	5.91	1.16	2.62	10.24	0.50	2.59	0.95	9.42
RMSE	514,994	575,225	600,227	995,674	582,069	228,178	85,171	92,489	357,445	22,365	102,095	36,616	627,065
RMSE Naïve	465,722	84,820	351,984	513,028	100,353	184,380	337,574	78,290	261,370	162,579	147,085	49,386	346,936
Ten years													
MPE (%)	23.89	34.06	23.84	37.13	20.32	8.83	3.71	(1.47)	(9.32)				24.26
MAPE (%)	23.89	34.06	24.43	37.13	20.32	8.83	4.32	3.89	9.32				28.33
MdAPE (%)	20.76	32.50	29.89	42.78	21.90	8.99	4.75	3.04	10.11				23.39
RMSE	961,809	1,265,388	879,576	1,290,937	709,716	336,487	191,381	184,829	381,507				1,152,530
RMSE Naïve	603,597	338,116	356,029	453,081	235,066	368,031	447,665	278,243	204,582				495,138
Fifteen years													
MPE (%)	37.32	42.07	27.06	39.39	21.31	8.50	2.37	(3.46)					26.20
MAPE (%)	37.32	42.07	27.46	39.39	21.31	8.50	4.08	5.08					29.39
MdAPE (%)	30.94	53.59	31.75	43.50	22.42	9.32	4.32	6.88					28.19
RMSE	1,431,970	1,558,015	1,008,076	1,406,844	766,643	332,862	180,409	226,634					1,324,306
RMSE Naïve	724,396	301,200	291,210	372,706	333,660	505,362	610,939	268,718					577,229
Twenty years													
MPE (%)	42.94	46.22	28.26		18.48	5.86	0.76						24.58
MAPE (%)	42.94	46.22	28.56		18.48	6.89	4.07						26.79
MdAPE (%)	57.53	57.13	31.99		21.23	7.27	4.24						17.05
RMSE	1,644,427	1,731,781	1,067,397		696,644	291,309	175,922						1,360,758
RMSE Naïve	681,917	263,650	281,452		494,633	627,910	662,275						610,736

Source: Population Projections Program, Population Division, U.S. Census Bureau: May 2000

Table 5. Error Statistics for the Forecasted Crude Birth Rates for the Total U.S. Resident Population: 1963 to 1999

[Rate per 1,000 persons. Resident population]

Forecast Period	Individual Series (By Base Year)												Multiple Series
	1963	1966	1969	1970	1972	1974	1976	1982	1986	1991	1992	1994	
Five years													
MPE (%)	12.15	14.52	14.70	28.98	18.25	8.05	2.74	2.28	(8.51)	0.20	2.52	0.21	12.17
MAPE (%)	12.15	14.52	15.86	28.98	18.25	8.05	3.30	2.28	8.51	0.49	2.52	0.99	13.99
MdAPE (%)	14.13	13.48	17.98	33.26	19.68	7.58	2.34	2.50	10.15	0.29	2.34	0.97	10.52
RMSE	2.52	2.70	2.83	4.63	2.77	1.28	0.61	0.37	1.44	0.10	0.38	0.15	2.90
RMSE Naïve	3.22	0.85	2.27	2.99	0.77	0.39	1.01	0.25	0.60	1.15	1.06	0.60	2.05
Ten years													
MPE (%)	22.79	32.14	23.47	36.10	20.92	10.87	6.13	(1.45)	(8.89)				23.79
MAPE (%)	22.79	32.14	24.05	36.10	20.92	10.87	6.40	3.73	8.89				27.32
MdAPE (%)	19.62	31.37	29.21	40.84	22.69	11.09	7.45	3.04	9.39				23.65
RMSE	4.40	5.57	3.95	5.68	3.26	1.78	1.14	0.70	1.41				4.94
RMSE Naïve	4.43	2.67	2.65	3.01	0.59	0.71	1.09	0.37	0.73				3.00
Fifteen years													
MPE (%)	34.53	38.46	26.10	37.10	21.50	10.50	4.82	(2.99)					23.64
MAPE (%)	34.53	38.46	26.48	37.10	21.50	10.50	5.46	4.51					25.69
MdAPE (%)	29.40	47.98	30.56	39.75	22.10	11.20	5.10	5.55					26.03
RMSE	6.12	6.44	4.30	5.84	3.36	1.73	1.02	0.78					5.02
RMSE Naïve	5.47	2.81	2.51	2.86	0.49	0.86	1.33	0.65					3.04
Twenty years													
MPE (%)	38.44	40.65	26.44		18.35	7.99	3.51						20.68
MAPE (%)	38.44	40.65	26.72		18.35	8.00	4.21						20.83
MdAPE (%)	48.89	47.94	30.54		21.59	9.37	2.21						14.15
RMSE	6.61	6.69	4.32		3.01	1.50	0.88						4.54
RMSE Naïve	5.66	2.82	2.40		0.58	0.96	1.23						2.94

Source: Population Projections Program, Population Division, U.S. Census Bureau: May 2000

Table 6. Percent Error for the Fertility Forecasts of the U.S.: 1963 to 1999

[In percents. Resident population]

Base Year	Percent Error (%) of Forecast Period				
	1st	5th	10th	15th	20th
Births					
1963	0.80	20.76	56.93	65.88	61.80
1966	8.73	24.68	62.03	57.22	58.24
1969	(2.96)	28.10	31.68	36.41	26.25
1970	8.62	42.99	42.56	41.10	
1972	10.70	21.08	23.39	19.97	7.84
1974	2.78	8.53	13.26	1.80	(1.99)
1976	(2.98)	4.32	6.81	(4.25)	(4.64)
1982	2.83	1.67	(7.21)	(8.50)	
1986	(3.88)	(10.31)	(11.13)		
1991	(0.66)	0.50			
1992	2.14	2.32			
1994	1.06	(0.95)			
Crude Birth Rate					
1963	0.87	19.99	53.63	57.87	50.21
1966	8.44	23.80	57.40	48.75	44.96
1969	(2.88)	27.71	30.71	33.23	21.29
1970	9.14	41.93	39.75	35.03	
1972	11.37	21.78	23.65	18.83	7.01
1974	3.69	10.88	15.49	3.63	1.06
1976	(1.39)	7.00	9.35	(1.46)	(0.47)
1982	3.02	1.34	(6.38)	(6.76)	
1986	(4.11)	(10.15)	(9.52)		
1991	(0.73)	0.29			
1992	1.95	2.22			
1994	1.20	(0.77)			

Source: Population Projections Program, Population Division, U.S. Census Bureau: May 2000

Table 7. Error Statistics for the Forecasted Number of Deaths for the Total U.S. Resident Population: 1963 to 1999

[Resident population]

Forecast Period	Individual Series (By Base Year)												Multiple Series
	1963	1966	1969	1970	1972	1974	1976	1982	1986	1991	1992	1994	
Five years													
MPE (%)	1.67	2.71	3.28	7.04	7.07	9.88	4.62	(0.91)	0.75	(0.24)	(3.78)	1.17	4.51
MAPE (%)	1.75	2.77	3.28	7.04	7.07	9.88	4.62	0.91	1.25	0.93	3.78	1.29	5.05
MdAPE (%)	1.94	3.08	2.68	5.35	8.67	10.09	3.76	1.12	1.26	0.90	4.30	1.67	3.55
RMSE	36,877	61,565	73,416	145,294	151,414	190,722	92,515	21,869	29,102	23,844	90,270	34,328	128,743
RMSE Naïve	36,485	37,336	53,168	74,715	110,314	46,502	19,597	47,199	35,190	40,627	53,298	24,018	78,293
Ten years													
MPE (%)	2.88	5.70	7.84	10.45	9.73	10.35	5.64	(0.45)	(0.13)				8.71
MAPE (%)	2.92	5.73	7.84	10.45	9.73	10.35	5.64	0.96	1.13				8.73
MdAPE (%)	2.70	4.67	8.68	12.07	10.94	10.63	6.06	0.91	1.21				10.96
RMSE	64,361	134,071	176,634	215,724	202,806	204,429	116,562	24,133	27,557				200,461
RMSE Naïve	61,657	114,019	146,035	149,290	145,267	47,106	22,778	46,768	30,085				150,582
Fifteen years													
MPE (%)	6.11	8.62	9.69	11.92	10.47	10.13	6.39	(0.77)					10.87
MAPE (%)	6.14	8.65	9.69	11.92	10.47	10.13	6.39	1.11					10.97
MdAPE (%)	4.48	8.71	11.61	13.23	11.41	10.09	6.35	1.29					12.36
RMSE	150,102	195,863	211,644	246,975	218,858	205,103	137,807	28,141					241,556
RMSE Naïve	161,501	195,258	183,613	192,889	157,270	40,718	27,397	64,817					217,786
Twenty years													
MPE (%)	8.30	9.99	10.34		10.97	10.36	6.90						12.18
MAPE (%)	8.32	10.01	10.34		10.97	10.36	6.90						12.18
MdAPE (%)	7.05	12.79	12.21		11.46	10.24	6.86						13.15
RMSE	196,146	222,562	225,358		233,126	215,755	153,245						265,525
RMSE Naïve	233,990	252,523	205,619		172,701	41,704	42,064						278,889

Source: Population Projections Program, Population Division, U.S. Census Bureau: May 2000

Table 8. Error Statistics for the Forecasted Crude Death Rates for the Total U.S. Resident Population: 1963 to 1999

[Rate per 1,000 persons. Resident population]

Forecast Period	Individual Series (By Base Year)												Multiple Series
	1963	1966	1969	1970	1972	1974	1976	1982	1986	1991	1992	1994	
Five years													
MPE (%)	1.34	2.11	3.35	7.03	7.74	11.43	6.86	(1.15)	0.54	(0.28)	(4.02)	1.22	4.77
MAPE (%)	1.68	2.31	3.35	7.03	7.74	11.43	6.86	1.15	1.21	0.93	4.02	1.34	5.58
MdAPE (%)	2.09	2.48	2.76	5.64	8.98	11.56	6.32	1.68	1.30	0.97	4.46	1.73	3.14
RMSE	0.17	0.24	0.36	0.67	0.74	1.00	0.60	0.12	0.12	0.09	0.36	0.13	0.63
RMSE Naïve	0.17	0.14	0.26	0.33	0.55	0.34	0.19	0.20	0.15	0.16	0.22	0.09	0.38
Ten years													
MPE (%)	1.96	4.21	7.43	9.58	10.27	12.31	8.08	(0.31)	0.30				8.54
MAPE (%)	2.13	4.31	7.43	9.58	10.27	12.31	8.08	1.18	1.15				8.54
MdAPE (%)	2.17	3.33	8.55	10.77	11.31	12.82	8.56	1.49	1.02				8.88
RMSE	0.21	0.46	0.75	0.88	0.94	1.07	0.71	0.12	0.11				0.84
RMSE Naïve	0.20	0.37	0.61	0.57	0.69	0.36	0.18	0.19	0.11				0.57
Fifteen years													
MPE (%)	4.09	5.93	8.80	10.05	10.64	12.10	8.93	(0.26)					9.24
MAPE (%)	4.20	6.00	8.80	10.05	10.64	12.10	8.93	1.10					9.24
MdAPE (%)	2.46	7.03	10.52	10.83	11.25	11.77	8.88	1.02					9.43
RMSE	0.46	0.60	0.84	0.91	0.96	1.06	0.79	0.11					0.85
RMSE Naïve	0.51	0.59	0.72	0.65	0.69	0.33	0.17	0.21					0.65
Twenty years													
MPE (%)	5.04	5.92	8.74		10.82	12.54	9.84						9.70
MAPE (%)	5.12	5.97	8.74		10.82	12.54	9.84						9.70
MdAPE (%)	4.92	6.47	9.49		11.31	12.57	9.43						10.94
RMSE	0.53	0.59	0.82		0.97	1.10	0.88						0.90
RMSE Naïve	0.67	0.67	0.72		0.71	0.34	0.16						0.71

Source: Population Projections Program, Population Division, U.S. Census Bureau: May 2000

Table 9. Percent Error for the Mortality Forecasts of the U.S.: 1963 to 1999

[In percents. Resident population]

Base Year	Percent Error (%) of Forecast Period				
	1st	5th	10th	15th	20th
DEATHS					
1963	2.35	(0.21)	5.32	14.19	14.50
1966	3.08	4.94	12.74	14.59	12.96
1969	1.32	6.81	14.79	13.54	13.34
1970	5.35	11.89	12.95	13.23	
1972	1.15	11.41	13.97	11.50	13.87
1974	8.21	12.17	10.96	10.65	9.80
1976	3.61	5.86	6.35	9.66	8.58
1982	(0.34)	(1.12)	1.38	(0.41)	
1986	(1.26)	1.98	(0.10)		
1991	1.59	0.15			
1992	(5.14)	(1.91)			
1994	(0.31)	2.17			
CRUDE DEATH RATE					
1963	2.10	(0.85)	3.25	8.38	6.29
1966	2.48	3.98	8.80	8.42	3.48
1969	1.19	7.23	13.81	10.89	8.88
1970	5.64	11.00	10.72	8.36	
1972	1.92	12.37	14.21	10.45	13.00
1974	9.11	14.47	13.15	12.64	13.22
1976	5.61	8.58	8.88	12.86	13.33
1982	(0.24)	(1.68)	2.12	1.92	
1986	(1.69)	2.12	1.92		
1991	1.58	0.06			
1992	(5.20)	(2.28)			
1994	(0.31)	2.30			

Source: Population Projections Program, Population Division, U.S. Census Bureau: May 2000

Table 10. Percent Error for Net Immigration Forecasts of the U.S.: 1963 to 1999

[In percents]

Base Year	Percent Error (%) of Forecast Period				
	1st	5th	10th	15th	20th
IMMIGRANTS (net of emigration)					
1963	(5.36)	(24.62)	(26.29)	(48.63)	(49.34)
1966	(3.38)	(13.61)	(6.76)	(42.00)	(39.51)
1969	(22.18)	2.04	(35.06)	(32.11)	(43.82)
1970	(13.61)	(23.81)	(52.66)	(38.32)	
1972	(1.72)	(14.89)	(32.77)	(39.92)	(60.28)
1974	(23.81)	(35.06)	(32.11)	(43.82)	(50.66)
1976	(14.89)	(42.00)	(39.51)	(58.34)	(57.32)
1982	(24.01)	(32.41)	(55.31)	(53.94)	
1986	(10.11)	(44.89)	(48.31)		
1991	(12.62)	(6.09)			
1992	(0.29)	(9.92)			
1994	(4.42)	(4.22)			
NET IMMIGRATION RATE					
1963	(8.02)	(25.79)	(27.83)	(51.01)	(52.97)
1966	(4.35)	(14.31)	(9.57)	(45.12)	(44.59)
1969	(22.23)	1.94	(35.63)	(33.69)	(46.03)
1970	(13.28)	(24.13)	(53.60)	(40.97)	
1972	(1.11)	(14.40)	(32.63)	(40.48)	(60.59)
1974	(21.81)	(34.21)	(31.81)	(44.42)	(48.56)
1976	(15.62)	(39.98)	(38.13)	(57.90)	(54.67)
1982	(24.82)	(30.71)	(54.35)	(53.36)	
1986	(11.16)	(44.20)	(47.87)		
1991	(11.23)	(6.50)			
1992	(0.56)	(9.46)			
1994	(4.94)	(4.36)			

Source: Population Projections Program, Population Division, U.S. Census Bureau: May 2000

Table 11. Error Statistics for the Forecasted Number of Immigrants Net of Emigration for the Total U.S. Resident Population: 1963 to 1999

[Resident population]

Forecast Period	Individual Series (By Base Year)												Multiple Series
	1963	1966	1969	1970	1972	1974	1976	1982	1986	1991	1992	1994	
Five years													
MPE (%)	(22.23)	(10.07)	(7.14)	(7.47)	(9.03)	(22.41)	(35.22)	(28.52)	(21.59)	(1.58)	(1.04)	(8.38)	(20.79)
MAPE (%)	22.23	10.27	7.96	8.29	9.84	22.41	35.22	28.52	24.01	6.02	5.48	8.38	21.13
MdAPE (%)	24.62	11.70	2.04	2.04	6.76	23.81	35.06	30.61	17.84	6.09	6.09	4.67	19.35
RMSE	102,218	63,204	58,445	62,782	65,542	142,743	271,040	184,491	276,493	70,267	59,906	92,414	189,197
RMSE Naïve	54,944	41,180	49,723	91,459	64,866	149,788	245,622	49,605	210,064	91,180	128,113	100,299	145,237
Ten years													
MPE (%)	(27.33)	(8.09)	(14.77)	(17.82)	(23.91)	(30.40)	(35.13)	(33.69)	(31.24)				(36.53)
MAPE (%)	27.33	8.59	15.18	18.23	24.32	30.40	35.13	33.69	32.45				36.53
MdAPE (%)	26.91	5.07	14.25	14.25	27.66	32.28	33.92	31.98	38.53				35.06
RMSE	130,256	60,460	109,067	174,352	205,406	222,383	246,596	293,748	329,232				321,813
RMSE Naïve	78,158	48,212	78,830	132,608	204,651	229,378	219,918	183,427	222,725				244,045
Fifteen years													
MPE (%)	(30.30)	(17.13)	(22.65)	(23.73)	(28.10)	(33.68)	(37.28)	(38.92)					(44.64)
MAPE (%)	30.30	17.47	22.92	24.00	28.37	33.68	37.28	38.92					44.64
MdAPE (%)	30.07	13.61	23.81	31.51	32.45	35.06	38.32	36.80					42.91
RMSE	153,830	164,087	184,684	193,284	215,297	239,604	280,327	352,272					357,351
RMSE Naïve	101,711	134,951	148,298	127,309	214,459	246,936	254,173	230,613					304,553
Twenty years													
MPE (%)	(36.28)	(21.61)	(27.04)		(32.48)	(37.77)	(41.77)						(50.16)
MAPE (%)	36.28	21.86	27.25		32.69	37.77	41.77						50.16
MdAPE (%)	34.64	22.99	32.28		33.92	38.91	39.74						50.00
RMSE	231,952	179,534	209,551		279,627	313,470	349,784						423,619
RMSE Naïve	183,119	143,108	168,448		278,807	320,495	323,787						400,816

Source: Population Projections Program, Population Division, U.S. Census Bureau: May 2000

Table 12. Error Statistics for the Forecasted Crude Net Immigration Rates for the Total U.S. Resident Population: 1963 to 1999

[Rate per 1,000 persons]

Forecast Period	Individual Series (By Base Year)												Multiple Series
	1963	1966	1969	1970	1972	1974	1976	1982	1986	1991	1992	1994	
Five years													
MPE (%)	(22.95)	(10.62)	(7.09)	(7.40)	(8.46)	(21.28)	(34.54)	(28.06)	(21.61)	(1.56)	(1.20)	(8.41)	(20.51)
MAPE (%)	22.95	10.62	7.86	8.19	9.53	21.28	34.54	28.06	24.03	5.76	5.40	8.41	20.83
MdAPE (%)	25.79	10.90	1.94	1.88	6.18	21.81	34.21	30.13	16.63	6.50	6.50	5.66	14.63
RMSE	0.52	0.31	0.28	0.29	0.29	0.62	1.16	0.75	1.09	0.26	0.23	0.33	0.77
RMSE Naïve	0.22	0.19	0.28	0.50	0.26	0.62	1.02	0.14	0.76	0.45	0.58	0.31	0.56
Ten years													
MPE (%)	(28.08)	(9.30)	(14.97)	(18.26)	(23.49)	(29.22)	(34.26)	(33.51)	(30.90)				(36.82)
MAPE (%)	28.08	9.30	15.35	18.65	24.03	29.22	34.26	33.51	32.11				36.82
MdAPE (%)	28.12	6.96	14.43	14.56	27.21	31.59	33.48	30.78	38.62				36.18
RMSE	0.65	0.30	0.50	0.79	0.90	0.95	1.04	1.16	1.25				1.30
RMSE Naïve	0.29	0.29	0.31	0.56	0.80	0.91	0.85	0.62	0.72				0.82
Fifteen years													
MPE (%)	(31.56)	(18.87)	(23.12)	(24.74)	(27.92)	(32.49)	(36.41)	(38.73)					(45.87)
MAPE (%)	31.56	18.87	23.38	25.00	28.28	32.49	36.41	38.73					45.87
MdAPE (%)	32.63	14.31	23.99	32.48	32.48	34.21	37.07	37.47					37.95
RMSE	0.75	0.77	0.83	0.87	0.93	0.99	1.15	1.36					1.46
RMSE Naïve	0.33	0.49	0.53	0.48	0.78	0.91	0.92	0.70					0.97
Twenty years													
MPE (%)	(37.93)	(24.08)	(27.89)		(32.49)	(36.63)	(40.65)						(51.24)
MAPE (%)	37.93	24.08	28.09		32.76	36.63	40.65						51.24
MdAPE (%)	34.91	24.24	33.73		33.66	37.28	38.07						50.94
RMSE	1.08	0.85	0.92		1.16	1.25	1.36						1.67
RMSE Naïve	0.62	0.46	0.54		0.96	1.12	1.10						1.24

Source: Population Projections Program, Population Division, U.S. Census Bureau: May 2000

Table B-1. Error Statistics of the Forecasted Annual Total U.S. Resident Population: 1947 to 1999

[Resident population]

Forecast Periods	Individual Series (By Base Year)																	Multiple Series
	1947	1949	1953	1955	1957	1963	1966	1969	1970	1972	1974	1976	1982	1986	1991	1992	1995	
Five year horizon																		
ME	(2,653,600)	(1,412,200)	(1,930,400)	(2,067,000)	(762,400)	542,400	1,175,600	(122,400)	(107,200)	(1,340,500)	(3,517,600)	(4,802,940)	126,786	646,645	105,850	676,293	(97,464)	(1,203,381)
MPE (%)	(1.45)	(0.88)	(1.13)	(1.17)	(0.42)	0.27	0.58	(0.06)	(0.05)	(0.62)	(1.59)	(2.13)	0.05	0.26	0.04	0.26	(0.04)	(0.65)
MAPE (%)	1.72	0.88	1.13	1.17	0.42	0.28	0.58	0.13	0.29	0.62	1.59	2.13	0.07	0.29	0.04	0.26	0.05	0.97
MdAPE (%)	1.57	0.76	1.09	1.21	0.43	0.20	0.68	0.10	0.34	0.62	1.59	2.18	0.07	0.37	0.03	0.28	0.03	0.58
RMSPE (%)	1.92	1.03	1.21	1.22	0.42	0.38	0.62	0.14	0.31	0.62	1.62	2.15	0.07	0.32	0.06	0.28	0.06	1.33
RMSE	2,984,054	1,664,135	2,078,416	2,151,035	763,120	758,642	1,268,742	286,784	655,490	1,341,315	3,579,455	4,872,272	178,658	793,121	147,287	747,607	171,484	2,557,152
Ten year horizon																		
ME		(4,201,500)	(3,294,100)	(2,976,726)	(52,400)	1,790,000	2,824,800	591,500	1,536,554	(1,199,654)	(4,192,503)	(5,249,307)	(173,598)	(818,134)				(1,176,976)
MPE (%)		(2.46)	(1.83)	(1.61)	(0.05)	0.87	1.33	0.27	0.68	(0.54)	(1.84)	(2.27)	(0.07)	(0.30)				(0.60)
MAPE (%)		2.46	1.83	1.61	0.45	0.87	1.33	0.36	0.85	0.54	1.84	2.27	0.16	0.58				2.04
MdAPE (%)		2.15	1.86	1.72	0.41	0.87	0.96	0.22	0.60	0.58	1.96	2.37	0.10	0.42				2.02
RMSPE (%)		3.04	2.00	1.69	0.53	1.10	1.62	0.46	1.07	0.56	1.87	2.29	0.24	0.70				2.42
RMSE		5,277,357	3,656,339	3,146,095	1,028,969	2,296,240	3,472,895	1,013,425	2,408,016	1,228,706	4,273,748	5,300,538	608,709	1,838,340				4,900,383
Fifteen year horizon																		
ME			(4,262,267)	(2,794,844)	1,968,333	3,955,400	5,366,153	1,665,598	3,747,046	(375,240)	(4,227,848)	(5,548,590)	(1,310,154)					2,485,715
MPE (%)			(2.27)	(1.48)	0.94	1.84	2.43	0.72	1.61	(0.19)	(1.82)	(2.34)	(0.50)					1.13
MAPE (%)			2.27	1.48	1.27	1.84	2.43	0.78	1.73	0.54	1.82	2.34	0.56					2.96
MdAPE (%)			2.57	1.57	0.44	1.30	2.07	0.62	1.44	0.58	1.79	2.38	0.13					2.83
RMSPE (%)			2.45	1.57	1.80	2.41	3.01	1.06	2.23	0.58	1.84	2.36	0.83					3.35
RMSE			4,660,989	2,969,778	3,702,819	5,242,754	6,743,310	2,453,023	5,212,838	1,322,853	4,282,265	5,607,485	2,194,717					7,670,206
Twenty year horizon																		
ME			(4,721,050)		5,105,900	6,683,213	8,626,396	3,356,864	6,250,390	348,874	(4,763,038)	(6,554,870)						6,296,665
MPE (%)			(2.44)		2.38	2.99	3.77	1.41	2.61	0.11	(1.99)	(2.67)						2.71
MAPE (%)			2.44		2.63	3.00	3.77	1.45	2.70	0.66	1.99	2.67						4.75
MdAPE (%)			2.85		1.49	2.41	3.38	0.93	2.25	0.62	1.95	2.46						4.14
RMSPE (%)			2.58		3.75	3.87	4.71	1.97	3.41	0.72	2.03	2.75						5.54
RMSE			5,060,642		8,070,557	8,777,442	10,977,023	4,761,589	8,252,013	1,710,431	4,917,348	6,840,970						13,164,322

Source: Population Projections Program, Population Division, U.S. Census Bureau: May 2000

Table B-2. Error Statistics of the Forecasted Annual Growth Rate for the Total U.S. Resident Population: 1947 to 1999

[Per 100,000. Resident population]

Forecast Periods	Individual Series (By Base Year)																	Multiple Series
	1947	1949	1953	1955	1957	1963	1966	1969	1970	1972	1974	1976	1982	1986	1991	1992	1994	
Five year horizon																		
ME	(0.53)	(0.28)	(0.25)	(0.23)	0.01	0.14	0.13	(0.03)	0.12	(0.01)	(0.21)	(0.23)	0.03	(0.09)	0.02	0.07	(0.02)	(0.08)
MAE	0.53	0.32	0.25	0.23	0.03	0.15	0.13	0.15	0.20	0.04	0.21	0.23	0.03	0.10	0.02	0.07	0.03	0.20
MdAE	0.50	0.31	0.26	0.22	0.04	0.15	0.08	0.19	0.21	0.03	0.20	0.28	0.03	0.05	0.03	0.09	0.04	0.15
RMSE	0.57	0.37	0.28	0.24	0.03	0.18	0.16	0.17	0.22	0.05	0.21	0.25	0.04	0.14	0.03	0.08	0.04	0.30
RMSE Naïve	0.22	0.05	0.11	0.10	0.19	0.30	0.12	0.17	0.21	0.11	0.12	0.12	0.07	0.10	0.10	0.12	0.05	0.18
Ten year horizon																		
ME		(0.53)	(0.26)	(0.17)	0.16	0.21	0.29	0.06	0.22	0.03	(0.10)	(0.10)	(0.06)	(0.18)				0.07
MAE		0.55	0.26	0.19	0.17	0.22	0.29	0.15	0.26	0.06	0.15	0.14	0.09	0.18				0.30
MdAE		0.69	0.26	0.20	0.08	0.18	0.28	0.16	0.28	0.03	0.19	0.07	0.05	0.24				0.28
RMSE		0.61	0.28	0.21	0.25	0.25	0.34	0.17	0.27	0.09	0.17	0.18	0.13	0.21				0.37
RMSE Naïve		0.05	0.11	0.25	0.42	0.33	0.16	0.13	0.19	0.10	0.12	0.10	0.08	0.10				0.30
Fifteen year horizon																		
ME			(0.17)	(0.02)	0.30	0.32	0.36	0.14	0.30	0.10	(0.06)	(0.10)	(0.12)					0.23
MAE			0.20	0.22	0.30	0.32	0.36	0.20	0.33	0.12	0.12	0.13	0.14					0.35
MdAE			0.17	0.22	0.31	0.33	0.44	0.19	0.32	0.08	0.09	0.07	0.19					0.30
RMSE			0.24	0.24	0.39	0.37	0.40	0.23	0.36	0.15	0.14	0.17	0.18					0.39
RMSE Naïve			0.30	0.45	0.54	0.38	0.15	0.12	0.22	0.13	0.10	0.09	0.07					0.38
Twenty year horizon																		
ME			(0.12)		0.43	0.37	0.43	0.19	0.30	0.07	(0.11)	(0.14)						0.22
MAE			0.18		0.44	0.37	0.43	0.23	0.33	0.12	0.15	0.17						0.36
MdAE			0.17		0.47	0.44	0.48	0.22	0.33	0.10	0.19	0.23						0.27
RMSE			0.22		0.54	0.42	0.47	0.26	0.35	0.15	0.18	0.20						0.43
RMSE Naïve			0.39		0.63	0.39	0.18	0.11	0.24	0.12	0.11	0.09						0.46

Source: Population Projections Program, Population Division, U.S. Census Bureau: May 2000

Model 1. The Conceptual Model Depicting the Level of Measurement of the Forecast Error Statistics.

Forecast Period	Forecast Base Year S_1	$S_2=S_1+3$	S_F	Average Error in Forecast Period (Year)	(Value)
(1)	(2)	(3)	(F+1)		
S_1					
S_1+1	$\left(Y^{s_1}_{s_1+1} - Y^{a}_{s_1+1}\right)$				
S_1+2	$\left(Y^{s_1}_{s_1+2} - Y^{a}_{s_1+2}\right)$				
S_1+3	$\left(Y^{s_1}_{s_1+3} - Y^{a}_{s_1+3}\right)$				
S_1+4		$\left(Y^{s_2}_{s_2+1} - Y^{a}_{s_2+1}\right)$			
S_1+5		$\left(Y^{s_2}_{s_2+2} - Y^{a}_{s_2+2}\right)$			
S_1+6					
S_1+7					
S_1+8				t	$\frac{1}{F}\sum_{i=1}^{F}\left(Y^{s_i}_{s_i+t} - Y^{a}_{s_i+t}\right)$
S_1+9					
.					
1-T	$\frac{1}{T}\sum_{t=1}^{T}\left(Y^{s_1}_{s_1+t} - Y^{a}_{s_1+t}\right)$	$\frac{1}{T}\sum_{t=1}^{T}\left(Y^{s_2}_{s_2+t} - Y^{a}_{s_1+t}\right)$			

$Y^{s_i}_t$ Forecast population for series S_i at time t.

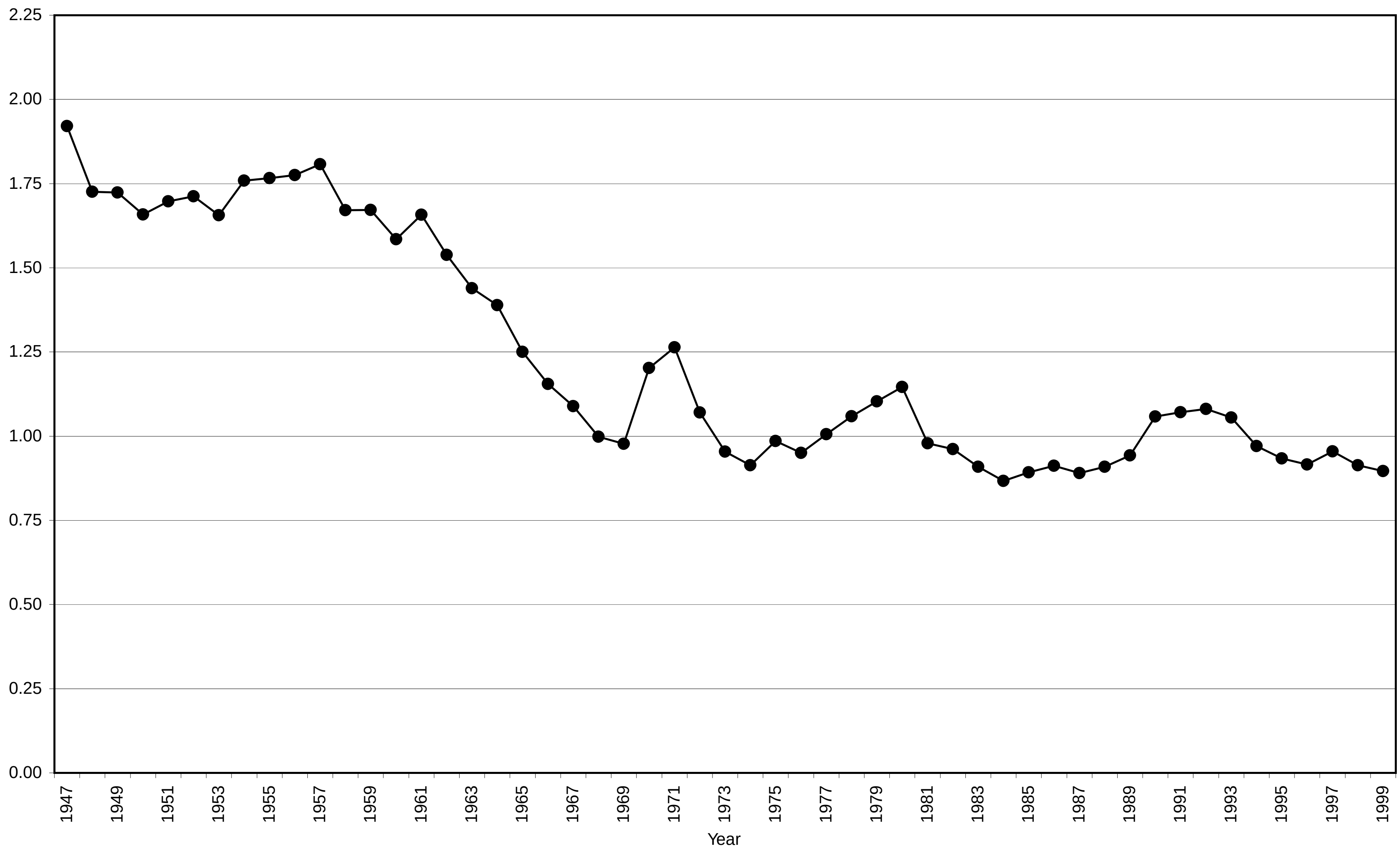
Y^{a}_t Observed population at time t.

T Number of forecast periods.

F Number of forecast series.

Graph 1. The Annual Growth Rates for the Total Population of the United States: 1947 to 1999

[Resident population]



Source: Population Estimates Program, U.S. Census Bureau: 1999

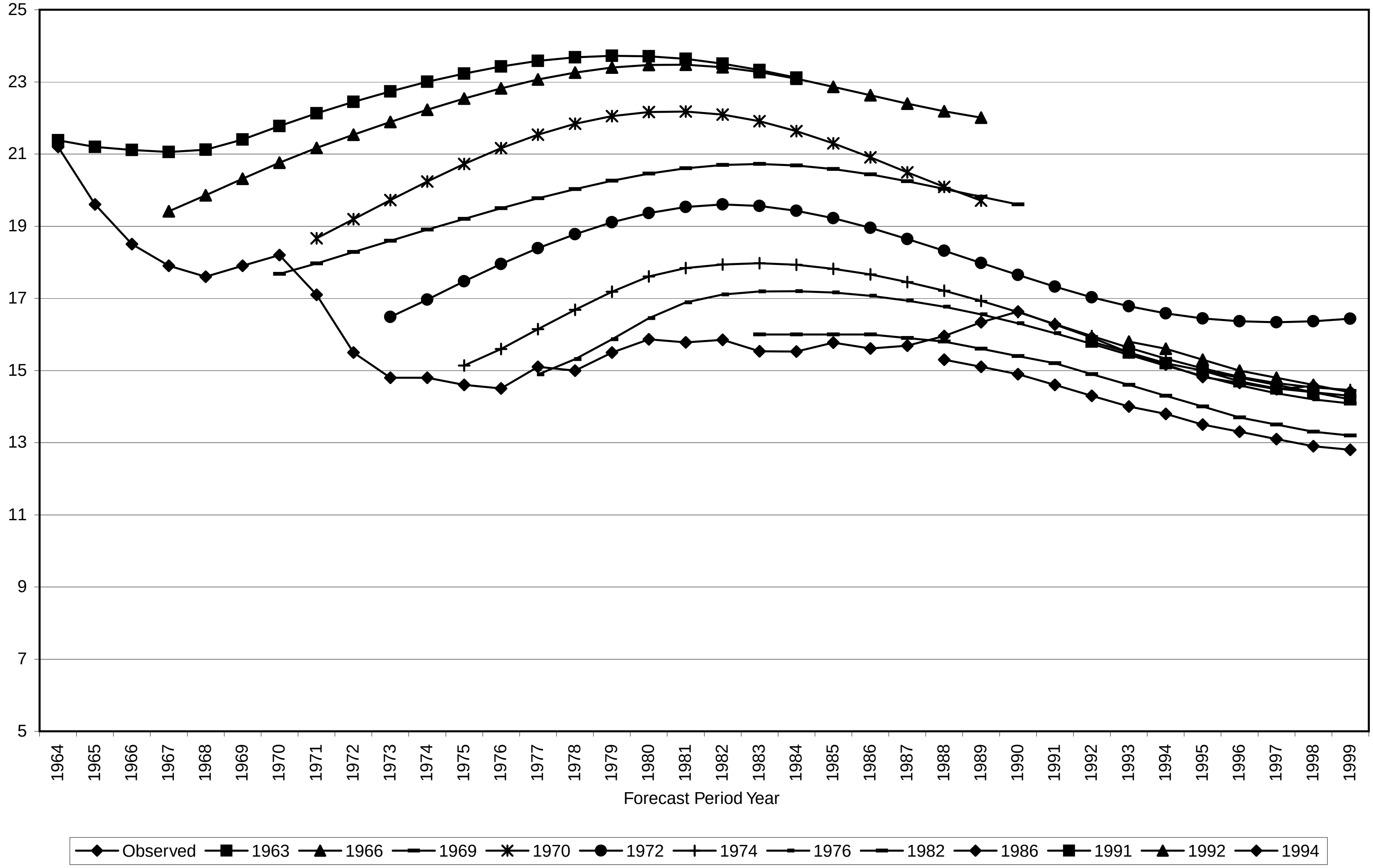
Graph 2. The General Fertility Rate for the Total Population of the United States: 1943 to 1998

[Rate per 1,000 women aged 15-44]

Source: National Center for Health Statistics. 1993. *Vital Statistics of the United States, Vol. 1, Natality* Hyattsville, MD.

Graph 3. The Observed and Forecasted Crude Birth Rates for the Total Population of the United States: 1964 to 1999

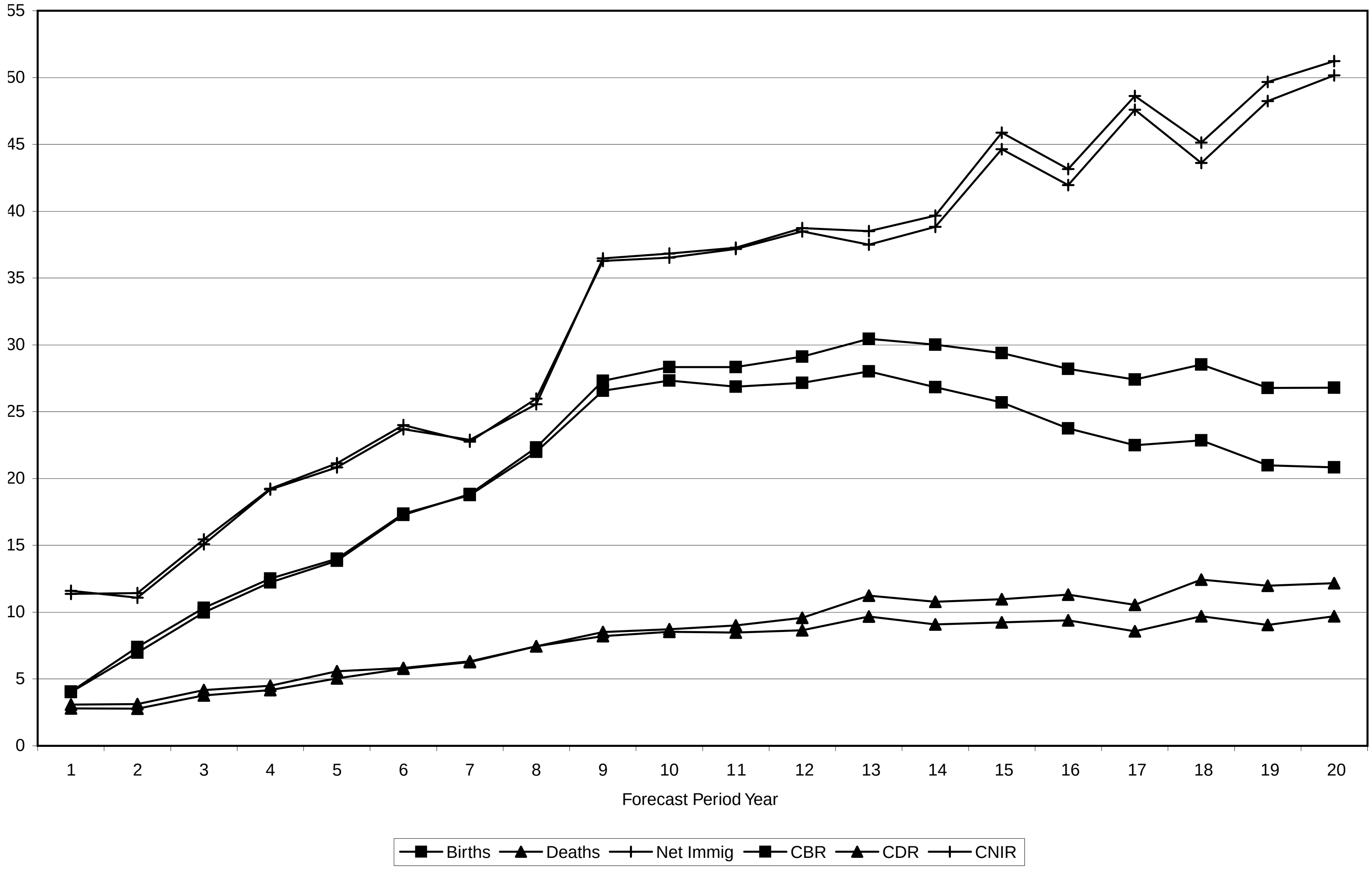
[Rate per 1,000 persons]



Source: Population Projections Program, U.S. Census Bureau: May 2000

Graph 4. The Multiple Series MAPE by Single Year for Each Component of Change and the Respective Crude Rates for the Total Population of the United States

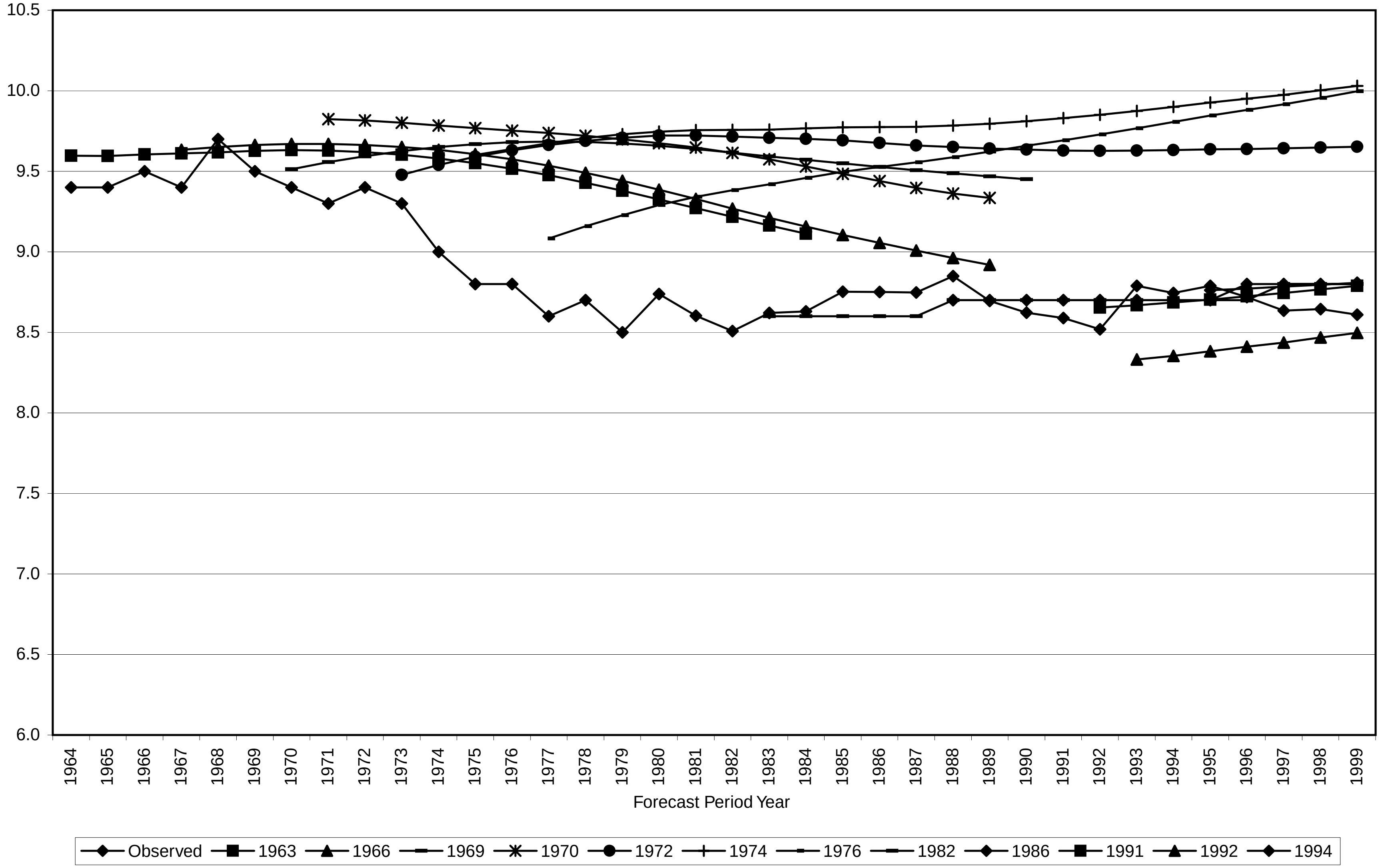
[In percents]



Source: Tables (2), (4), (5), (7), (8), (11), and (12).

Graph 5. The Observed and Forecasted Crude Death Rates for the Total Population of the United States: 1964 to 1999

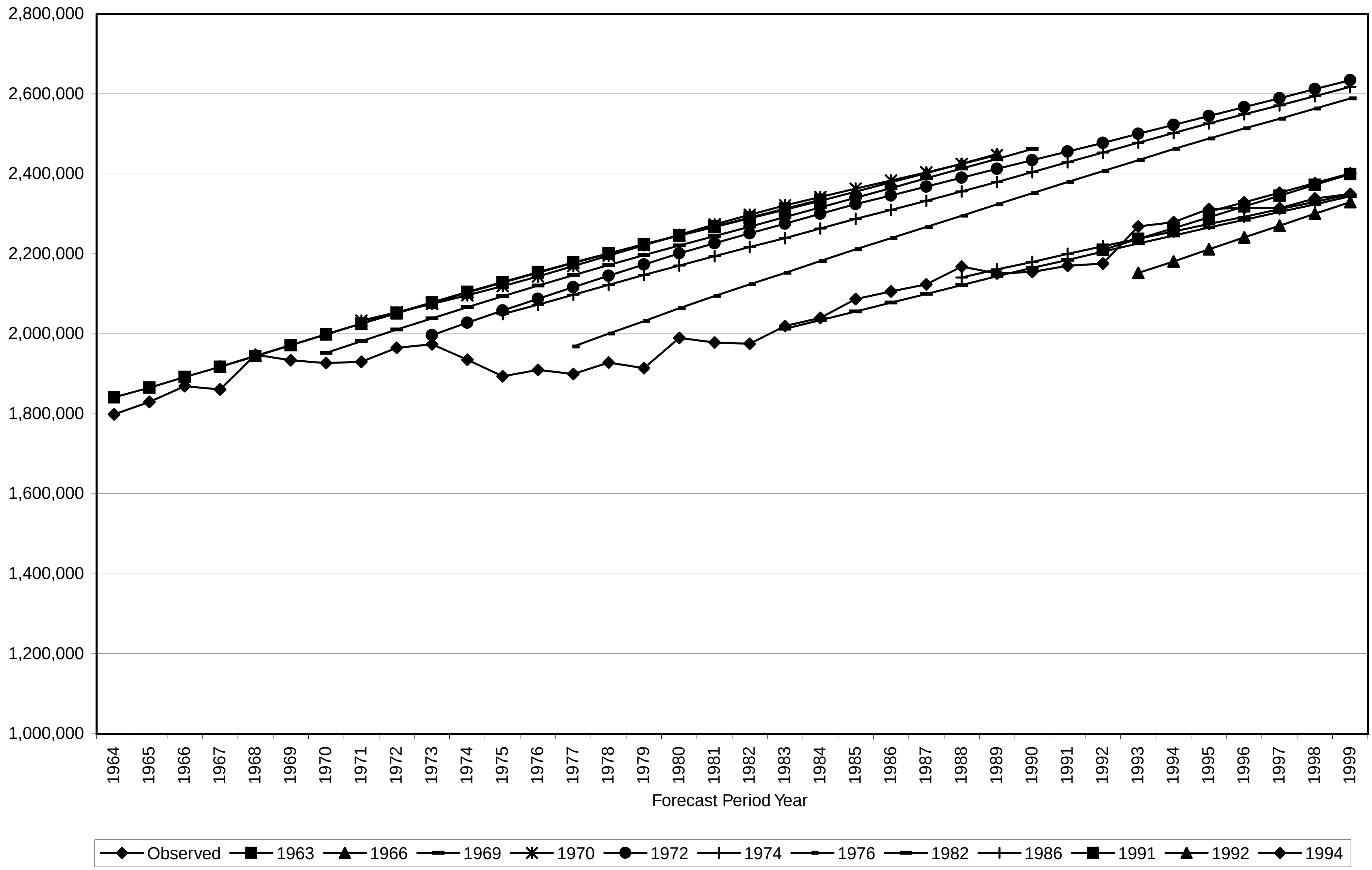
[Rate per 1,000 persons]



Source: Population Projections Program, U.S. Census Bureau: May 2000

Graph 6. The Observed and Forecasted Number of Deaths for the Total Population of the United States: 1964 to 1999

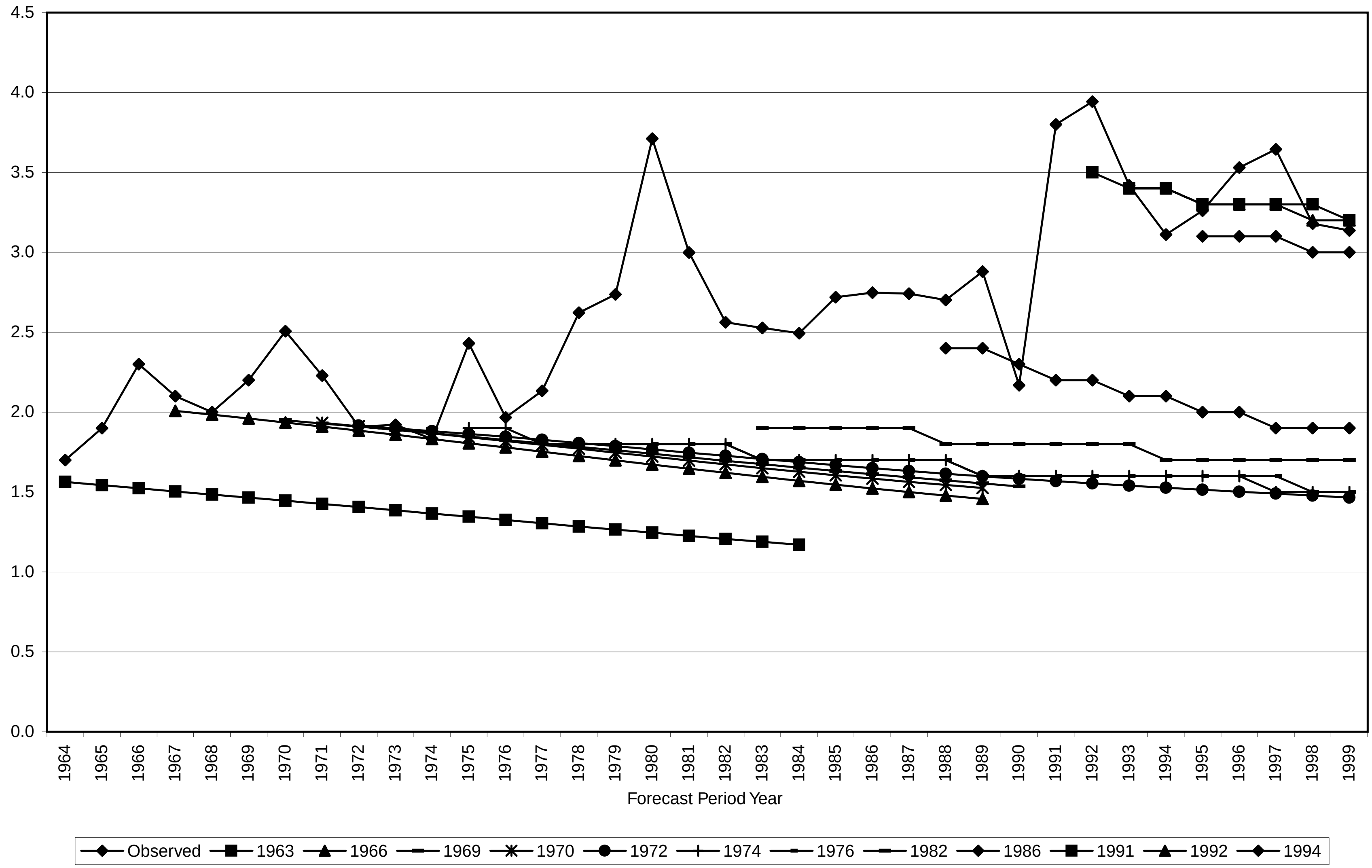
[Number of Deaths]



Source: Population Projections Program, U.S. Census Bureau: May 2000

Graph 7. The Observed and Forecasted Crude Net Immigration Rates for the Total Population of the United States: 1964 to 1999

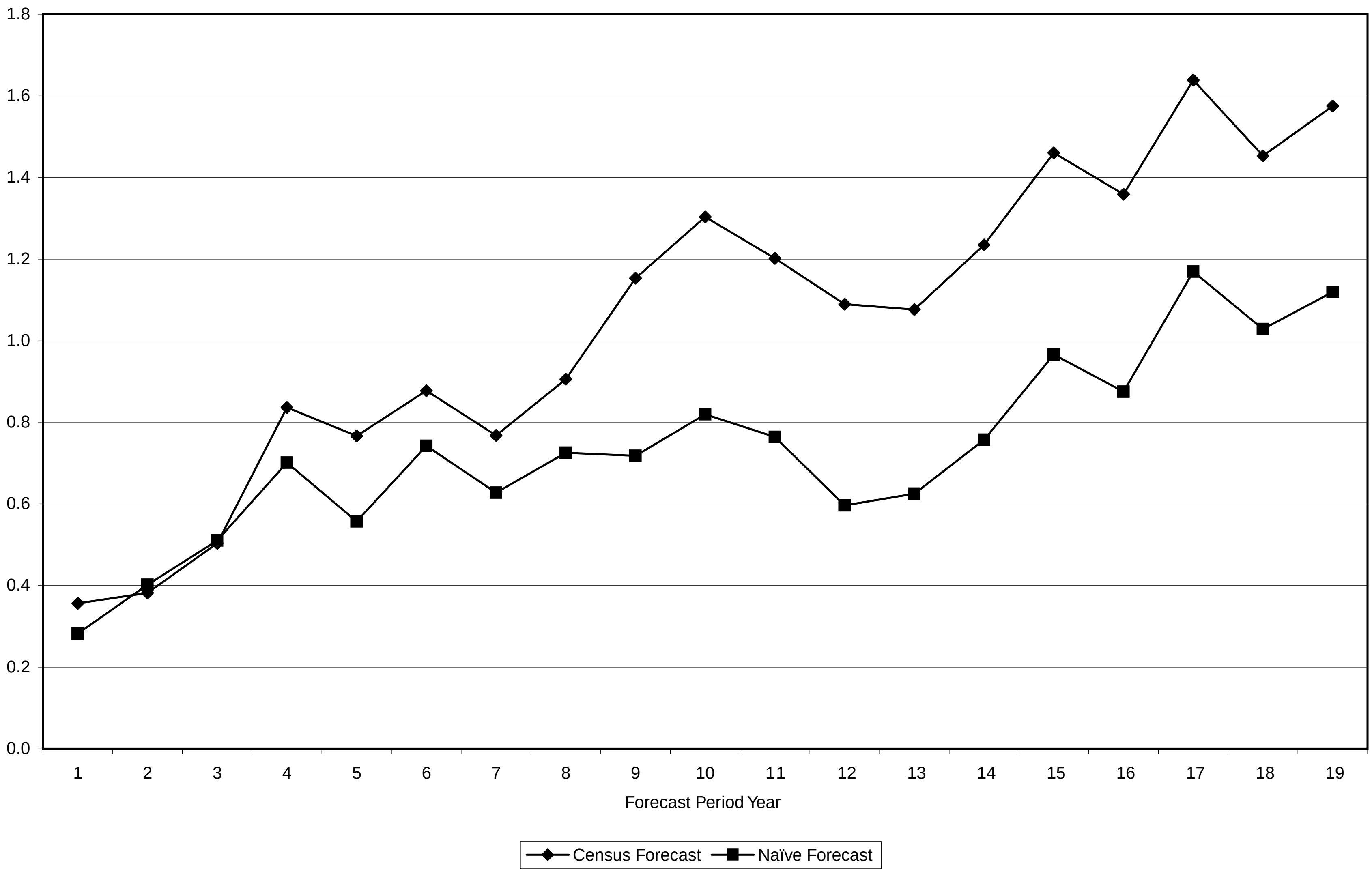
[Rate per 1,000 persons]



Source: Population Projections Program, U.S. Census Bureau: May 2000

Graph 8. Comparison of the Multiple Series RMSE for the Crude Immigration Rate for the Total Population of the United States

[Rate per 1,000 persons]



Discussion

Hania Zlotnik

This paper presents a careful and thorough evaluation of the series of population projections for the United States population prepared by the U.S. Census Bureau. The approach taken to evaluate past projections is based on the availability of a complete series of independent estimates of the parameters considered. This point needs to be underscored. One of the most difficult tasks in evaluating the performance of population projections is to obtain reliable estimates of the indicators being projected. In the case of the United States, estimates of fertility and mortality indicators are generally reliable though not necessarily free of error. Estimates of net international migration, in contrast, are difficult to obtain and subject to unknown levels of error. Mulder's paper indicates that past projections have been particularly deficient in projecting net international migration. This deficiency is largely due to the fact that, at the time many of those projections were being prepared, the estimates of international migration that Mulder uses as basis for comparison were not available. It is only with the greater availability of data on international migration that began in the late 1980s and the 1990s that researchers have been able to establish retrospectively annual estimates of international migration (see, for instance, the work of Bean, Passel or Warren). That is, those evaluating projections have the unfair advantage of having access to estimates that were not available at the time the projections were made and by this I mean not only estimates referring to periods beyond the base year of the projections but also, and more importantly, to estimates referring to periods before the base year of the projections. It is clear from the comparisons presented by Mulder that such was the case of early projections, since the trends the projections extrapolate seem to start at the wrong level because current estimates (i.e. those referring to the base year) were probably not available. Lack of current data at the time when the projections were prepared can account therefore for an important part of the "projection error" detected. In fact, I wonder whether the fairly good agreement that Mulder finds between more recent projections and the estimates available for the 1990s will persist once the estimates for the 1990s are adjusted in light of the results of the 2000 Census and the ACE survey.

Another important conclusion that can be drawn from Mulder's analysis is that the use of sophisticated forecasting methodology based on statistical models does not necessarily perform better than other methods. Although Mulder does not describe in detail the differences in the projection methodology used in preparing the various projection sets, she does comment that in preparing the 1986 projections the Box-Jenkins time series model was used to make short-term forecasts of fertility. Yet, in 1986 the forecast error over the first five years is higher than in other projections for the 1980s and 1990s. One problem with the use of time-series methods is that they are based on the assumption that the data vary cyclically and actual data may not conform to such a model of change.

In fact Mulder concludes that a "naive" model of forecasting tends to perform better than other methods used. It must be noted that although for projections elaborated before 1990 that seems to be the case, the errors associated with the "naive" model are not necessarily small. Furthermore, Mulder's data suggest that projections have been getting better and that recent projections have outperformed the "naive" model.

Lastly, it should be noted that Mulder assesses the accuracy of projections by comparing the values of demographic indicators such as the crude birth rate, the crude death rate, the number of births and the number of deaths with estimated numbers. Although it is interesting to know what level of error affects these indicators, it would also be important to have information on the errors associated with indicators such as total fertility and the expectation of life which, aside from being independent of the age structure of the population, are normally the indicators used to formulate assumptions about future trends. If an extension of the paper were possible, inclusion of error analysis with respect to those indicators would be very useful.

Jeffrey Passel

It is true that almost invariably the turning points were missed. But so were periods of stabilization. It seems to take a long time to even recognize that there was a turning point.

Some of the charts show a declining birthrate trend, yet the projection still shows an upward trend. The base data are often not up to date.

I want to commend the Bureau for its most recent projections in the area of immigration. I think the degree to which the last set was revolutionary was really not recognized. Not only did the Bureau switch to modeling individual components of migration, but it also tried to take into account the changes built into the legal system. I think this approach is an important beginning. The next step is to investigate the causes of the errors.

There are statements in the paper that the "naive" model was more accurate at various points in the past. But it is not always clear to me what is the naive model—or what it should be. Is it constant trends or constant levels or what?

The statement is made that doing projections by race is a necessity. But Social Security does not use race—so it is not a necessity. I feel the Census Bureau uses race in the belief that this will improve the performance of its projections.

It may be more important in the future to include generation as a variable—that is, to separate immigrants by first, second, or higher order generation of birth. The Census Bureau did do their current projections by nativity, but incorporated no native/foreign-born fertility differentials.

The paper understates the problems of measurement errors in both the historical data and in the development of the base population. In fact, I will take issue with the notion that the projections of the 1990's were much more accurate—I don't think that will be true after taking the 2000 Census into account. The problem is that we correct the error in our standard, but we have no way of correcting the inputs to the projections.

Passel concludes that the Census Bureau may have underestimated international migration, primarily undocumented flows. It also overestimated emigration. But this was probably not fixable in the projections which used reasonable assumptions when they were made. He has a hard time determining the extent of the errors because of Mulder's use of crude net immigration rates—which he hardly ever sees used. However, he estimates that the most recent immigration projections were off by about 50%.

Why do we miss turning points? And why take so long to recognize they were missed. Also we should define what we mean by a "naive" model. It may be time to look for other variables besides race, such as generation or class. Finally, there is a need to focus more on problems of measurement and error in both estimates and projections.

Ahlburg, *Chair*

The Mulder paper continues to use the tried and true measures of accuracy. We ignore the fact that there is a whole world of forecasting out there where different measures are used—people care more about absolute level of error than the percent error as measured in the MAPE. He thinks we need to look at these alternatives. Also why do we use this strange cohort-component thing that no one else does? Keilman also made a point in a book about the importance of controlling for the difficulty of the forecast—we may not be getting better—it may be the world has just been more stable? Ahlburg is surprised that the naive model was not a 3 or 5 year average. He is not surprised by the failure to project turning points as he argues that we never are looking for them. He has seen nothing written on leading indicators in literally decades. Ahlburg argues for combining series to improve forecasts. Stan Smith has done some good work on this in small-area forecasts. Ahlburg says he doesn't know who our "experts" are who advise us. We need to structure the meetings with experts—just sitting around drinking coffee won't work. Scott has shown that high school kids do better than experts. We need to find out how to use experts better. Mulder implies a simple model will outdo the cohort-component model. Ahlburg predicts we won't use a simple trend model in our next projections.

Floor Discussion

Goldstein

Just a comment that a "unitless absolute measure" sounds like a contradiction.

Ahlburg

There are some modifications of the classic ones we use. Some papers by Armstrong present a whole array of options. The best may be the relative mean absolute percent error, which is just a modification of the MAPE.

Mulder

There are also problems with different indicators in terms of the unit of analysis—how large is your population and how does it respond to error. Has been work with "M" estimators, but a lot more needs to be done.

Bongaarts

He enjoyed the paper, but wonders why there was no discussion of error in terms of population size—That is, how close do past projections come to the census 2000 total of 280 million? He thinks we probably did pretty well—not as bad as the impression given by the individual component tables. He also thinks the difference between baseline error and growth error was not explicated well enough. He wonders why there was no evaluation of the total fertility rate—only crude events. Similarly, nothing was done on life expectancy. Also we should say more about the interaction of these errors—there is no discussion of how there often are offsetting predictions of too many births but not enough migrants or deaths. Finally, he would like more information about how the projections were made.

Mulder

Measures such as TFRs and life expectancy are often not available in historical series. She just didn't have the time to construct these measures from the data in the older projections.

Masnick

One of the tensions in his mind is long-term versus short-term. We don't want to be embarrassed in the long run, but we tend to get evaluated in the relative short-run. He doesn't know the resolution—maybe we need to aim at the middle term. He suggests that just evaluating these projections across all different durations probably doesn't help us do a better job.

McDonald

You might think of error as the cost of making an error for each person. This is why we do cohort-component—because we get most age groups right for next 5 to 10 years. Errors are usually much bigger for the young and the old, so maybe cost of an error would be a useful way to think about error.

Coleman

He is interested in Ahlburg's idea of model-based simple forecasts, even things like the logistic model. He is also interested in Lutz' probability-based forecasts and if Mulder thinks it is a reasonable way to attach an error range to forecasts. Coleman was also surprised to hear that race did not make a difference in accuracy. Finally, what did you do with mixed-origin people in our projections.

Mulder

She didn't use race because only had white and black groups until very recently. She has no idea where Coleman got the impression that she looked at racial groups.

Passel

The Census Bureau did not recognize mixed race persons before the 2000 census. Those who reported mixed-race in earlier data were placed in a single-race category by use of a number of different rules.

Haub

He thinks ethnicity is a very important factor, especially Hispanics as their fertility is so high—in some cases even higher than in the sending countries.

Passel

Most non-immigrants are only here temporarily and shouldn't be included in our data. But groups such as foreign students and those with H1-B visas should be included.

Van de Kaa

Mulder could have given more attention to the different forecast requirements of the components. People have little control over mortality so projections tend to be pretty accurate. International migration should be hard to do because most people within the country aren't interested in it—it mostly depends on behavior of people outside of your country. Fertility involves only a small part of the population, so it is intermediate between the other components. Maybe these relative differences should be taken into account in your evaluations?

He also thinks that the components are interrelated—a decline in mortality is accompanied by a decline in fertility, and so on. In Europe today, the very low fertility has led to increased migration. Van de Kaa thinks that this is also happening in the U.S. and that it might be worth study.

Keilman

This is a very useful update for an area where we know a lot more about conditions in Western Europe. He recommends that Mulder look at errors in age structures as they are more important. He suggests that she ignore Ahlburg and don't worry about measurement of errors. Keep it simple. Let the user know how far off your projections were for various age groups.

In earlier work he found high positive errors for the young because of over-estimated fertility—also negative errors for the old because of underestimated mortality improvement. Is this still true?

This work is extremely important because it gives us an idea of the size of historical errors. Such information will serve as a benchmark for assessing the expected errors produced by other types of forecasts.

Keilman then commented that use of crude birth rates will not yield very different errors from those of the TFR for quite a long-time—but life expectancy errors will quickly yield very different errors from crude death rate errors. Actually there cannot be much error in life expectancy at birth—he would like to see a study of error at age 60 or 80.

Passel asked why it takes so long before we recognize turning points that have been missed. In the 1970s Ascher introduced the idea of "assumption drag" which explains it. It takes about 10 years for demographers to recognize a change in trend because we have to wait that long to be sure. This is because of the extrapolative nature of our assumptions

Ahlburg says we never forecast turning points. That is because we never look for them. I think a careful analysis of period and cohort data might lead to turning points—for example if the age at first birth changes direction.

John Long did an analysis showing that target TFR levels were strongly correlated with current levels. Does this still seem true for the U.S. in your more extensive data set?

Spar

Have you done any work looking at the relationship between demography and changes in the economy? I am thinking of the work of Easterlin that the Census Bureau experimented with at one time.

Mulder

Yes, at the state level. Our current state projections include an alternative domestic migration series which uses Bureau of Economic Analysis forecasts of Gross Domestic Product.

O'Connell

The trouble with adding economic variables is then you have to forecast them also.

Spar

He suggested that such difficulties would require us to work much more closely with other Federal agencies.

Tuesday Afternoon

Session II

U.S. Fertility in International Comparison: An Exploration to Aid Projections

Tomas Frejka
(Sanibel, Florida)
and
W. Ward Kingkade
(U.S. Census Bureau)

CONTENTS

Abstract

Introduction

1. Fertility trends

1.1 The United States

1.2 The United States fertility trends in international comparison

1.3 Age patterns of fertility

1.3.1 Completed fertility of cohorts from the 1930s to the 1950s

1.3.2 Fertility age patterns of most recent cohorts

1.4 U.S. age patterns in international comparison

1.5 Parity distribution in the U.S.

1.6 Parity distribution in international comparison

1.7 Childlessness

1.8 Parity progression

1.9 Central findings

2. Fertility differentials

2.1 Introduction

2.2 Fertility trends of the White and non-White U.S. populations

2.3 Fertility differentials of cohorts concluding childbearing at the end of the 20th century

2.3.1 The principal race, ethnic and nativity differentials

2.3.2 Income and educational attainment characteristics of U.S. women

2.3.3 Fertility differentials by income

2.3.4 Fertility differentials by education

2.3.5 Fertility differentials by income and education

2.3.6 Fertility differentials by nativity, education and income

2.3.7 Income and educational differentials in childlessness of racial and ethnic groups

2.3.8 Childlessness differentials by education and income

2.4 Principal findings

3. Differentials in planned and unplanned pregnancies and births

3.1 The core of the issue

3.2 Unintended pregnancies and unintended births: U.S. differentials

3.2.1 The overall view

3.2.2 A more profound insight

3.2.2.1 Unintended pregnancy rates

3.2.2.2 Intended pregnancy rates

3.2.2.3 Mistimed and unwanted pregnancy and fertility rates

3.2.2.4 Contraceptive failure rates

3.3 Pregnancy and fertility rates: U.S. trends

3.4 Pregnancy and fertility rates: U.S. in international comparison

4. The evidence and hypotheses elucidating why U.S. fertility is relatively high

4.1 Proximate determinants affecting U.S. fertility

4.1.1 U.S. contraception practices and the provision of services in international comparison

4.1.2 Induced abortion practices in the United States in international comparison

4.1.3 Marriage and cohabitation

4.2 Indirect determinants affecting U.S. fertility

4.2.1 Race, ethnicity, nativity and culture

4.2.2 Education

4.2.3 Income

4.2.4 More comprehensive reflections on economic factors

4.2.5 Religion

4.2.6 Physical abuse of women

4.2.7 Lack of inter-generational communication

4.3 Summary

5. Brief conclusions

References

Tables

Figures

Discussion

U.S. Fertility in International Comparison: An Exploration to Aid Projections

Tomas Frejka
Sanibel, Florida
and
W. Ward Kingkade
U.S. Census Bureau

Abstract

The joint objectives of this study are to investigate the causes for relatively high United States fertility and to provide insights aiding the projection of fertility.

In the last quarter of the 20th century U.S. fertility trends have been relatively stable: Total period fertility fluctuated between 1.8 and 2.1; estimated completed fertility for cohorts born in the 1950s and 1960s was around 2.0; cohorts that were at the onset and in the middle of their reproductive careers were experiencing fertility patterns distinct from other Western developed countries—higher and more stable. This implies U.S. fertility could remain relatively high for some time.

There were many differentials: Fertility of White non-Hispanics was high by European standards and that of Blacks 15 percent and Hispanics 35 percent higher than the former; the poorest segments of any racial/ethnic group had comparatively high fertility and even more so those with less than a high school education. Fertility of the wealthy and well educated differed very little between racial/ethnic categories. U.S. women had high rates of unintended, mistimed and unwanted, pregnancies and births. Hispanics had twice the rate of unintended pregnancies of Whites and Blacks three times higher. Hispanics had twice the rate of planned pregnancies of Whites and Blacks. Rates of unintended pregnancies changed very little during the 1980s and early 1990s.

High rates of unplanned pregnancies compared to other Western countries were in part caused by the relatively deficient family planning delivery system characterized by reliance on medical specialists, expensive services, inadequate effective choice of contraceptives, and limited sex education. A relatively low level of functional literacy among the young and least educated coupled with various manifestations of poverty, especially among the Blacks and Hispanics, such as grave existential concerns, could be another part of the explanation for high rates of contraceptive failure. None of the causes for high fertility are likely to change rapidly. In sum, the analyses point in the direction of stability or a slight fertility decline.

Introduction[1]

At the end of the 20th century fertility of the United States population measured by the total period fertility rate (TPFR) was unquestionably the highest among the low-fertility countries with a generally comparable social and economic history and conditions[2] (see Table 1 and Figure 1). In 1999, compared to the average of 18 "West" European countries it was higher by 29 percent, and in comparison to the other three large low-fertility overseas countries, Australia, Canada and New Zealand, it was higher by 17 percent. The comparatively high level of U.S. fertility is a phenomenon of the last 15 years of the 20th century. Prior to that, during the

[1] This document is longer than a standard paper. The senior author, when given the assignment early in 2001, considered it essential to deal with the numerous issues that are covered below. Even so many topics remain that would have provided additional insights, but have not been dealt with.

[2] The U.S. TPFR was also significantly higher than that of a number of low-fertility populations of East and South-East Asia, such as that of Japan, China, Taiwan, Hong Kong, South Korea, Thailand and Singapore. International comparative analyses in this paper will be made with the European and overseas "developed" populations and not with those of East and South-East Asia.

1970s, it was below the average level of the other low-fertility countries (see Figure 1). Since the mid-1970s the TPFRs in the other countries were on a continuous moderate downward slope, while in the U.S. fertility was quite stable. From 1974 to 1986 the TPFR in the United States was around 1.8 children per woman. In the late 1980s it increased to 2.1 by 1990 and has fluctuated mildly above 2.0 throughout the 1990s.

The study seeks to make a modest contribution to clarifying how this situation developed taking a U.S. perspective. There is a substantial body of literature that explores the causes of low fertility in the developed countries during the last quarter of the 20th century (for instance, Foster 2000, Lesthaeghe and Moors 1996, McDonald 2001, Van de Kaa 1987) much of which is relevant for the United States. Thus far, however, no attempt has been made to explore causes for relatively high fertility in the United States. Given the vast amount of issues that would need to be dealt with in order to present a comprehensive analysis, our study is undoubtedly unbalanced and incomplete. Furthermore, because the primary purpose of the Conference is to conduct analyses helpful in preparing fertility projections, the guiding principle for including topics and analyses was our judgment with respect to their relevance for such projections. This paper makes no pretense to being a comprehensive review and analysis of U.S. fertility determinants. It is a selection of topics and analyses the results of which might help to obtain a better understanding for likely future directions of fertility and thus might be helpful in preparing projections. In sum, the paper has a two-fold objective: 1. Investigate causes for relatively high fertility in the United States; and 2. Provide insights aiding the projection of U.S. fertility.

The first section illustrates various aspects of U.S. fertility of the recent past and how these differ from the other developed countries. The analysis deals with fertility trends, changes in the age patterns of fertility, parity distributions and parity progression ratios. Even though basically descriptive, this section provides initial insights on the causes for the relatively high U.S. fertility. Furthermore, a certain understanding is gained as to the degree of stability of U.S. fertility in the near future.

The second section examines racial, ethnic, income, and educational differentials of the U.S. population, and some international comparative analyses are explored. It is largely descriptive, nevertheless the exposition reveals further insights into the factors contributing to elevated U.S. fertility. The analysis in the first two sections provides a solid base for the sections which follow.

The third section focuses on the extent to which conceptions and births are intended. The relatively high proportions of unintended, mistimed and unwanted, pregnancies and births among various age, racial, ethnic and socio-economic groups of the U.S. population are a central link in the chain leading to the explanation, admittedly partial and incomplete, of why U.S. fertility is high. At the same time, there were also racial/ethnic groups with relatively high rates of intended births which contributed to high U.S. fertility.

Finally, section four provides selected explanations for relatively high U.S. fertility. The interpretations and elucidations arrived at in this section of the paper will range from reasonably firm conclusions to hypotheses and possibly even speculations.

It is mainly the last two decades of the 20th century that will be analyzed in detail. However, to understand developments at the end of the century a longer-term time horizon helps to put these developments into perspective. When appropriate, and depending on data availability, trends since the middle of the century or even for the entire century will be explored. Data suitable for international comparison are available mostly for the second half of the century.

1. Fertility trends

1.1 The United States

The 20th century is marked by considerable fertility fluctuations. During the first 3 decades a steep decline took place from a total period fertility rate of close to 4 children per woman around the year 1900 to a trough in the 1930s of about 2.2 children per woman in the early 1930s (see Figure 2). Given the mortality conditions at the time, this was below the replacement level. The notorious baby-boom ensued. By the year 1957 its peak was reached with a TPFR almost as high as at the beginning of the century, 3.8 children per woman. As rapidly as fertility increased, the TPFR declined during the next 20 years and reached 1.8 by 1976.

During the last quarter of the 20th century U.S. fertility has been remarkably stable. From 1975 to 1986 at around a TPFR of 1.8 followed by a moderate increase to 2.1 in 1990. Throughout the last decade of the century the TPFR has been between 2.0 and 2.1 (see Figure 2).

Completed cohort fertility followed a similar path, albeit at more moderate levels. Women born at the beginning of the century had on average about 2.5 children. The total cohort fertility rate (TCFR) was at its lowest for the cohorts born between 1905 and 1912 at 2.3 children per woman. It then increased steadily and the cohorts of the early 1930s bore on average 3.2 children per woman. Subsequently the TCFR declined to reach its lowest level ever at 2.0 with the 1950 cohort.[3] The cohorts of the 1950s will apparently remain at that level as might the 1960s cohorts, as indicated by the analysis below, for instance, on the age patterns of fertility of cohorts born during the 1960s.

Throughout there was a considerable fertility differential by racial origin (Figure 3). Among the cohorts of the early 1930s the White population had a TCFR of 3.1 compared to the non-White population with a TCFR of 3.8 children per woman, a difference of about 25 percent. Subsequently, fertility of the two basic ethnic groups declined almost in parallel to reach 1.9 for the White population and 2.5 for the non-White population among the cohorts of the early 1950s. The differential had, however, increased to over 30 percent. Among the cohorts of the 1950s and early 1960s, total cohort fertility for the Whites was essentially stable, but among the non-Whites cohort fertility was declining steadily. A more detailed exposition of White and non-White fertility levels and trends is pursued in section 2. Fertility differentials.

1.2 The United States fertility trends in international comparison

The basic differences in period fertility trends were illustrated in the *Introduction*. For the sake of simplicity, only cohort fertility trends are compared in the remainder of this section. The long-term trends in the U.S. population starting with the cohorts born early in the 20th century through those born in the 1940s were quite similar to trends in Canada, Australia and New Zealand. Starting with the cohorts born around 1950 what stands out is the stability of U.S. fertility, whereas in the other countries there was a continued decline (see Figure 4).

Compared to Western Europe, the baby boom cohorts born in the 1920s and 1930s in the U.S. had considerably higher fertility (see Figure 5). The outstanding difference among the cohorts of the 1950s and early 1960s is a continued fertility decline in West European countries compared to the stable trend in the U.S.

In Southern Europe, the cohorts of the 1930s and 1940s experienced a fertility decline similar to the comparable cohorts in the United States (see Figure 6). Contrary to the U.S., this decline continued with the cohorts born during the 1950s and early 1960s.

It was only in the populations of the Nordic countries that fertility levels and trends of the cohorts born during the 1950s were similar to those of the United States (see Figure 7). In Sweden and Finland, however, signs of an incipient decline were evident.

1.3 Age patterns of fertility

1.3.1 Completed fertility of cohorts from the 1930s to the 1950s

Underlying the completed cohort fertility trends, whether declining or unchanging, were considerable changes in the age structures.

As shown in Figure 2, the TCFR declined between the 1930–31 and the 1940–41 cohorts from 3.2 to 2.7 children per woman. Fertility change was, however, not identical at all ages. Actually, fertility of young women up to age 24 in the 1940–41 cohort was relatively high compared to the older cohort, but this gain was more than offset by relatively low fertility once these women were between the ages of 24 and 49 (see Figures 8 and 9 and Table 2). In the 1940–41 cohort compared to the 1930–31 cohort, fertility of women in their late twenties was lower by more than 30 percent and that of women in their thirties was lower by 50 percent. The fertility surplus[4] early in the reproductive period of the 1940–41 cohort was outweighed by the fertility deficit later in life so that the net deficit was 0.5 of a child (see Table 2). The average age of cohort childbearing declined from 26.2 to 24.7 (see Figure 10).

The fertility decline continued among the cohorts born during the 1940s from 2.7 in the 1940–41 cohort to 2.0 in the 1950–51 cohort. The age structural changes for these two cohorts a decade apart were, however, very different than among the previous two cohorts. Young women were delaying childbearing. When in their teens and twenties, the 1950–51 cohort had almost 0.8 children less than women ten years older (see Figures 8 and 9 and Table 2). In particular, fertility in the prime childbearing ages between 20 and 24 was more than 43 percent lower. There was a slight shift of fertility into the higher ages. When these women were in their thirties and forties they had

[3] Minor proportions of total cohort fertility rates (TCFRs) of those born in the 1950s and early 1960s were estimated. For details see Frejka and Calot (2001).

[4] Changes in the age structure of cohort fertility can be observed by comparing age-specific fertility rates of one cohort with that of another. In this study as a rule cohorts born 10 years apart are compared. When the age-specific fertility rates of a cohort born later (a younger cohort) is *higher* than that of a cohort born earlier (an older cohort), the difference is considered a *surplus*. When the age-specific fertility rates of a cohort born later (a younger cohort) is *lower* than that of a cohort born earlier (an older cohort), the difference is considered a *deficit*. Frequently whole age ranges in a younger cohort tend to be either higher or lower compared to the older cohort. Thus, for instance, when comparing the U.S. cohort born in 1940–41 with that born in 1930–31, the age group 14–23 experienced a surplus – fertility had increased in this age group; and the age group 24–49 experienced a deficit – fertility had declined (See Table 2).

slightly more children than the cohort ten years older—0.1 of a child—but by far not enough to catch up. Altogether, the average age of childbearing increased quite considerably from age 24.7 to 26.0 (see Figure 10).

Completed cohort fertility did not continue to decline among the cohorts of the 1950s. The TCFR of the 1960–61 cohort was equal to that of the 1950–51 cohort at 2.0 children per woman, but the age structure did change (see Figures 8 and 9 and Table 2) as did the average age of childbearing. When these women were young their fertility was relatively low, mainly between the ages of 18 to 20. On the other hand, later in life mainly once they were in their thirties, their fertility was somewhat higher. The fertility increase of these cohorts of women when they were older was sufficient to compensate for their fertility deficit earlier in life, ipso facto, the TCFR remained the same. The average age at childbearing among the cohorts of the 1950s continued its increase from 26.0 to almost 27.0 (see Figure 10).

1.3.2 Fertility age patterns of most recent cohorts

The cohorts born during the 1960s and 1970s, those that were at the onset or in the middle of their reproductive periods as of the late 1990s, appeared even more stable in their reproductive behavior than the older cohorts (see Figure 11 and Table 3). The age patterns of fertility of these cohorts differ very little from one another. This implies that fertility could remain roughly at the same level, i. e. around, or only slightly below, replacement in the United States in the foreseeable future. As a matter of fact, there is the possibility of a fertility increase, as cumulated fertility rates up to the known ages for the cohorts of the mid-1960s to the mid-1970s were consistently on a moderately increasing trend.

A detailed examination of cumulated fertility of cohorts born from 1950 through 1975 confirm the above. Figure 12, which depicts only registered data, demonstrates that the 1960–61 compared to the 1950–51 cohort had accumulated a fertility deficit of about 0.14 children by age 27. After that age women of the 1960–61 cohort were having somewhat more children than the cohort born ten years earlier at the same ages. As a result by age 37 the former cohort had almost eliminated the deficit accumulated earlier in life and the cumulated cohort fertility rates of the two cohorts were almost identical. Adding the estimated fertility rates for women of the 1960–61 cohort when they will be in their late thirties and forties results in identical total cohort fertility rates for these two cohorts ten years apart.

The cohorts of the 1960s, exemplified by the 1965–66 cohort, follow the fertility age pattern of the 1960–61 cohort quite closely, albeit with minor deviations. Apparently the cohorts of the late 1960s and the early 1970s, as far in life as they have come up to the year of observation, are experiencing moderate increases in fertility compared to the 1965–66 birth cohort (see Figure 12).

1.4 U.S. age patterns in international comparison

Changes in the age patterns of the U.S. population of the 30 cohorts of the 1930s through those of the 1950s were similar to those in other western countries, yet there were also many differences.

The principal similarity was that among the birth cohorts of the 1930s—the peak baby-boom mothers—there was a continuous shift of fertility into the younger ages, whereas in the subsequent cohorts the fertility shift was reversed and, gradually and continuously, more children were born by women when they were in their late twenties and thirties in practically all the western countries (see Figure 13).

One significant difference was that among the 30 cohorts of the 1930s through the 1950s the age patterns of fertility were younger in the U.S. than elsewhere. Larger proportions of children were borne by younger women than in other countries. There were several other notable differences between the age patterns of fertility of the U.S. and most other western countries.

The proportion of total fertility that was realized before exact age 27 in the United States increased from 60 percent in the 1930–31 cohort to 71 percent in the 1940–41 cohort, an increase of 19 percent. The increase was even larger in most of the other countries. In Norway, for instance, this proportion increased from 46 to 60 percent, between the 1930–31 and the 1940–41 cohorts, an increase of 33 percent. In no other country, however, was the proportion of children born before exact age 27 as high as in the U.S. In the other countries this proportion was around 60 percent or less compared to the U.S. 71 percent (see Figure 13).

In subsequent cohorts fertility shifted to the older ages. In the 1960–61 cohort in the United States barely one half of all children were borne by women in their teens or early to mid-twenties. Nevertheless, this proportion was still comparatively higher than in the other western countries. In the U.S. the proportion of children born before exact age 27 in the 1960–61 birth cohort was 53 percent compared to a range of between 42–47 percent in the other countries (see Figure 13). In the U.S. only 47 percent of all children were borne by women in their late twenties and in their thirties and forties. In all the other western countries higher proportions of children were borne by women after exact age 27. In Norway it was 56 percent, in Western Germany 58 percent and, for instance, in the Netherlands it was as many as 69 percent (not shown in Figure 13).

The magnitude of the shift in fertility from the younger to the older ages was considerably smaller in the U.S. compared to the other western countries. This can be observed, for instance, when comparing the fertility age patterns of the 1960–61 and the 1950–51 cohorts in a number of countries. The fertility decline among the women up to exact age 27, the fertility "deficit" among young women, was much smaller in the United States than in any other country (see Table 4). This deficit was 0.14 children in the U.S., whereas in many other countries it was about double that size and as high as almost four times that size in New Zealand.

The other outstanding difference was that in the U.S. the entire fertility deficit was compensated when women reached their late twenties and thirties, whereas in other countries only around 50 to 65 percent of the fertility deficit was compensated when the respective women were older. One extreme case was Italy with only 17 percent of the deficit replaced (see Table 4). On the other hand, in the Nordic countries the degree of compensation was similar to the U.S. as exemplified by Norway.

Fertility behavior of young women in the United States continued to be remarkably different compared to other Western countries among the cohorts born during the 1960s (see Table 5). In the other countries the decline of fertility experienced by young women born during the 1950s continued, whereas in the U.S. the cumulated cohort fertility rate (CCFR) of the 1970–71 cohort was moderately higher than that of women born in 1960–61. Consequently, by exact age 27 U.S. women of the 1970–71 cohort bore on average 1.1 children, whereas, for instance, in the other non-European countries 0.6 children and in Italy as few as 0.4 children.

The relative stability of U.S. fertility is also apparent among the cohorts of the mid-1970s which initiated their childbearing careers in the 1990s (see Table 6). Whether the comparison is made over 25 or 15 cohorts, it is obvious that the youngest cohorts in the United States have trends and levels different from the other low-fertility countries.

In comparison to the 1950–51 cohort, the CCFR up to exact age 22 of the 1975–76 cohort in the U.S. was lower by 12 percent, whereas in the other countries the decline was at a minimum 40 percent and as high as between 60 and 70 percent. Compared to the 1960–61 cohort U.S. fertility of the 1975–76 cohort increased by 7 percent and in all the other countries this cohort experienced a decline, on average of over 30 percent.

The difference between the U.S. and other countries is notable in the absolute numbers. The 1975–76 cohort in the United States had borne 0.5 children per woman, whereas in the other countries it was between 0.1 and 0.2 children (see Table 6).

1.5 Parity distribution in the U.S.

The parity distribution of women age 40[5] in the United States underwent a major change between the cohorts born in the early 1930s and those born around 1950. Among women (couples) of the 1930 birth cohort there was a clear preference for large families. Thirty-seven percent of women had 4 or more children and an additional 22 percent had three, for a total of almost 60 percent having more than 3 children. In contrast, in the 1950 birth cohort only 31 percent of all women had 3 or more children. The largest decline was for parity 4+, from 37 to 11 percent (see Figure 14). Family size preferences had clearly changed. The proportion of women (couples) with 2 children had increased from 22 percent in the 1930 cohort to 35 percent among women born in the late 1940s. There was also a notable increase in the proportions of women with only 1 child and those remaining childless.

The parity distribution remained remarkably stable among the cohorts of the 1950s. The parity distribution of the 1950 and the 1960 birth cohorts were almost identical. Over one third of all women had 2 children, another third had 3 or more, and among the remaining one third those with only one child or no children were roughly evenly distributed (see Figure 14).

1.6 Parity distribution in international comparison

The main difference between the U.S. parity distribution of the 1960 birth cohort and that of selected West European countries is that there are relatively more women with four or more children in the United States (see Table 7). If one takes all women with three or more children, the proportion of such women is practically the same as in Norway and Finland.

[5] The parity distribution of women age 40 rather than 50 are considered appropriate for the present research purposes because this enables the inclusion of those born around 1960 in the time series. It is justified by the fact that parity distribution between the ages of 40 and 50 changes only marginally as no more than 1 to 2 percent of all children were born after age 40 in all the cohorts born between 1930 and 1960.

Another difference is that U.S. women have the smallest proportion of parity two. U.S. women are roughly in the middle of the range of parity one proportions and at the lower end of childlessness.

In terms of trends, the U.S. parity distribution experienced relatively large changes between the 1930 and the 1950 birth cohorts in comparison to most other countries for which sufficient data were available (see Figures 14 to 19). Trends of the U.S. parity distribution have also been more stable than in any of the other countries for the cohorts born between the late 1940s and that of 1960.

In all the Western countries, women with two children represent the largest proportion (see Figures 14 to 19). In most countries the peak of this proportion was reached either by the 1945 or by the 1950 cohort and it has been declining among subsequent cohorts. Compared to most of the other countries, in the U.S. this decline has been almost imperceptible so far.

1.7 Childlessness

Intentional or voluntary childlessness has again become a subject of thorough inquiry as part of the concern with declining and low fertility in the developed countries (for instance, Dorbritz, Schwarz 1996; Foster 2000; Golini 1998; Rowland 1998). Therefore a separate small section on this issue is included.

In the United States women remaining childless represented around 10 percent of the cohorts born during the 1930s (see Table 8 and Figure 14). Among the cohorts of the 1940s there was a perceptible increase of childless women, from 11 percent in the 1940 birth cohort to 17 percent in the 1950 cohort. For subsequent cohorts the proportion of childless women remained stable.

In other western countries the percentage of childless women among the 1960 birth cohort was at a similar level as in the United States (see Table 8). The outstanding difference between U.S. women and those in most of the other western countries was that elsewhere the proportion of childless women appeared to be on an upward trend whereas U.S. cohorts apparently reached a plateau from around the 1950 birth cohort (see Table 8 and Figures 14 to 19). In some countries, such as Norway, the increase in the share of childless women was moderate and gradual. In others, the increase has been quite rapid. In the Netherlands this proportion increased from 11–12 to 20 percent between the cohorts of the early to mid-1940s to those born around 1965. In Finland and West Germany the percentage of childless women more than doubled from the cohort of 1945 to that of 1965, and between 1940 and the 1960 birth cohort, respectively. In Italy this proportion was not very different from one cohort to the next, but at a rather high level, between one sixth to one fifth of all women remaining childless (see Table 8 and Figure 17).

1.8 Parity progression

Parity progression ratios for the United States illustrate the fact that there was a decline of the ratios to the first, second, third and fourth births between the cohorts born in the early 1930s to those born in the early 1950s (see Figure 20). Parity progression ratios of the cohorts born during the 1950s have been quite stable. The ratios from parity zero to parity 1, and from parity 2 to parity 3 even displayed mild increases.

There are only two "West" European countries with parity progression ratio data for comparison, Finland and the Netherlands (see Figures 21 and 22, and Table 9). Restricting the comparison to the cohorts born in the late 1950s and early 1960s, as a rule, U.S. women have higher parity progression ratios, except for the progression from parity 1 to parity 2 (see Table 9).

1.9 Central findings

At the turn of the 21st century U.S. fertility was relatively high compared to the other western type developed countries. It was the result of divergent trends. In the United States fertility had been quite stable during the last 10 to 20 years of the 20th century, whereas in most of the other western countries it was declining.

This is demonstrated by exploring various facets of fertility. Total period and cohort fertility have been relatively stable in the U.S. and declining almost everywhere else. Among cohorts born during the 1950s that have reached the end of their reproductive periods, declining fertility of women when they were in their teens and early twenties was offset by sufficiently high fertility when these women were older, not so in most other developed countries. In the younger cohorts, those that are in the middle or at the beginning of their reproductive life who were born during the 1960s and 1970s, in the United States fertility has been level or even slightly increasing, everywhere else it has been declining. Among the youngest women, as shown for women born in the mid-1970s and below age 22, fertility in most of the other countries was at levels so low that it can hardly decline more. Among these women fertility in the U.S. was at least twice as high, and as much as ten times higher than elsewhere.

The relatively high and stable fertility of the U.S. cohorts born during the 1960s—cohorts which had already realized approximately one half of their eventual childbearing—was an indication that U.S. fertility is likely to remain among the highest in the developed countries in the near future.

In terms of clues for fertility projections it appears useful to follow several different aspects of fertility behavior. The mapping of fertility behavior of cohorts which are at the onset or in the middle of their reproductive periods seems to be among the most relevant information for fertility projections. Important knowledge is generated by comparing the experience of birth cohorts within countries and by making international comparisons.

2. Fertility differentials

2.1 Introduction

In section 1 on U.S. fertility trends, the difference between the White and the non-White population was illustrated (see Figure 3). As there is a general belief that U.S. fertility distinguishes itself from that of other developed countries in large part because of racial and ethnic fertility differentials, this avenue of exploration will now be pursued.

To begin with, the fertility levels and trends of the White and non-White U.S. populations will be analyzed in sub-section 2.2. The analysis is based on the classic cohort fertility calculations prepared and published by R. L. Heuser[6] (1976) and an extension prepared specifically for this study by W. Ward Kingkade.

Subsequently, sub-section 2.3 will deal not only with racial and ethnic fertility differentials, but also with those of nativity (native- or foreign-born), age, income and education, including cross-classifications of these characteristics. The obvious reason for doing so is to explore the importance of relevant associations of fertility and the respective characteristics. The data used in this section are those collected by the United States Bureau of the Census approximately every two years in the June Supplement to the Current Population Survey.[7]

2.2 Fertility trends of the White and non-White U.S. populations

As referred to above, trends of total cohort fertility rates (TCFRs) of the White and non-White populations experienced similar directions during the 20th century, but started to diverge with the cohorts of the 1950s and early 1960s (see Figure 3). The former stabilized, whereas the TCFRs of the non-White population continued to decline. Underlying these trends were differing changes in the age patterns of fertility (see Table 10). At the same time, the changes of the fertility age patterns between cohorts of both the White and the non-White populations were in similar directions (see Figures 23 and 24).

The lifetime fertility pattern of the non-White population was much younger. By age 27 in the 1950 cohort, for example, 68 percent of all children had been born in the non-White population, compared to 60 percent in the White population. In both populations fertility shifted into the older ages, but the non-White childbearing age pattern continued to be significantly younger. In the 1960 cohort approximately 61 percent of children in the non-White and 51 percent in the White population were born before age 27 (these numbers are not shown in Table 10).

Between the cohorts born around 1930 and the 1940 cohort there was a fertility increase early in the reproductive period in both populations and a considerable decline of fertility when the 1940 cohort was in its late twenties and thirties (see Figures 23 and 24). Among subsequent cohorts fertility was declining when they were young and increasing when they were older; the fertility increase at the older ages was more substantial in the White than in the non-White population.

Differences in the mean ages of childbearing for the respective populations were considerable and also illustrate that the directions of change were similar but not identical (see Figure 25). The shift of childbearing into older ages between the 1940 cohorts and those of the early 1960s was more pronounced among the

[6] The composition of the U.S. population changed over time and especially during the past two decades became more heterogeneous. In 1980, the non-White population constituted 14.3 percent of the total and over 84 percent of that were Blacks. In 2000 non-Whites constituted 17.5 percent of the total, and the proportion of Blacks in this category declined to 74 percent. Among the White population in 1980, 85.7 percent of the total, the proportion of Hispanics was 7 percent, whereas in 2000, when Whites constituted 82.5 percent of the total, it had increased to 11 percent (http://eire.census.gov/popest/estimates.php; and the Census 2000 Summary file).

[7] The Current Population Survey (CPS) has been conducted by the U.S. Census Bureau since the early 1940s for the principal purpose of measuring the national unemployment rate and characteristics of the labor force. The CPS is based on a sample designed to represent the population of the United States and that of each state. Monthly state and national estimates of the population by age, sex, race and Hispanic origin are taken into account. Sample housing units are interviewed for 4 consecutive months, twice over a 16-month period. Currently, most CPS interviews are conducted by computer-assisted telephone interviewing, although personal interviews may still be conducted in some instances. In addition to basic data on the labor force and unemployment, the CPS collects information on a range of supplemental topics of national interest. This includes the periodic, usually every two years, June Fertility Supplements, which contain birth histories, responses on birth expectations, and characteristics of children for women ages 15–44. The June 1998 Fertility Supplement is the source of the data described in the present report.

White than among the non-White population. The difference in the mean age of childbearing between the two was about half a year in the 1940 cohort and over 1.5 years in the cohorts of the early 1960s.

There was a significant difference between the age patterns of fertility of the White and the non-White population comparing the 1950 and the 1960 cohorts. Among the Whites the shortfall or deficit of fertility when the cohorts were young was fully compensated when this cohort was older. Among the non-White population the corresponding deficit was compensated only to a minor degree (see Table 10). By age 27 the White population had 0.165 less children than the 1950 cohort, but after age 27 the 1960 cohort had 0.175 more children than the previous cohort. The 1960 cohort caught up fully and more (106 percent) with the 1950 cohort despite low fertility early in the reproductive period. The non-White population experienced a decline of 0.296 children when young, but had only 0.086 children more when older; not even one third (29 percent) of the deficit early in the reproductive period was compensated when these women were older.

We now turn to the cohorts which are at the onset or in the middle of their childbearing period, the cohorts of the 1960s and the 1970s (see Table 11). The fertility differential continues to be large between the White and the non-White population. By age 27 the 1970 White cohort had on average borne 1.0 child per woman, the non-White 1.4 children. The White vs. non-White differential by age 22 for the 1975 cohort is even larger. While there were some differences in the rates of fertility change by these ages, they were relatively small, especially when compared to other countries. The decline of the cumulated cohort fertility rates (CCFRs) by age 27 between the 1960 and 1970 cohorts was about 4 percent for both U.S. populations, compared to 17–40 percent for the populations selected in Table 5.

The closeness of the curves in Figures 26 and 27 illustrates the relative stability of fertility during the 1980s and 1990s, in particular of the White population.

The curves in Figures 28 and 29 are the cumulative cohort fertility differences between the cohorts 1956 through 1976, and the base 1951 cohort. Figure 28 illustrates the comparatively low fertility of the 1956 and the 1961 White cohorts when they were young and the catching up to the base cohort when they were in their late twenties and thirties. The relatively low fertility of the 1956 and 1961 non-White cohorts was more pronounced. By age 27 the 1961 non-White cohort had borne 0.3 children per woman less than the 1951 cohort (see Figure 29 and Table 10). As the 1956 and the 1961 non-White cohorts are aging, a tendency of catching up with previous, older cohorts can be observed, but it appears practically impossible that they will succeed.

The cumulative fertility of the 1960s and the 1970s cohorts, White and non-White appeared to be relatively stable. The 1971 cohorts by age 27 displayed minor declines. The CCFR of the 1976 non-White cohort by age 22 declined slightly and the 1976 White cohort experienced an increase (see Figures 26, 27, 28 and 29 and Table 11).

Trends of the long-term parity distributions displayed in Figures 30 and 31 reveal changes that were probably unique among the developed countries. For the purposes of this study, however, we will draw attention mainly to recent developments.

The older White cohorts, including those of the 1940s had experienced major changes in the parity distribution (see Figure 30). Starting with the cohorts born around 1950 stability not seen before has been established. In the cohorts of the 1950s and early 1960s, around 35 percent of all White women had two children, the proportions of women with one, or three children were between 15 and 20 percent, as were those of childless women, and about 10 percent had 4 children or more.

The long-term parity distribution trends of the non-White population were quite extraordinary (see Figure 31). A degree of stability could be observed among this population starting with the cohorts of the 1950s. There were around 32–33 percent of non-White women with 2 children, between 16 and 24 percent of women with one, three, or more than 4 children. The proportion of childless women even in the cohorts of the mid-1960s was under 10 percent, but displaying an increase.

The parity progression ratios of White women also displayed relative stability among cohorts born during the 1950s and early 1960s (see Figure 32). About 82–83 percent proceeded to have a first child and of these slightly under 80 percent had a second child. Among non-White women, the progression to the first and to the second birth among the cohorts of the 1950s and early 1960s were on the decline (see Figure 33). Over 90 percent were having a first birth, however, of these only around 75 percent were having a second birth. Progression ratios to third, fourth and fifth order births appeared to be stabilizing and were about ten percentage points higher than among White women.

2.3 Fertility differentials of cohorts concluding childbearing at the end of the 20th century[8]

The main groups of the population with respect to race, ethnicity, and nativity in the following analysis were the White, Hispanic, Black, Asian and Pacific Islander, and Other (which include Native Americans, Alaska Natives and other groups) and Native- and Foreign-born. The exposition aspired to be relevant to the understanding of research and policy issues. Regrettably, the analysis faced some limitations due to the relatively small size of certain categories in the sample. Frequently the category of "Other" could not be included because it is not sufficiently represented when classifications of income, education, and nativity were introduced. Special care had to be taken even with other race and ethnic categories in the more detailed analysis. The actual race and ethnic categories used in this study (most of the time) are:

- White non-Hispanic" rather than "all White," because the latter category includes the majority of Hispanics, namely White Hispanics;
- All Hispanics," even though this category includes a small number of Hispanics of Black and Asian/Pacific Island origin;
- All Blacks," even though a small number of Hispanics are included;
- Asian and Pacific Islanders," even though a small number of Hispanics are included.

Regarding the other characteristics, major attention is devoted to income and education because these are among the more important ones modifying or even determining levels of living, careers, and life styles.

2.3.1 The principal race, ethnic and nativity differentials

The analysis focuses on women 40–44 years of age in 1998. That provides a picture of fertility differentials of cohorts with virtually completed fertility. This group consisted of approximately 72 percent White non-Hispanic, 13 percent Black, 10 percent Hispanic, 5 percent Asian and Pacific Islander women and one percent Others (see Table 12). Fourteen percent of all women were foreign-born and 86 percent had been born in the U.S.

On average, U.S. women in the 40–44 age group in 1998 had had 1.88 children ever born. Fertility of Black women was higher than that of the White non-Hispanic by 15 percent, Asian and Pacific Islander fertility was 16 percent higher and that of Hispanic women was higher by 35 percent.

Nineteen percent of the women age 40–44 in 1998 had remained childless. The proportion childless among White non-Hispanic women was 20 percent, compared to 15 percent among Hispanic, 17 percent of Black, and 18 percent of Asian and Pacific Islander women by the time they practically reached the end of their reproductive period.

2.3.2 Income and educational attainment characteristics of U.S. women

Prior to analyzing fertility differentials it is useful to introduce background knowledge about income and educational attainment characteristics of U.S. women. Among the White non-Hispanic women age 40–44 there were relatively few with an annual household income under $20,000, even under $30,000 (see Table 13 and Figure 34). Twenty-two percent of White non-Hispanic women lived in households with less than $30,000 compared to a full 50 percent of Hispanic women and almost as many Black women, 49 percent. A considerably larger proportion of foreign-born, 40 percent, lived in poor households (<$30,000) compared to native-born women, 26 percent.

On the other end of the income distribution spectrum, 27 percent of White non-Hispanic women were in households with an annual income of over $75,000 compared to 10 percent of Hispanic and 9 percent of Black women. Asian and Pacific Islander women had a relatively large proportion in high income households, 30 percent.

White non-Hispanic women age 40–44, the age at which income is relatively high in the life cycle, lived in households with an estimated average income of $52,000 per year compared to $35,000 in Hispanic and Black households. Women of Asian and Pacific Islander origin lived in households with an income almost equal to the White non-Hispanic women.

In comparison to other developed countries, income in the United States is higher and the income distribution more unequal than in most other developed countries. In 1998, the United States had the highest per capita GNP expressed in purchasing power parity (PPP) among the developed countries, $29,240, followed by Switzerland, $26,876 and Norway, $26,196 (U.S. Statistical Abstract, 2000). In 1995, the Gini coefficient[9] for Denmark and Sweden was about 22 and for the United States and the United Kingdom about 35 (UNECE, 2000).

[8] For a description of the data source for this section see Bachu and O'Connell (2000).

[9] Higher values indicate a more unequal income distribution.

Differences between the race and ethnicity groups are particularly evident when comparing the proportions of those with a limited amount of education, and, on the other hand, those with advanced education (see Table 14 and Figure 35). Thirty-nine percent of Hispanic women 40–44 years old had less than a complete high school education compared to 14 percent among Black women, 13 percent of Asian and Pacific Islanders, and seven percent of White non-Hispanic women. By far the highest proportion of women with more than a bachelor's degree was among the Asian and Pacific Islanders, 41 percent, compared to 30 percent among the White non-Hispanic women, 17 percent among the Black and 12 percent among Hispanic women.

A comparison of educational attainment of the U.S. population with that of other developed countries is not a simple matter. Conventional statistical data indicate that the United States population is among the best educated. In 1998, the U.S. had the lowest percentage of population with less than a lower secondary education, 14 percent, compared to 20 percent in Canada, 17 percent in Norway, 19 percent in the United Kingdom and 35 percent in the Netherlands (U.S. Statistical Abstract, 2000). In the United States, 35 percent of the population had a university or non-university tertiary education, compared to 39 percent in Canada, 27 percent in Norway and 23 percent in the United Kingdom.

In addition to how much education was attained, the quality of learning outcomes is a critical aspect when evaluating any educational system. Generally these two aspects tend to be positively correlated. Higher educational attainment generates youngsters better capable of dealing with the complexities of contemporary life. The United States is no exception to this global regularity. What differentiates the United States from many low-fertility countries of Western Europe is the lower overall quality of education among the most poorly educated and young segments of the population. These are precisely those population groups in which unwanted pregnancies are most prevalent. This topic will be dealt with in greater detail in section 4.2 *Indirect determinants,* subsection 4.2.2 *Education.*

2.3.3 Fertility differentials by income

Turning to fertility differentials by income, altogether there was no clear-cut association between income and fertility (Table 15 and Figure 36). In the larger racial/ethnic categories women whose families had the lowest income tend to have the highest fertility, but beyond that the differentials were relatively minor. In the smaller racial/ethnic groups women in families with middle income had the highest number of children, and the upper income categories had relatively low fertility.

The smallest fertility differentials by income are among the White non-Hispanic women (see Table 15 and Figure 36). The lowest income category, <$10,000, had about 10–20 percent more children than the other income groups. The next lowest income group of women, $10–20,000, had about as many children on average as women in middle and high income families. It is interesting to note that women in families with high and upper middle income have marginally higher fertility than women in the $10–30,000 annual income categories. There is an almost imperceptible indication of an inverted "J" type of association.

The nature of the income differentials among Black women is not much different from the non-Hispanic White, except that the women of the poorest families have about 30–35 percent more children than the other low, middle and high income groups.

Among Hispanic women the relationship between income and fertility was rather uneven. In general, women in families with income up to $20,000 had 20–35 percent more children than the middle and upper middle income families. Women in the $50–75,000 income category had only 10 percent lower fertility than the poorest women, but 10 percent higher fertility than women in families with $20–50,000 annual income. Women in the wealthiest Hispanic families had the lowest fertility of all, about 40 percent less children than the poorest women.

Among the Asia and Pacific Islander women those in families with middle income had by far the largest number of children. Women in families with $20–30,000 annual income had about 60 percent more children than women in families with more than $50,000.

The fertility differentials by income were the largest among the "Other" women. The poorest women had twice as many children as the wealthiest, and women in families with $20–30,000 had almost three times as many children than those with over $75,000.

The fertility differentials among foreign-born women were larger than among the native-born. Among the foreign born women those in families with income below $30,000 had about 20–30 percent more children than those in families with $30–75,000 annual income, and around 45–60 percent more children than the wealthiest women.

Fertility differentials by income of native-born women between the income groups above $10,000 were relatively small. They all had on average between 1.7 and 1.9 children per woman. In comparison, the poorest women had about 20 percent more children.

An important fact is that foreign-born women in families with $10–30,000 had over 40 percent higher fertility than native-born with the same income. We will deal with fertility differentials of native- and foreign-born women in greater detail below.

2.3.4 Fertility differentials by education

Almost without exception in all the race, ethnic and nativity groups the association between fertility and educational attainment was negative (Table 16 and Figure 37). Women with no high school education had the highest fertility and those with a graduate or professional degree had the lowest number of children.

Overall women without a completed high school education had almost twice as many children as women with graduate or professional degrees. Among the women without a completed high school education, Hispanic and Black women had the largest average number of children per woman, 3.1 and 2.8, respectively. White non-Hispanic women without a completed high school education had only 2.1 children per woman. Among women with graduate or professional degrees there were relatively small differences between the race and ethnic groups. White non-Hispanic and Hispanic women had 1.3 children per woman; the "Other" women had the lowest fertility, 1.0 child per woman; followed by Black women with 1.1 child per woman; Asian and Pacific Islander women, who had the highest number, namely 1.4 children per woman.

As with fertility differentials by income, the smallest fertility differentials by education were among the non-Hispanic White women. Nevertheless, even in this ethnic category women without a completed high school education had 60 percent more children than women with graduate and professional degrees.

Asian and Pacific Islander women had the second smallest difference between women with the lowest and highest education. Women without completed high school education had about 80 percent higher fertility. A somewhat peculiar feature among these women was the relatively high fertility of those with completed high school education and with some college education, including associate degrees after two years of college. Their fertility was only about 10 percent lower than that of women without a completed high school education and 60 percent above that of women with bachelor's and graduate degrees.

Among Hispanic women not only those with graduate degrees but also those with bachelor's degrees had comparatively low fertility. The differentials were rather large between those with a college and university education and women without a completed high school education. The latter had more than twice the number of children compared to the former.

Among Black women those without a completed high school education had 2.5 times as many children as those with a graduate or professional degree.

Among "Other" women there were two clusters. Women with bachelor's and graduate degrees had very low fertility and all other women, including those with a completed high school education and those with some college education, including associate degrees, had relatively high fertility. The latter 2.5 times that of the former.

2.3.5 Fertility differentials by income and education

The above analysis of fertility differentials by educational attainment and by income has yielded an interesting finding, namely that fertility differentials by income are considerably weaker and less systematic than fertility differentials by educational attainment. This finding is generally valid for almost all race, ethnic and nativity categories. There are, however, exceptions. A detailed cross-classification of children ever born by education and income for White non-Hispanic women demonstrates that for the whole group the differentials by income are very small, but among women with a graduate or professional degree there is a notable positive association between fertility and income. Such women with the lowest annual income of up to $20,000 had on average 0.3 children per woman, whereas the wealthiest women had on average 1.6 children (see Table 17 and Figure 38). There was a similar, albeit weaker, association among women with some college education, including associate degrees, and women with bachelor's degrees had an inverted "J" type of association.

2.3.6 Fertility differentials by nativity, education and income

In concurrence with conventional wisdom, the principal data in Table 12 demonstrated that foreign-born women had higher fertility than native-born. That was basically the case for Hispanic women (see Table 18). Furthermore, among the foreign-born and among the native-born Hispanic women the association between education and fertility was much stronger than that between fertility and income. At the same time, it was women with bachelor's degrees that had the largest fertility differential between foreign- and native-born.

Contrary to expectations, among White non-Hispanic women the foreign-born had lower fertility than the native-born (see Table 19). But the strong negative association between education and fertility, and the weak association between income and fertility held for both the native- and the foreign-born women in this racial/ethnic group.

2.3.7 Income and educational differentials in childlessness of racial and ethnic groups

The basic differentials of proportions of women without any children among the 40–44 years old between the race/ethnic groups were similar to those of children ever born. White non-Hispanic women had the highest proportion childless, 20 percent, and Hispanic women the lowest proportion, 15 percent (see Table 20). Black women and those of Asian and Pacific Islander descent were in between.

The more detailed classification by income illustrates that there was no systematic association of income and childlessness (see Table 20 and Figure 39). Among White non-Hispanic women proportions childless in the individual income categories were close to the average of 20 percent. Among Black women there was a loose positive association between income and proportions childless. Among Hispanic women the deviations from the average were large. Low income women had small proportions childless and middle and the highest income women had large proportions childless, however, the "upper middle class" had a low proportion childless. Most Asian and Pacific Islander women of different income groups had similar proportions childless with the exception of the $20–30,000 category.

The income differentials in proportions childless among native born women as a whole were not very large. There may well be distinct differentials within native born racial/ethnic groups, but for the most part our data, due to insufficient sample size, do not permit such a detailed investigation.

The associations between proportions childless and educational attainment were more systematic and the differentials large (see Table 21 and Figure 40). In general, among women with higher education large proportions remained childless. The largest differentials and the most direct association were among Black women. Among the least educated Black women about 10 percent remained childless, whereas close to 50 percent of Black women with graduate and professional degrees were without children. Among White non-Hispanic and Hispanic women the association of childlessness and education was slightly "J" shaped and among Asian and Pacific Islander women "U" shaped.

Among native born women the positive association of education and proportions childless is straightforward, whereas among foreign born women except for the women with highest education the differentials were negligible.

2.3.8 Childlessness differentials by education and income

As with the differentials for children ever born by income and by education, the differentials in proportions childless by education (Table 21 and Figure 40) are much clearer than the differentials by income (Table 20 and Figure 39). The fertility differentials are, however, somewhat different than the proportion childless differentials, in other words, the two types of differentials are not mirror images.

White non-Hispanic women was the one racial/ethnic group, where a more detailed analysis by income and by education yielded intriguing insights. The sample was of adequate size to be analyzed in greater detail.

The differentials of proportions childless by income for *all* White non-Hispanic women were small (see Table 22 and Figure 41). The range was between 19 and 25 percent childless and there was no systematic pattern. Within each educational attainment category, however, the differentials were much larger and in principle there was a negative association between income and the proportions childless. The more pronounced association of this type was among those women who either had some college education and even more so women with graduate and professional degrees. For women with bachelor's degrees, those in households with the lowest incomes had relatively low proportions of childlessness. The highest proportion was in the $20–30,000 category and from thereon the proportions declined with increasing income. Even among White non-Hispanic women with a high school education or less there was a slight negative association between proportions childless and increasing income.

2.4 Principal findings

The United States population is the wealthiest among the developed countries. Its income distribution is relatively unequal. In terms of formal educational attainment, the U.S. population is among the best, but apparently the quality of basic education leaves much to be desired. The younger population of the United States, in particular those with less than a tertiary education, exhibits lower levels of functional literacy in comparison to their counterparts in Western Europe.

Income differentials in the United States of families with women in their early forties between racial/ethnic groups and within these groups were considerable, but it is not known how these compare with income differentials in other countries. There were also major differences in educational attainment between racial/ethnic groups and again we have no comparison with other countries.

When summarizing the analysis of fertility levels and trends of the White and non-White populations one has to keep in mind that both these categories are heterogeneous aggregates. Conceivably trends in the components of the aggregates could be canceling each other out. The category "White" contains a considerable proportion of Hispanics; the latter contains not only a majority of Blacks, but also many people of Asian and Pacific Island origin.

Be that as it may, the non-Whites had considerably higher fertility than the Whites among the cohorts of the 1950s and early 1960s, and at the same time the differentials were diminishing. That process seems to have halted among the cohorts, which were in the middle or at the onset of their childbearing at the end of the 20th century. Taking the example of the 1970 cohort, by age 27 the Whites had borne 1.0 child per woman and the non-Whites 1.4. Further, in both populations there had been little change compared to the 1960 cohort. It is important to point out, that fertility up to age 27 in the 1970 non-White cohort was double or more compared to European populations and fertility of the White 1970 cohort was also higher, by 40–60 percent, in comparison to the Europeans. The differentials between the cohorts born in the mid-1970s were even much larger.

When dealing with the more detailed racial/ethnic breakdowns, the largest fertility differential was observed between White non-Hispanic women and Hispanic women, the latter had 35 percent more children ever born than the former. Fertility of Black as well as Asian and Pacific Islander women was 15 percent higher than that of White non-Hispanic women. White non-Hispanic women, which constituted almost three quarters of the total, had 1.8 children ever born which roughly compares to the average of West European total cohort fertility rates of the cohorts born around 1960 (see, for instance, Figure 5). Since the other racial/ethnic groups had higher fertility, this is one formal explanation for the overall higher fertility of the U.S. population compared to the other developed countries.

The association between income and fertility was weak. The absolutely poorest women did have more children than other income categories, but beyond that the effect of income on fertility of the cohorts at the end of their childbearing in the late 1990s was marginal. Income, for the most part, also had little effect on childlessness, although it is worth noting that among the poorest women large proportions of White non-Hispanic women and of women of Asian and Pacific Islander descent were childless.

The association of education and fertility was substantial. In the main racial/ethnic groups basically fertility declined step-wise with each higher category of educational attainment. The average number of children ever born of women with graduate or professional degrees was 1.3 and there was only a small difference between the racial/ethnic groups. Black women in this educational category had only 1.1 children ever born.

The effect of education with childlessness was also sizable. Not only in the graduate degree category but likewise in the bachelor's degree category, close to a third of Hispanic women remained childless. Almost as many women remained childless in the bachelor's degree category among White non-Hispanic women. Among Black women more than a quarter with bachelor's degrees were childless and almost one half with graduate/professional degrees.

In a more detailed analysis of White non-Hispanic women, a positive association of income and fertility within educational categories appeared, although on average for the whole age group of women 40–44 years old there is no sign of any association between fertility and income. The positive association was especially apparent among those with a graduate or professional degree, however, to a lesser extent also among women with some college education and those with bachelor's degrees. Similarly, there was a distinct negative association of childlessness and income within some educational categories, although no such association was detected among all women in the 40–44 age group. The negative childlessness and income association was again most obvious among White non-Hispanic women with graduate and professional degrees, but also among those with only partial college education.

Another intriguing finding regarding fertility differentials among White non-Hispanic women was a larger number of children ever born among native-born in contrast to foreign-born, even though generally it is the foreign born women that have higher fertility.

The various types of fertility differentials analysis conducted above can also provide clues for fertility projections. Depending on the relative size of a particular subgroup its distinctive fertility levels and trends can have important implications for overall levels and trends. Additional insights would be gained by following fertility differentials through time. This can be done with data from the Current Population Surveys from 1980 to 2000. To undertake such analyses in this study would have required more time and space.

3. Differentials in planned and unplanned pregnancies and births

The well known Davis-Blake (1956) conceptual framework which was further developed and quantified by Bongaarts (1978, 1982) documented the importance of contraception and induced abortion in determining fertility levels and trends. A number of researchers from the Alan Guttmacher Institute in New York investigated the effect of inefficient use of contraception and the prevalence of induced abortion on fertility in the United States and conducted national analyses as well as an international comparative study. The research relies on surveys and is performed with a definition of unintended (or unplanned) pregnancies which included mistimed and unwanted pregnancies and births.[10] The data are influenced by ex-post evaluations of the women involved, which casts some doubts on their objectivity. While we recognize the difficulties in distinguishing between unwanted, mistimed and intended pregnancies and births, we considered it appropriate to incorporate results of these analyses in our study, because analogous concepts, definitions and methods were used for all U.S. subgroups, and for the international comparative study of countries and Canadian provinces.

Reactions to this section are likely to differ depending on the profession, experience and beliefs of the readers. Many economists, sociologists, as well as others are likely to doubt the validity of the investigations, basically because the way in which people answer a "yes or no" question about whether a birth was planned or unplanned, mistimed or unwanted, can conceal a complex spectrum of possibilities. In the extreme, some may claim that in modern societies the likelihood of unwanted births is small, because they can be avoided by extant contraceptive options, which are backed up with the availability of induced abortion.

In sum, while we recognize the difficulties associated with the concepts and the surveys, we believe these instruments capture important events and developments in women's and couples' lives. Furthermore, the surveys conducted in different countries or provinces (Canada) utilized comparable definitions and methods. Thus, if the results differ by orders of magnitude, there is reason to consider the differentials as real, and a comparative perspective is provided.

3.1 The core of the issue

Our investigation indicates that there is a complex interplay of numerous factors generating relatively high U.S. fertility. Based on a credible body of evidence (primarily Henshaw 1998, 2001; Jones et al. 1989; Ranjit et al. 2001) one important aspect of the story establishing why U.S. fertility is relatively high can be summarized as follows.

During the 1970s, 1980s and 1990s a considerable proportion of pregnancies in the United States were unintended, 49 percent in the early 1990s (see Table 23). There were almost as many unintended as intended pregnancies. Since only about half of the unintended pregnancies were terminated by induced abortion, 54 percent in the early 1990s, a significant proportion of births were unintended, 31 percent of the total number of births in the early 1990s.

The proportion of unintended or unplanned pregnancies in the U.S. was higher by a degree of magnitude than in many other western countries. The U.S. proportion of unintended pregnancies was, for example, 60 percent higher than in Belgium, Canada and Sweden, more than twice as high compared to Great Britain and five times as high as in the Netherlands around 1980 (see Tables 24 and 25), countries with cultural, social and economic features similar to those of the United States (Jones et al. 1989). If all unintended pregnancies and births would have been avoided, in the U.S. as well as in the other western countries, levels of fertility around 1980 would have been of a similar order of magnitude in many of the countries concerned, including the United States (see Tables 24 and 25).

If one takes into consideration that a (unknown) proportion of the mistimed pregnancies and births could be considered as wanted, the basic conclusion is modified, however, the result would be similar. The U.S. total fertility rate would still be relatively high. The rate of unintended pregnancies would be somewhat diminished yet still relatively high, and in addition a relatively high rate of intended pregnancies, which would include some of the mistimed births, would be contributing to high fertility.

The basic conclusion is that U.S. fertility is relatively high in part due to the large proportion of unintended—mistimed and unwanted—pregnancies and births.

In order to understand the mechanisms that underlie the above conclusion, two types of explorations will be pursued:

> Analyze variations in the proportions and rates of intended and unintended—mistimed and unwanted—pregnancies and births in different segments of U.S. women and in comparable countries or provinces of Canada (in this section 3); and
>
> Determine, hypothesize and/or speculate about, the reasons for the above variations (in section 4).

[10] Unintended (unplanned) pregnancies are estimated on the basis of data on unplanned births from the National Surveys of Family Growth and estimated numbers of induced abortions. Unintended pregnancies are calculated as the total of induced abortions plus births to women who had not wanted children at that time (mistimed births) plus births to women who had wanted no more children ever (unwanted births) (Cf. Henshaw 1998).

3.2 Unintended pregnancies and unintended births: U.S. differentials

3.2.1 The overall view

Unintended pregnancy and unintended birth differentials can be appraised by a number of characteristics: age, marital status, income, race and ethnicity[11] (see Table 23).

Age differentials are considerable. The highest proportion of unintended pregnancies was among teenagers, i.e. women 15–19 years of age, 78 percent—which was 60 percent above the average for all ages (49 percent). Their proportion of unintended births among all births was 66 percent, more than double the average, because teenagers ended relatively few unintended pregnancies by induced abortions.

Women in their early twenties also had a high rate of unintended pregnancies—59 percent. Because they terminated 55 percent of these unintended pregnancies by induced abortion, their proportion of unintended births was much lower than that of teenagers, 39 percent, still above average.

The lowest proportions of unintended births were among women in their late twenties and early thirties—around 20 percent. Their proportions of unintended pregnancies were relatively low and, in addition, they terminated relatively many of these by abortion.

Unintended age-specific pregnancy rates (pregnancies per 1000 women of the respective age group) were the highest for the older teenagers, i.e. those 18–19 years old: 105 per thousand, and for women aged 20–24 years old: 96 per thousand.

Pregnancy rates by age enable the calculation of life-time numbers of unintended pregnancies, which is the total unintended pregnancy rate.[12] If the rates of the early 1990s were experienced by a birth cohort of women, they would on average have had 1.4 unintended pregnancies per woman during their reproductive period.

The differentials by *marital status* are what one would expect. The proportion of unintended pregnancies of never-married women is 2.5 times that of married women and the proportion of unintended births is 2.7 times that of married women.

A number of important features are revealed by the *income* differentials.[13] The proportion of unintended births among low-income women is more than twice that of women with above median income, 45 versus 21 percent. The latter not only have a considerably smaller proportion of unintended pregnancies, but are also more motivated to terminate these by induced abortion. The overall pregnancy rate of the low-income women is twice that of women with relatively high income, 144 versus 71 pregnancies per 1000 women of reproductive age.

Low-income women also have the highest rate of intended pregnancies, however the differential in comparison to women with higher income is relatively small, 33 percent. It is the difference in the unintended pregnancy rate which is substantial. It is three times higher than among women with above median income (see Table 23; 88.3 and 29.2, respectively, in last column).

Race differentials are also notable. Black women have an overall pregnancy rate 65 percent higher than White women. The intended pregnancy rate of Black women is below that of White women and that of women of other races, 38 versus 47 pregnancies per 1000 women (White and other) of reproductive age. On the other hand, the unintended pregnancy rate of Black women is almost three times higher than that of White women, and more than twice that of women of the other races.

The proportions of unintended pregnancies and unintended births among Hispanic women do not differ from the average of other women. This applies also to the proportions of pregnancies terminated by induced abortions. Note that pregnancy rates of Hispanic women are about 70 percent above that of other women. The differential for intended births is even slightly higher than for unintended ones. The high rate of intended pregnancies among Hispanic women stands out in comparison to White and to Black women, 74 versus 47 and 38 per 1000 women of reproductive age, respectively.

[11] All of these characteristics are important and useful. It is, however, regrettable that the investigators concerned with unintended pregnancies and unintended births did not examine the relationships with educational attainment, given its critical role in the nature of fertility differentials as discussed in section 2.

[12] The total (unintended—mistimed and unwanted—or intended) pregnancy rate is analogous in its concept and construction to the total fertility rate. The value corresponds to the number of pregnancies a woman would have during her reproductive life if she were to experience the 1994 age-specific pregnancy rates. This differs from the general pregnancy rate (last three columns in Table 23) which is the total number of all pregnancies divided by the total number of women 15 to 44 years old and thus dependent also on the age structure.

[13] The three income categories are defined on the basis of the federal poverty level at the time of the interview. In 1994, the federal poverty level was $17,020 for a family of four. The first category is below that level, the second between $17,020 and $34,040 and the third above $34,040.

It is noteworthy that the total pregnancy rate was almost identical for Black and Hispanic women, 137 compared to 143 per 1000 women of reproductive age, respectively. The rate of unintended pregnancies among Black women, however, was considerably higher than among Hispanic ones (over 40 percent), whereas the rate of intended pregnancies was almost double (94 percent) for Hispanic women compared to Black women (more on that below in 3.2.2).

In sum, the *highest proportions* of unintended births and unintended pregnancies were among women 15–19 years old, never-married, with income below the poverty level ($17,020 for a family of four in 1994), and Black. The *highest rates* (per 1000 women of the respective category) of unintended pregnancies were also for the never-married, poor and Black women, however, these rates were the highest for women 20–24 years old.

3.2.2 A more profound insight[14]

This section is based on cross-tabulations of age-specific (unintended—mistimed as well as unwanted—and intended) pregnancy rates and total pregnancy rates by income and race/ethnic categories. White women constitute 72 percent of the total, Black women 13 percent and Hispanic women 11 percent (a category of "Other" which constitutes the remaining 4 percent is not taken into account). Among White women over 72 percent are in the highest income category and only 11 percent in the lowest; Black women have 40 percent in the highest category and 35 percent in the lowest; finally, among Hispanic women 38 percent are in the highest income category and 32 percent in the lowest

3.2.2.1 Unintended pregnancy rates

Altogether Hispanic women had twice as many unintended pregnancies as White non-Hispanic women and Black women had three times the number compared to White women (see last column of Table 26).

Among Black women the incidence of unintended pregnancies was quite close to the average in all income categories. Among Hispanic women there was a clear negative association of unintended pregnancies with increasing income (see Table 26). White women in the lower income categories had twice the number of unintended pregnancies compared to the majority in the higher-income category.

The even more detailed breakdown by age, income and race/ethnicity reveals that within age groups the association of unintended pregnancy rates with income tends to be negative and the highest incidence of unintended pregnancies is in the 20–24 age group among all the race/ethnic groups (see Table 25). There are a few exceptions. For instance, in the 20–24 age group Hispanic women in the higher-income group also have a high incidence of unintended pregnancies; and the poorest Black women in the 25–29 age group have the lowest unintended pregnancy rate.

3.2.2.2 Intended pregnancy rates

The intended pregnancy rates are similar for White non-Hispanic and for Black women (see Table 26). If anything, they are slightly lower for Black women. Among White and Black women the intended pregnancy rates are also fairly evenly spread out over the income categories, i. e. there is no clear-cut association between the incidence of intended pregnancies and income.

Intended pregnancy rates for Hispanic women are almost twice as high compared to White and Black women. Just as with unintended pregnancies, there is a negative association between the incidence of intended pregnancies and increasing income among Hispanic women.

Among White women intended pregnancy rates are the highest in the 25–29 age group as well as among women age 20–24 in the two low income categories and in the higher-income group of women age 30–34 (see Table 27). Among Black women the highest intended pregnancy rates are in the 20–24 and the 25–29 age groups.

In sum, these data provide some additional insights. For instance, the associations between unintended as well as intended pregnancies and income, especially among Black women, are similarly weak as associations between income and fertility (see section 2.3.3 above). Further, the differentials in age intensity of intended pregnancies between the different race/ethnic groups of women are proof of the late age schedule of fertility among White women and early age fertility schedule of Black and Hispanic women.

3.2.2.3 Mistimed and unwanted pregnancy and fertility rates

It is instructive to investigate the composition of unintended pregnancies when divided into unwanted and mistimed. Presumably women feel more profoundly about the undesirability of unwanted pregnancies than that of mistimed pregnancies. In the perception of women, the latter pregnancies occur earlier than they would have otherwise, whereas the former were never

[14] The calculations for this section were provided by S. K. Henshaw of the Alan Guttmacher Institute in New York at the specific request of the authors. We wish to express our gratitude for preparing these data. Note that the income categories are identical to section 3.2.1, but the race/ethnic groups correspond to the analysis in section 2 and differ somewhat from the categories in section 3.2.1; the White and the Black in the former include Hispanic women, whereas in this section Hispanics are excluded from these groups.

supposed to have happened at all. Furthermore, it is even more enlightening to explore how women decide about the outcome of either type of pregnancies, whether to carry to term or terminate them.

In all race/ethnic groups and income categories more than half and up to three quarters of the unintended pregnancies were mistimed ones (see Table 28). Black and Hispanic women had higher proportions of unwanted pregnancies than White women, around 40 per cent compared to about 30 per cent of the latter. The poorest women tended to have the highest proportions of unwanted births.

That about as many mistimed as unwanted pregnancies end as abortions is a counterintuitive finding (see Table 29). Theoretically it could be assumed that most mistimed pregnancies would not be terminated, because women want to have these children—only later. In reality this is not so. In the United States, judging by the proportions of mistimed pregnancies ending in induced abortions, many women consider these pregnancies a burden so severe that they resort to abortions. These proportions are as large, or even larger than, among women with unwanted pregnancies. Among White women slightly less than half of mistimed pregnancies are terminated compared to over half of the unwanted ones. Among Black and Hispanic women higher proportions of the mistimed pregnancies compared to the unwanted ones are terminated; among the mistimed and the unwanted pregnancies it is about one half or more that end as abortions.

Finally, mistimed fertility rates were higher than the unwanted fertility rates among all the race/ethnic and income categories (see Table 30). The relative magnitudes of the fertility rates of the various categories were not very different from those of the pregnancy rates, because most of the proportions of pregnancies ending in abortions were within a relatively narrow range (see Table 29).

3.2.2.4 Contraceptive failure rates

Nineteen percent of U.S. women after starting to use reversible contraceptives became pregnant during the first two years. That is the principal finding of a research project conducted with a sample of over 13,000 (Ranjit et al. 2001:24). The findings of the project demonstrate failure rate differentials by race, ethnicity, union status, income and length of use of contraceptives. Basically these are in line with what one would expect given the differentials in the proportions of unintended pregnancies by the same characteristics discussed above. For instance, around 30 percent of women with income below the poverty line of *all* race/ethnic subgroups—Blacks, Hispanics and Whites—experience contraceptive failure. Also, contraceptive failure rates are high among all income categories of Black women. For those with incomes 100 to 250 percent above the poverty line it is 30 percent and even for those with incomes 250 percent and more above the poverty line the failure rate is 23 percent.

3.3 Pregnancy and fertility rates: U.S. trends

The situation in the early 1990s can be compared with that of the late and the early 1980s (see Table 31). On the aggregate level, trends of unintended pregnancies and unintended births can be measured either with the general (per 1000 women 15 to 44 years old) or with the total rates (sum of the age-specific fertility rates). The general unintended pregnancy (and fertility) rates are given in the first line in Table 31. According to these, the general unintended pregnancy rate hardly changed during the 1980s, and declined considerably during the early 1990s—by about 17 percent. If, however, the 1981 total unintended pregnancy rate calculated on the basis of the age-specific data in the table is compared with the 1994 total rate, no decline is observed. The total unintended pregnancy rate was 1.4 pregnancies per woman in 1981 and 1994, respectively. A similar relationship is valid for the unintended fertility rate. In other words, the decline of the general rates is due to changes in the age structure of women in the childbearing ages, not due to changes of the age-specific rates.

3.4 Pregnancy and fertility rates: U.S. in international comparison

A highly regarded comprehensive international comparative inquiry was carried out in the late 1980s using data and information mainly from the late 1970s to the early to mid-1980s. The volume reporting on the results provides data on births, induced abortions and pregnancies by intention collected from vital statistics and surveys. A small number of country and province (Canada) studies exploring in detail factors influencing contraceptive use, namely laws and policies, service delivery and information delivery, was conducted as a part of the project. Some aspects of this work might be outdated, nevertheless they are informative and relevant for our purposes.

Two tables provide an initial orientation. Among nine western countries with available data the United States had the highest total pregnancy rate and the second highest total unplanned pregnancy rate (see Table 24). The U.S. total unplanned pregnancy rate was 1.5 times higher than the unweighted average of the other countries and it was about 60 percent higher than in Belgium, Canada and Sweden, over twice the rate in Great Britain and almost five times higher than in the

Netherlands. In contrast, the United States had the median total planned pregnancy rate, i.e. its total planned pregnancy rate equaled the unweighted average of the other countries. For the U.S. this measure was actually lower than in Denmark, the Netherlands, Sweden, and Great Britain.

Table 25 summarizes the pregnancy, abortion and fertility rates in the countries/provinces where the more profound investigations were undertaken. Similarly as in the comparison with the larger number of countries, the United States had the highest total unplanned pregnancy rate and the highest total pregnancy rate. In comparison with the units under consideration, the United States together with Quebec had the lowest total planned pregnancy rate.

The U.S. total fertility rate was not much higher than that of the other entities because about one-third of the unintended pregnancies were terminated by induced abortions. This proportion was higher than in any of the other countries—the U.S. had the highest total abortion rate (see Table 25). Despite the high total abortion rate, its total unplanned fertility rate was still the highest.

The above data are from the early 1980s, however, as observed above, the total unplanned pregnancy rate in the United States hardly declined between the early 1980s and the early 1990s. It remained at 1.4 unintended pregnancies per woman. Detailed calculations based on data in Table 31 show that this rate was 1.443 in 1981, it increased to 1.548 in 1987 and subsequently declined to 1.407 in 1994.

This analysis implies that a "proximate" cause of high fertility in the United States was the high rate of unintended pregnancies and unintended births, possibly coupled with a moderate level of intended (wanted) pregnancies and births, some of which were contained among mistimed pregnancies and births. In the early 1980s, if all unintended pregnancies and births had been prevented, the United States total fertility rate would have been equal to or lower than in most other western countries. The total planned pregnancy rate in the U.S. was of a similar order of magnitude as in most of the other countries. Also the wanted pregnancy and fertility rates might have been relatively high if large proportions of the mistimed pregnancies and births would have been wanted pregnancies and births. This does not appear to be the case because of the high proportion of mistimed births that were terminated. And the fact that the level of unintended pregnancies has not diminished over time is an important aspect to watch closely.

The relative magnitude of intended and unintended pregnancies and births by different groups (race/ethnic, income, age) and the changes in time add useful knowledge for fertility projections. The interaction of contraception and contraceptive failure with induced abortion the outcome of which are intended and unintended births provides another aspect for understanding fertility trends. The analyses distinguishing between mistimed and unwanted pregnancies and births provide more refined knowledge.

4. The evidence and hypotheses elucidating why U.S. fertility is relatively high

For the purpose of organizing the various reasonably documented or hypothesized factors modifying U.S. fertility, the conceptual framework of direct (proximate) and indirect (social, economic, cultural, ethnic, religious, etc.) determinants will be loosely applied (Davis and Blake 1956; Bongaarts 1978 and 1982). First, relevant aspects of the proximate determinants will be examined, mainly those regarding contraceptive behavior, as well as abortion practices, together with a brief note about marriage and cohabitation. Subsequently a number of different indirect fertility determinants, which might have been important in shaping U.S. fertility, will be explored.

4.1 Proximate determinants affecting U.S. fertility

4.1.1 U.S. contraception practices and the provision of services in international comparison

The evidence supporting the findings which follow can be found in Jones et al. (1989). The qualitative analysis of the "country" studies in that volume reveals the public health circumstances influencing the relatively high incidence of unintended pregnancies and unintended births in the United States.

In the 1980s the proportion of women of reproductive age using the three main modern effective methods of contraception—IUDs, hormonal contraception and sterilization combined—was relatively low in the United States compared to the two Canadian provinces, the Netherlands and Great Britain. There was also a relatively high proportion of women in the U.S. who did not use any method of contraception.

Use of modern methods of contraception requires medical intervention thus reliance on these methods depends on access to appropriate services. Factors that play an important role are cost, proximity and familiarity. In principle, larger proportions of women in the U.S. are at a disadvantage in comparison to women in the other countries with respect to these factors. The main reason is that contraceptive/family planning services as part of the general health care delivery system in the U.S. are provided through medical specialists (mainly obstetricians and gynecologists, or at specialized clinics for the disadvantaged population, see below) rather than being integrated in the primary health care services and delivered by general or family practitioners.

The United States was less successful than other countries in providing appropriate services and counseling to first time users, which are typically young people.

In contrast to the U.S., health care services (and contraceptive health services) in the other countries are relatively inexpensive, at times free of charge. Therefore there is no need for special services for the economically disadvantaged. Family planning clinics in the U.S. exist primarily to serve the low-income population. Many people, especially the young, find them unattractive and assume they provide a lower standard of medical care.

The effective choice of contraceptive methods in general, but especially for low-income strata, is relatively narrow in the United States.

The dissemination of information about contraceptives in the U.S. is more limited than in other countries. Apparently this is at least in part due to more limited sex education in schools, less information about contraceptives in the media, and less tolerance in attitudes about sex.

The Jones et al. volume (1989: 224) comes to the following overall conclusion:

> In sum, how family planning services are provided seems to have a substantial impact on the pattern of contraceptive practice. The service delivery system in the United States is different from other Western countries in ways that make it less conducive, on the whole, to use of the modern, highly effective methods of contraception. In addition, readily available information about contraceptive methods and services is lacking in the United States, especially simple, objective materials in the mass media. The high U.S. incidence of abortion and unplanned births can be attributed at least partially to these circumstances.[15]

4.1.2 Induced abortion practices in the United States in international comparison

In the mid 1990s induced abortions were more widely practiced in the United States than in any other western country, even though the prevalence of abortions had declined considerably since the early 1980s (by more than 20 percent). Among 14 western developed countries with available data, the United States had the highest total abortion rate, 0.69 abortions per woman of reproductive age (Henshaw et al. 1999 a,b). The unweighted average of the other 13 countries was 0.39 abortions per woman of reproductive age, i.e. the U.S. rate was 75 percent above that of the other countries.

As demonstrated above, this high U.S. abortion rate was not sufficiently powerful to offset the high rate of unintended pregnancies in the United States. Despite the widespread application of induced abortions, fertility was relatively high in the U.S.

4.1.3 Marriage and cohabitation

Exposure to conception by dwelling in a sexual union (marriage, consensual union) tends to be an important factor in determining the level of fertility (Bongaarts 1978 and 1982). In the developed countries at the end of the 20th century, given the diversity in family formation patterns and the extent of extra-marital sexual activities, its importance in modifying fertility has probably diminished.

Even though the importance of this factor might be questionable, a crude comparison of the U.S. with other developed countries is presented. The total first marriage rate (TFMR), a measure which can be applied to compare the general nature of marriage patterns between countries, will be used. In 1985, the latest year with data available for the U.S., the TFMR was considerably higher than in the other western developed countries with the exception of Greece and Portugal (Sardon 2000). The unweighted TFMR for 21 other developed countries was 658 per 1000 women compared to 776 in the United States, a difference of 18 percent.

Furthermore, the average age of entry into unions is lower in the United States than in other developed countries, particularly among Hispanics and Blacks. Also pre-union fertility among these segments is probably higher than in the other countries.

4.2 Indirect determinants affecting U.S. fertility

The basic criterion for including a topic—factor or mechanisms—in this section was that it has a relatively significant impact on one or several proximate determinants of fertility and thus on fertility and on the fertility

[15] The Alan Gutmacher Institute conducted a comprehensive cross-national study *Teenage Sexual and Reproductive Behavior in Developed Countries* during 1998–2001. The project updates a significant part of the presented analysis and provides numerous insights and conclusions relevant to the understanding of relatively high U.S. fertility (Darroch et al. 2001 a and b, and Singh et al. 2001). Results from the AGI study were published in December 2001; too late for them to be incorporated in the present report.

differentials analyzed in previous sections. An important additional consideration was that the respective factor/mechanism is instrumental in assessing the differences between U.S. fertility levels and trends and those of other countries. Almost as a rule the discussions are very concise and not exhaustive and frequently specific relationships are hypothesized and advanced as a conjecture. Thus this section is not a thorough and comprehensive treatment of the full range of indirect fertility determinants. It would be impossible to cover so much in one study. Numerous important topics, such as gender roles, the changing status of women, family values, international migration, are not even mentioned.

4.2.1 Race, ethnicity, nativity and culture

There is an extensive literature on racial and ethnic fertility differentials, however, in a recent authoritative review paper it is stated that "(L)ittle is known about the mechanisms through which race and ethnicity influence fertility behaviors." (Forste and Tienda, 1996).

The most recent data continue to show that on average Hispanic, Black as well as Asian and Pacific Islander women had higher fertility than White non-Hispanic women. Furthermore, foreign-born Hispanic (and possibly also Black as well as Asian and Pacific Islander foreign born) women had considerably higher fertility than native born. Simplistic explanations would include the consideration that these communities have relatively high fertility because their proportions of poor and uneducated are relatively high. In addition, traditions of relatively high fertility in the recent past are built into the value systems of these communities.

Those are no doubt factors contributing to relatively high U.S. fertility. It is, however, important to note that the fertility of White non-Hispanic women was as high, usually higher than that of other developed country populations to begin with. The relatively high fertility of the other racial and ethnic groups, in particular of the foreign born, was on top of the relatively high fertility of White non-Hispanic women.

The association of high fertility and racial/ethnic background disappears with advanced education (see section 2). Fertility among highly educated Hispanic women is the same as that of White non-Hispanic women, and fertility of Black women with graduate or professional degrees is lower than that of Hispanics and White non-Hispanic women. Among women with bachelor's degrees, fertility of Black women is the lowest and that of White non-Hispanic women is the highest, with Hispanics in between, but altogether the differentials are small. In other words, advanced education appears to override cultural fertility differentials between racial/ethnic groups.

4.2.2 Education

The examination of fertility differentials of the U.S. population demonstrated the conspicuous negative association between education and fertility. The notable negative association was valid for all racial/ethnic groups, and was particularly pronounced for the Blacks. Fertility of well-educated Blacks is about one half compared to those with a high school education and even less compared to those without such education. Almost one half of Black women with graduate or professional degrees remained childless. It would be useful to explore the reasons for this state of affairs in a comprehensive fashion.

Since the apparent attainment of formal education in the United States was as high or higher than in the other developed countries one would expect fertility in the U.S. to be as low or lower than elsewhere. The quality of learning outcomes and the competence needed to cope adequately with the complexities of everyday life, particularly among the younger and lesser educated segments of the population appear to be a factor more important than formal education with respect to contraceptive behavior. In a major cross-national study of adult functional literacy these segments of the U.S. population scored poorly relative to their counterparts in Western Europe. This might be an important circumstance explaining high rates of contraceptive failure and high rates of unintended, mistimed and unwanted, pregnancies and births in the United States.

Our assessment of cross-national differences in the quality of education is based on the results of the International Adult Literacy Survey (IALS) developed and managed by Statistics Canada and the Educational Testing Service of Princeton, New Jersey[16] (OECD, 2000). The Survey was conducted in 20 countries[17] in the 1990s using a standard instrument developed originally in English and translated under careful controls into other languages as appropriate for the participant countries. The instrument differentiates three domains of "Adult Literacy": Prose Literacy, Document Literacy, and Quantitative Literacy. *Prose Literacy* is defined as "the ability to understand and use information contained in various kinds of text," measured by using selections from newspapers, magazines, and

[16] This is the organization responsible for development of the well-known Scholastic Achievement Tests (SATs) and Graduate Record Examinations (GREs), as well as a multiplicity of more specialized educational tests.

[17] These were: Australia, Belgium(Flanders), Canada, Chile, the Czech Republic, Denmark, Finland, Germany, Hungary, Ireland, the Netherlands, New Zealand, Norway, Poland, Portugal, Slovenia, Sweden, Switzerland, the United Kingdom, and the United States.

brochures. *Document Literacy* is defined as ability to process information in documents such as schedules, charts, tables, maps, and forms; success involves ability to locate information in a variety of displays and, if necessary, transfer information from one source to another. *Quantitative Literacy* is defined as ability to perform arithmetic operations involving data and instructions embedded in documents of various types; one such task at the median difficulty level requires respondents to look at two graphs containing information about consumers and producers of energy and calculate the amount of total energy in quadrillion BTUs consumed by Canada, Mexico, and the United States.

The detailed findings of the IALS reveal that U.S. respondents ages 20–25 with less than a college or other tertiary-level education score consistently lower than their counterparts in West European countries such as Sweden and the Netherlands. These findings hold across all domains under standard criteria of statistical significance, but are especially pronounced among the least educated respondents (those who have not completed high school). Among respondents with a tertiary-level education, the United States does not fare worse than the majority of IALS participant countries, although United States respondents do remain outperformed by respondents in Sweden.[18] A similar gradation of results obtains with respect to age. United States respondents ages 16–25 are consistently outperformed by respondents of the same age bracket in the vast majority of European participant countries. The only consistent exceptions are three formerly socialist countries of Eastern Europe (Hungary, Poland, Slovenia), whose scores frequently do not differ under conventional standards of statistical significance from those obtaining for the United States.

The IALS went to great lengths to standardize instrumentation and coding procedures. Its methodology can surely be improved upon in the future, but for the moment it represents an admirable effort to address the vital, albeit somewhat sensitive, topic of international variability in adult functional literacy.

The principal implication we draw from the results of the IALS is that the educational systems of most West European countries appear to impart a higher quality of elementary and secondary education than the educational system of the United States. As a result, IALS respondents at the high school and pre-high school levels register higher levels of adult functional literacy than their counterparts at these education levels in the United States. Insofar as family planning falls within the domain of adult functions, this may offer one explanation of why rates of unintended pregnancies and births are higher in the United States than in Western Europe despite the fact that formal levels of educational attainment in the United States are higher than in West European countries.

4.2.3 Income

At the level of the total population and that of individual racial/ethnic groups, it was the poorest women who had the highest fertility but beyond that there was hardly any association between income and fertility.

We are assuming that the high fertility of the poor contributed to the relatively high U.S. fertility because the relative proportions of poor in other developed countries were most likely smaller than in the United States.

Conceivably high fertility of the poor is directly linked to an association of poverty with low quality education, and with lower proportions of functionally literate young people.

There are numerous other factors which may contribute to relatively high unintended fertility of the poor.

- The poor are continuously preoccupied with numerous existential needs and contraception gets to be neglected. Many families earn too little to cover "basic necessities like food, housing, health and child care" (Bushey and Gundersen 2001; Ehrenreich 2001).
- Physical abuse of women (which is associated with unintended pregnancies) is probably more prevalent among the poor (see below).
- Fertility of young teenagers is high among the poor (Kirby et al. 2001).

The poor had the highest proportion of unintended pregnancies and a relatively small proportion of these were terminated by induced abortion.

As already mentioned above, the association between fertility and income is weak on an aggregate level. However, among better educated women a positive relationship between income and fertility can be detected, especially among those with graduate and professional degrees, but to a limited extent also among women with some college education and those with a bachelor's degree. This is probably a contributing factor to the relatively high fertility of White non-Hispanic women. A reasonably large proportion of White non-Hispanic families enjoyed a level of income which per-

[18] Respondents in the Netherlands also score higher than their United States counterparts, but the differences are not statistically significant.

mits the purchase of various household services as well as more and better quality child care. Regardless of whether these women are employed outside the household, a considerable part of the work needed for childrearing is done by others thus providing them with the option of having one or more children without a large demand on their time.

4.2.4 More comprehensive reflections on economic factors[19]

Economists tend to agree that the perceived costs of children exert some influence on fertility decisions. However, there is considerable disagreement over what these costs are, how they are distributed, and how they are perceived.

Microeconomic theory as developed by Becker (1981) and others emphasizes the time costs of children as well as direct expenditures, but most empirical estimates are based simply on cash expenditures. Haveman and Wolfe (1995) provide some basic estimates of time costs but these are based on estimates of opportunity costs of foregone earnings, with no consideration of foregone leisure.

A growing literature documents costs imposed on mothers that go beyond direct earnings losses, and are not completely attributed to reduced labor force experience (Joshi, 1990, 1991; Waldfogel, 1997; Budig and England, forthcoming). These are of particular relevance in a context in which mothers face a high risk of becoming primary income-providers (as well as care providers) for their children. Furthermore, these costs are significantly affected by a number of policy variables, including the relative cost of child care, the prevalence of part-time work and the social safety net.

Another vein of analysis emphasizes the ways in which individuals perceive costs. Following Richard Easterlin, some economists like Diane Macunovich (1999) emphasize the influence of relative incomes across generations, which may be affected by cohort size. That is, individuals hope to provide their children with a standard of living that is equivalent to or better than their own.

All three approaches to examining costs (direct expenditures, time costs, relative costs) seem relevant to a consideration of possible factors contributing to high fertility in the U.S.

It has often been suggested that women with low human capital and therefore relatively low wages are likely to have higher fertility than those for whom the "opportunity cost" is higher, and this seems to borne out in terms of relative fertility in the United States. Less consideration has been given to reasons why relatively highly educated women in the U.S. probably have higher fertility than their counterparts in Europe. One reason may be that the costs of purchasing child-care from relatively low wage women are lower in this country precisely because the distribution of income is more unequal. The paradigmatic case is one in which a professional/managerial woman is able to hire a full-time nanny who is a recent, frequently illegal, immigrant willing to work long hours for relatively low pay. Casual observation of the characteristics of women in charge of baby strollers in Central Park in New York City in the middle of the day supports this hypothesis. Data on income differentials in subsection 2.3.5 above support this hypothesis.

Relatively low minimum wages in the United States also lower the relative cost of childcare provided in family day care and child care centers, perhaps partly countervailing the relatively low level of public subsidy. Variation in income inequality and minimum wages across states offer some potential for empirical exploration of this issue.

Relatively low minimum wages apparently lower the relative costs of many other goods and services in the United States. Therefore, the extent to which the preparation of meals, laundry and cleaning of clothes, household cleaning and other services are purchased is probably significantly larger in the U.S. than in any other "western" country. This seems to be the case for all social classes and ethnic groups. It would appear that, for instance, the utilization of fast food services by lower and middle income strata is extensive, probably much more so than in other developed countries. The relatively substantial shift of work out of the household and its substitution by the purchase of services might be facilitating higher fertility.

Another factor relevant to the costs of substitutes for maternal care is the prevalence of shift work in the United States, where substantially less than half of all full-time workers work a regular 9–5 schedule Monday to Friday. Work and business hours are much less stringently regulated in the United States, relative to Europe. A surprisingly large percentage of so-called blue collar workers are able to arrange shifts in ways that allow both parents to schedule employment at different times of the day or week, making it possible to combine dual earner responsibilities with family care (at the expense of shared spousal time) (Presser, 1995). Shift work represents another U.S.-specific factor that may mitigate against relatively low levels of public subsidy for child care.

[19] This section benefitted immensely from discussions with Nancy Folbre of the University of Massachusetts

Public policies in the United States are also relevant to a consideration of costs of children. It is sometimes assumed that the lack of an explicit family allowance in this country implies less public support than is typical in European countries. However, many implicitly pronatalist policies are embodied in provisions of the tax code. The Earned Income Tax Credit, sometimes described as a wage subsidy, more closely resembles a means-tested family allowance. It was increased substantially in the 1980s. The maximum credit for families with two or more children, $3,656, is comparable to the size of some European family allowances. The value of other tax-related benefits, such as the exemption for dependents, increases with family income, as a result of the structure of the progressive income tax system.

The combined effect of the dependent tax exemption and the child tax credit for a family in the United States in the 31% tax bracket (for a married couple earning more than about $120,000) is $2,704 a year for a family with two children and $4,056 for a family with three children. The changes implemented by the Bush administration will phase in another $500 in tax credits per child for virtually all families except those in the bottom 20% of the income distribution, for whom refundability is limited (England and Folbre, 2001). The total tax benefits for a family with three children, in the 31% bracket could reach $5,556 per year.

Changes in the social safety net are also relevant. It has been argued that the public assistance program in effect before 1996 encouraged non-marital fertility. Whether this is the case or not, some researchers have reported small declines in response to strictly enforced work requirements, time limits, and so-called family caps. Furthermore, increasingly strict enforcement of paternal child support responsibilities, particularly among low-income fathers, has been linked to declines in fertility (Case, 1998). Here again, the variation in policies across states offers some potential for empirical hypothesis testing.

Finally, the potential importance of shifts in relative income suggests that patterns of economic growth deserve consideration. Following Macunovich's reasoning, the economic boom of the 1990s might be expected to have positive effects on U.S. fertility. Also relevant should be the growth of women's earnings relative to men during that period. While the price effects might be expected to be negative (due to opportunity costs) the income effects might be positive–that is, women might perceive greater ability to raise children on their own. This hypothesis could be examined through analysis of a longitudinal data set such as the Panel Study of Income Dynamics of National Longitudinal Survey of Youth.

4.2.5 Religion

There are probably relatively more population strata in the U.S. than in any European country in which religion and religiosity are a factor of high fertility, for instance, the Pennsylvania Dutch, Hutterite, and Mormons. Even so, the weight of these groups in the total U.S. population is small and thus their impact on overall U.S. fertility is likely to be negligible. Religiosity might also differ among ethnic/race groups and may underlie some of the differentials analyzed in previous sections.

4.2.6 Physical abuse of women

Physical abuse of women is widespread in the United States. According to the report of the surgeon general, an estimated 22 percent of women have been victims of rape (Satcher 2001). Goodwin et al. (2000) provide evidence that physical abuse of women leads to unintended pregnancies.

4.2.7 Lack of inter-generational communication

There is evidence that inter-generational communication can make a difference in the risk for sexual intercourse among young people (Blake et al., 2001). Is there any evidence that the lack of communication is worse in the U.S. than elsewhere?

4.3 Summary

Compared to most other developed countries, rates of unplanned pregnancies and births are high in the United States. In part various aspects of the relatively deficient family planning delivery system appear to be contributing factors, such as principal reliance on medical specialists, expensive services, inadequate effective choice of contraceptives, and limited sex education. Consequently also rates of induced abortions are high, however not sufficient to offset the high rates of unintended pregnancies. Furthermore, women in the U.S. have a longer exposure to conception in sexual unions in part due to earlier ages of entry as well as larger proportions of women entering into unions than in other Western countries.

A wide range of social and economic factors are contributing to high U.S. fertility. A relatively low level of functional literacy among the young and less educated population segments in the U.S. could be a part of the explanation for high rates of contraceptive failure and high rates of mistimed and unwanted pregnancies. Various manifestations of poverty, especially among the Blacks and Hispanics—such as lack and poor quality of education as well as grave existential concerns—facilitate high fertility. Childlessness also tends to be lower among groups with higher income and advanced education. These can afford to purchase services—household help, quality child care, food preparation by others, etc.—which provides an atmosphere where childbearing and childrearing are less stressful. In general, there are indications that the cost of child-care was relatively low not only for the higher income families, but often also for the middle class due to low minimum wages. The high prevalence of shift work provided parents with flexibility to share child-care at different times of the day and the week. In the United States the dependent tax exemptions together with the child tax credit provide child financial support comparable to family allowances common in European welfare states. Also, the economic boom of the 1990s might have had a positive effect on U.S. fertility. Finally, some racial/ethnic groups, such as the Hispanics, have a preference for higher fertility.

Turning to the relevance of this section for preparing fertility projections: Which of the factors pertaining to the proximate determinants of fertility or the vast area of indirect fertility determinants should be followed as indicators useful for such projections? 1. Contraceptive failure rates for various race/ethnic and socio-economic groups with the goal of assessing how the respective population is approaching a perfect contraceptive state; and 2. Trends of the prevalence of induced abortions because these provide information on the extent to which contraceptive failure rates are compensated by induced abortions.

In dealing with the multiple circumstances that either facilitate or curtail childbearing, fertility projections can benefit from focusing on selected factors or mechanisms, which have a direct bearing on particular proximate determinants of fertility levels and trends. Two such mechanisms are relevant in the United States: The health and family planning system from the point of view of delivering contraceptives; and the educational system from the point of view of raising levels of functional literacy because that provides part of the explanation for rates of contraceptive failure. Finally, we believe that theories dealing with the economic, social, cultural, and other mechanisms shaping motivations for childbearing (for instance, Van de Kaa, 1987; Kohler et al., 2001) can provide guidance in preparing fertility projections.

5. Brief conclusions

What does the exposition and analysis add up to with respect to future directions of U.S. fertility in the foreseeable future?[20]

Examination of the facts has demonstrated that U.S. fertility has been relatively stable and high during the past decade or so. In particular, the cohorts that were at the onset and in the middle of their reproductive careers have experienced fertility behavior quite distinct from other developed countries—higher and more stable. This is true for the White as well as for the other ethnic and racial sub-groups of the population. More accurately, fertility of the White community is relatively high and the fertility of all other race and ethnic groups is higher than that thus resulting in overall high fertility.

Nevertheless, there are powerful factors putting downward pressure on U.S. fertility just as in the other developed countries: The enhanced status of women, increasing years of formal education, establishing oneself in a career, reaching and maintaining economic security, etc. (Van de Kaa 1987, Lesthaeghe and Moors 1996). Such developments will persist and will continue to exert downward pressure on fertility.

In the United States, however, there are counter-pressures keeping fertility relatively high. Many among the predominantly White well educated and prosperous people can afford goods and services, which alleviate the costs and tensions of child-raising. To a lesser degree, this probably applies to a part of the middle class.

On the other end of the income distribution spectrum there are large groups of all races and ethnic origin, with an over-representation of the Black population, with low income who are overwhelmed by their daily struggle for survival. For them planning family size and contraception are suppressed by existential concerns. In addition, due to their limited education, often of low quality, they tend to have a lesser degree of functional literacy. In addition, the structure and various facets of health and family planning care for these underprivileged strata make access to modern contraceptives difficult. Consequently there is a high degree of contraceptive failure and high rates of unintended pregnancies and births.

Also, cultural factors and tradition can be of importance, as typified by the Hispanic population in the U.S., which has a high rate of intended pregnancies and births, particularly so among recent immigrants.

It is difficult to foresee which pressures will dominate in the near future. Over the years U.S. demographers have made considerable errors in forecasting fertility, notably in predicting the high fertility of the late 1940s and early 1950s. Keeping in mind the uncertainty about the future, the above analysis supports the assumption that an increase in fertility in the near

[20] We are not dealing with possible policy implications of the above analysis, because that is not within the mandate of the U.S. Bureau of the Census.

future is unlikely. Much of the analysis points in the direction of either stability, i.e. negligible change in fertility trends in the foreseeable future, or a mild and gradual fertility decline.

What supports the assumption about reasonable stability or a mild decline in U.S. fertility? In general terms, none of the important proximate or indirect factors modifying U.S. fertility trends are likely to change rapidly. More specifically:

- The educational system is unlikely to undergo rapid improvements which would bring about major changes in the composition of formal educational attainment of the various race/ethnic groups, and increase levels of functional literacy at a rapid pace;
- The health care system is not expected to change rapidly and improve effective access to contraception in the broad sense of the concept. However, some simple changes could lower the incidence of unintended pregnancies and births and a mild fertility decline would result;
- Any change of the income distribution is not a short-term process. Therefore, considerable proportions of all race/ethnic communities, especially of the Black and Hispanic ones, will remain poor and with unsatisfactory education;
- There is no reason to assume that the relatively strong desire for large families among some race/ethnic and religious groups will diminish rapidly.

The soft "scientific" evidence points in the direction of stability or a slight fertility decline, even though surprises cannot be ruled out.

References

Bachu, A. and M. O'Connell, 2000, *Fertility and American Women: June 1998*. Current Population Reports. P20-526. U.S. Census Bureau. Washington, D.C.

Becker, G., 1981, *A Treatise on the Family*. Enlarged Edition. Cambridge: Harvard University Press.

Blake, S.M., L. Simkin, R. Ledsky, C. Perkins, and J.M. Calabrese, 2001, "Effects of parent-child communications intervention on young adolescents' risk for early onset of sexual intercourse" *Family Planning Perspectives*, 33(2): 52-61.

Bongaarts, J., 1978, "A framework for analyzing the proximate determinants of fertility," *Population and Development Review,* 4(1): 105-132.

Bongaarts, J., 1982, "The fertility-inhibiting effects of the intermediate fertility variables," *Studies in Family Planning*, 13(6-7):179-89.

Budig, M. and P. England, forthcoming. "The Wage Penalty for Motherhood," *American Sociological Review*.

Bushey, H. and B. Gundersen, 2001, *When Work Just Isn't Enough: Measuring hardships faced by families after moving from welfare to work*, Washington, D.C.: Economic Policy Institute.

Case, A., 1998, "The Effects of Stronger Child Support Enforcement on Nonmarital Fertility," 191-215 in Irving Garfinkel, Sarah McLanahan, D.R. Meyer, and J.A. Seltzer, editors, *Fathers Under Fire: The Revolution in Child Support Enforcement*. New York: Russell Stage.

Darroch J.E., J.J. Frost, S. Singh and Study Team, 2001a, *Teenage Sexual and Reproductive Behavior in Developed Countries*, Occasional Report No. 3, New York: The Alan Gutmacher Institute

———, 2001b, "Differences in teenage pregnancy rates among five developed countries: The roles of sexual activity and contraceptive use," *Family Planning Perspectives*, 33(6): 244-250.

Davis, K. and J. Blake, 1956, "Social Structure and Fertility: An Analytic Framework," *Economic Development and Cultural Change,* 4(3): 211-235.

Dorbritz, J. and K. Schwarz, 1996, "Kinderlosigkeit in Deutschland - ein Massenphänomen? Analysen zu Erscheinungsformen und Ursachen," *Zeitschrift für Bevölkerungswissenschaft*, 21 (3): 231-261.

England, P. and N. Folbre, 2001, "Reforming the Social Family Contract: Public Support for Child Rearing in the U.S.", forthcoming in Greg Duncan and Lindsay Chase-Lansdale, *For Better or Worse: The Effects of Welfare Reform on Children in the United States*, New York: Russell Sage.

Ehrenreich, B., 2001, *Nickel and Dimed: On (not) Getting By in America*,

Forste R. and M. Tienda, 1996, "What's behind racial and ethnic fertility differentials?" *Population and Development Review Supplement*, 22: 109-133.

Foster, C., 2000, "The limits to low fertility: A biosocial approach," *Population and Development Review*, 26 (2): 209-234.

Frejka, T. and G. Calot, 2001, Cohort reproductive patterns in low-fertility countries, *Population and Development Review*, 27(1): 103-132.

Golini, A., 1998, "How low can fertility be? An empirical exploration," *Population and Development Review*, 24 (1): 59-73.

Goodwin M. et al., 2000, "Pregnancy intendedness and physical abuse around the time of pregnancy ...," *Maternal and Child Health Journal* 4 (2): 85-92.

Haveman, R. and B. Wolfe, 1995, "The Determinants of Children's Attainments: A Review of Methods and Findings." *Journal of Economic Literature XXXIII*: 1829-1878.

Henshaw, S.K., 1998, "Unintended pregnancy in the United States," *Family Planning Perspectives* 30 (1): 24-46.

———, 2001, Personal communications, 5 September and 29 November.

Henshaw, S.K. and D.J. Feivelson, 2000, "Teenage abortion and pregnancy statistics by state, 1996," *Family Planning Perspectives* 32 (6): 273-280.

Henshaw, S.K., S. Singh, and T. Haas, 1999a, "The incidence of abortion worldwide," *International Family Planning Perspectives*, 25, Supplement, S30-S38.

———, 1999b, "Recent trends in abortion rates worldwide," *International Family Planning Perspectives*, 25 (1):44-48.

Heuser, R.L., 1976, *Fertility Tables for U.S. Birth Cohorts by Color: United States, 1917-1973*, Publication Number (HRA) 76-1152, U.S. Department of Health, Education and Welfare, Public Health Service, Health Resources Administration, National Center for Health Statistics, Rockville, Maryland.

Jones, E.F., J.D. Forrest, S.K. Henshaw, J. Silverman and A. Torres, 1989, *Pregnancy, Contraception, and Family Planning Services in Industrialized Countries*, New Haven and London: Yale University Press.

Joshi, H., 1990, "The Cash Opportunity Cost of Childbearing: An Approach to Estimation Using British Evidence," *Population Studies* 44: 41-60.

———, 1991, "Sex and Motherhood as Sources of Women's Economic Disadvantage, in *Women's Issues in Social Policy*, ed. D. Groves and M. McClean. London: Routledge.

Kirby, D., K. Coyle and J.B. Gould, 2001, "Manifestations of poverty and birthrates among young teenagers in California zip code areas," *Family Planning Perspectives* 33 (2): 63-69.

Kohler, H.-P., F.C. Billari and J.A. Ortega, 2001, Towards a theory of lowest-low fertility, *Max Planck Institute for Demographic Research Working Paper*, WP 2001-032.

Lesthaeghe, R. and G. Moors, 1996, "Living arrangements, socio-economic position, and values among young adults: A pattern description for France, West Germany, Belgium and the Netherlands, 1990," in D. Coleman, editor, *Europe's Population in the 1990's*, Oxford University Press.

Macunovich, D., 1999, "The Fortunes of One's Birth: Relative Cohort Size and the Youth Labor Market in the U.S., *Journal of Population Economics*.

McDonald, P., 2001, "Theoretical paradigms to explain low fertility," paper presented at IU.S.SP workshop on *International Perspectives on Low Fertility: Trends, Theories and Policies*, Tokyo.

Organization for Economic Co-operation and Development, 2001, Education Policy Analysis 2001, OECD: Paris

Presser, H., 1995, "Job, Family, and Gender: Determinants of Nonstandard Work Schedules Among Employed Americans in 1991." *Demography*, Vol. 32(4) :577.

Ranjit, N., A. Bankole, J.E. Darroch and S. Singh, 2001, "Contraceptive failure in the first two years of use: Differences across socio-economic groups," *Family Planning Perspectives* 33 (1): 19-27.

Rowland, D.T., 1998, "Consequences of childlessness in later life," *Australasian Journal on Ageing* 17(1): 24-8.

Sardon, J.-P., 2000, "Évolutions récentes de la démographie des pays développés," *Population*, 55 (4-5), 729-764.

Satcher, D., 2001, *The Surgeon General's Call to Action to Promote Sexual Health and Responsible Sexual Behavior*, Washington, DC: Department of Health and Human Services

Singh, S., J.E. Darroch, J.J. Frost and Study Team, 2001, "Socioeconomic disadvantage and adolescent women's sexual and reproductive behavior: The case of five developed countries," *Family Planning Perspectives*, 33(6): 251-258.

United Nations Economic Commission for Europe (UNECE), 2000, *Economic Survey of Europe*, No. 2/3, Geneva: United Nations

U.S. Census Bureau, 2000, *Statistical Abstract of the United States*, Washington, D.C.

Van de Kaa, D., 1987, "Europe's Second Demographic Transition," Population Bulletin, 42:1, 1-57.

Waldfogel, J., 1997. "The Effect of Children on Women's Wages," *American Sociological Review* 62: 209-217.

Table 1. Total period fertility rates, United States and selected low-fertility countries, 1970-1999

Country/Region	Total period fertility rate			
	1970	1980	1990	1999
United States	**2.43**	**1.85**	**2.07**	**2.05**
Canada	2.36	1.73	1.71	1.60
Australia	2.86	1.89	1.90	1.77
New Zealand	3.13	2.01	2.14	1.86
Norway	2.50	1.72	1.93	1.84
England & Wales	2.40	1.88	1.84	1.73
West Germany	1.99	1.45	1.45	1.40
Italy	2.43	1.64	1.33	1.19
"Western" Europe - 18 country average	2.43	1.88	1.70	1.58
Overseas countries - 3 country average	2.78	1.88	1.91	1.74
U.S. compared to average "Western" Europe (in percent)	0	-2	22	29
U.S. compared to average overseas country (in percent)	-13	-2	8	17

Sources: Sardon 2000; U.S. Census Bureau 2000

Table 2. Fertility deficits and surpluses comparing birth cohorts 10 years apart, United States, cohorts 1930-31, 1940-41, 1950-51, and 1960-61

Fertility	Cohort 1930-31 and 1940-41		Cohort 1940-41 and 1950-51		Cohort 1950-51 and 1960-61	
	Age group	Number of children	Age group	Number of children	Age group	Number of children
Deficit	24 - 49	-0.677	14 - 31	-0.763	17 - 27	-0.147
Surplus	14 - 23	0.153	32 - 49	0.101	14 - 16	0.004
					28 - 49a	0.146
Total		**-0.524**		**-0.662**		**-0.003**

Note : a = Includes estimated data for ages 37-49 in the cohort 1960-61

Table 3. Cumulated cohort fertility rates (CCFRs) at specified ages and relative changes compared to birth cohorts 10 years older, United States, cohorts 1960-61, 1965-66, 1970-71 and 1975-76

Exact age	Cumulated fertility of birth cohort				Change of CCFR compared to cohort ten years older (in percent)			
	1960-61	1965-66	1970-71	1975-76	1960-61	1965-66	1970-71	1975-76
37	1.895	n.a.	n.a.	n.a.	-1	n.a.	n.a.	n.a.
32	1.584	1.592	n.a.	n.a.	-5	0	n.a.	n.a.
27	1.057	1.062	1.072	n.a.	-12	-4	1	n.a.
22	0.491	0.472	0.513	0.528	-18	-7	4	12

Table 4. Cumulated cohort fertility rates (CCFRs), by completed age 27 and after age 27, selected countries, birth cohorts 1950-51, 1960-61, 1970-71 and 1975-76

Country	CCFR of birth cohort at completed age 27			CCFR of birth cohort after completed age 27			Measure of compensation of fertility after age 27
	1950-51	1960-61	1960-61 minus 1950-51	1950-51	1960-61	1960-61 minus 1950-51	
United States	**1.201**	**1.057**	**-0.144**	**0.814**	**0.960**	**0.146**	**101**
Canada	0.924	0.697	-0.227	0.982	1.114	0.132	58
Australia	1.349	0.918	-0.431	0.982	1.198	0.216	50
New Zealand	1.635	1.091	-0.544	0.911	1.246	0.335	62
Norway	1.258	0.908	-0.350	0.828	1.177	0.349	100
England & Wales	1.151	0.911	-0.240	0.901	1.034	0.133	55
Western Germany	0.940	0.662	-0.278	0.753	0.932	0.179	64
Italy	1.037	0.744	-0.293	0.819	0.868	0.049	17

Table 5. Cumulated cohort fertility rates (CCFRs) by exact age 27, United States and selected countries, birth cohorts 1950-51, 1960-61 and 1970-71

Country	CCFR of birth cohort by exact age 27			Difference of CCFRs between birth cohorts (in percent)	
	1950-51	1960-61	1970-71	1960-61/ 1950-51	1970-71/ 1960-61
United States	**1.201**	**1.057**	**1.072**	**-12**	**1**
Canada	0.924	0.697	0.581	-24	-17
Australia	1.349	0.918	0.646	-32	-30
New Zealand	1.635	0.794	0.639	-51	-19
Norway	1.258	0.908	0.743	-28	-18
England & Wales	1.151	0.911	0.755	-21	-17
West Germany	0.940	0.662	0.478	-30	-28
Italy	1.037	0.744	0.445	-28	-40

Table 6. Cumulated cohort fertility rates (CCFRs) by exact age 22, United States and selected countries, birth cohorts 1950-51, 1960-61, 1970-71 and 1975-76

Country	CCFR by exact age 22 in cohort				Difference of CCFRs between birth cohorts (in percent)				
	1950-1951	1960-1961	1970-1971	1975-1976	1960-61/ 1950-51	1970-71/ 1960-61	1975-76/ 1970-71	1975-76/ 1960-61	1975-76/ 1950-51
United States	**0.600**	**0.491**	**0.513**	**0.528**	**-18**	**4**	**3**	**7**	**-12**
Canada	0.307	0.201	0.176	0.177	-34	-13	1	-12	-43
Australia	0.385	0.212	0.152	0.148	-45	-28	-3	-30	-62
New Zealand	0.486	0.317	0.240	0.223	-35	-24	-7	-30	-54
Norway	0.362	0.216	0.146	0.116	-40	-32	-20	-46	-68
England & Wales	0.365	0.235	0.224	0.214	-36	-5	-5	-9	-41
West Germany	0.396	0.206	0.133	0.140	-48	-36	5	-32	-65
Italy	0.210	0.182	0.073	0.054	-13	-60	-25	-70	-74

Table 7. Parity distribution at age 40, United States and selected "Western" countries, birth cohort 1960

Country	Parity (in percent)				
	0	1	2	3	4+
United States	**17**	**18**	**34**	**20**	**11**
Norway[a]	12	14	40	25	8
West Germany	23	22	37	18	
Italy[b]	17	26	42	15	5
Netherlands	19	16	41	18	6
Finland	19	16	36	20	9

Notes: [a]1958; [b]1955

Table 8. Proportion of women childless, United States and selected “Western” countries at age 40, birth cohorts 1930 to 1970

Birth cohort	Country							
	United States	Norway	West Germany[a]	Italy	Ireland	Netherlands	Finland	Sweden
1930	**10.2**			20.1		14.9	11.3	
1935	**9.2**	9.8		17.3		12.4	9.1	
1940	**11.0**	9.7	10.1	22.5		11.4	9.0	
1945	**13.7**	9.2	13.3	17.2		11.6	9.2	
1950	**16.8**	9.7	14.9	17.2		15.2	11.8	
1955	**17.6**	11.6	19.4	13.8	14.1	17.7	16.5	13.7
1960	**16.9**	12.3[b]	23.2	17.1	16.6	18.7	18.8	14.3
1965	**16.3**				22.1	20.3	21.7	15.2
1967							22.8	16.9

Sources: Observatoire Demographique Europeen, for Norway Lappegard (2000); Dorbritz, Schwarz 1996 for West Germany

Notes: [a]Age 49; [b]1958

Table 9. Parity progression ratios, United States and selected “Western” countries, birth cohorts of the late 1950s and early 1960s

	Parity progression ratio (in percent)							
	United States							
	1950-51	1960-61	1961-62	1962-63	1963-64	1964-65	1965-66	1966-67
0 : 1	84	85	85	85	85	85	85	n.a.
1 : 2	79	78	78	78	78			
2 : 3	47	48	48					
3 : 4	38	37						
	Finland							
	1950-51	1958-1959	1959-1960	1960-1961	1961-1962	1962-1963	1963-1964	1964-1965
0 : 1	87	83	82	82	81	81	80	79
1 : 2	72	80	80	81	81	81		
2 : 3	39	46	46	46				
3 : 4	28	31						
	Netherlands							
	1950	1961	1962	1963	1964	1965	1966	1967
0 : 1	84	82	82	82	82	81	81	n.a.
1 : 2	79	81	80	80	79			
2 : 3	47	37	36	35				
3 : 4	38	26						

Table 10. Cumulative cohort fertility rates (CCFRs) by exact age 27 and after age 27, U.S. racial groups, birth cohorts 1950 and 1960

Racial group	CCFR of birth cohort at exact age 27			CCFR of birth cohort after exact age 27			Measure of compensation of fertility after age 27 (in%)
	1950	1960	1950-1960	1950	1960	1950-1960	
US All Races	1.247	1.072	-0.175	0.789	0.949	0.160	91
US White	1.168	1.003	-0.165	0.784	0.959	0.175	106
US nonwhite	1.728	1.432	-0.296	0.819	0.906	0.086	29

Table 11. Cumulative cohort fertility rates by exact ages 22 and 27, United States racial groups, birth cohorts 1950, 1960, 1970, and 1975.

Racial group	Cumulative cohort fertility rate				Differences in CCFRs between birth cohorts (in percent)		
	1950	1960	1970	1975	1960/ 1950	1970/ 1960	1975/ 1970
By age 22							
US Total	0.632	0.493	0.479	0.511	-22	-3	7
US White	0.561	0.431	0.414	0.453	-23	-4	9
US Nonwhite	1.065	0.818	0.765	0.738	-23	-7	-3
By age 27							
US Total	1.247	1.072	1.040	n.a.	-14	-3	n.a.
US White	1.168	1.003	0.966	n.a.	-14	-4	n.a.
US Nonwhite	1.728	1.432	1.370	n.a.	-17	-4	n.a.

Table 12. Children ever born, percent childless and numbers of women 40-44 years old of certain race, ethnicity, nativity, United States, Current Population Survey 1998.

Characteristic	Total	White	White non-Hispanic	Hispanic	Black	Other	Asian & Pacific Islander	Native	Foreign Born
Absolute measures									
Number of women (in thousands)	9,842	8,046	7,113	991	1,267	84	446	8,481	1,362
Children ever born (per woman)	1.88	1.84	1.77	2.39	2.04	2.12	2.05	1.82	2.21
Percent childless	19.1	19.4	20.1	14.5	17.3	19.3	18.1	19.7	14.9
Relative measures									
Proportion of category (in percent of total)	100.0	81.7	72.3	10.1	12.9	0.9	4.5	86.2	13.8
CEB (White non-Hispanic = 100)	106	104	100	135	115	120	116	103	125
Percent childless (White non-Hispanic =	95	97	100	72	86	96	90	98	74

Table 13. Distribution of women 40–44 years old of certain race, ethnicity, nativity by income category, United States, 1998

Income category	White non-Hispanic	Hispanic	Black	Asian & Pacific Islanders	Native	Foreign Born
<$10000	5.0	14.1	18.1	6.8	6.8	11.7
$10000-$19999	6.3	19.9	16.3	10.4	8.0	15.2
$20000-$29999	11.0	16.1	14.3	9.9	11.5	13.3
$30000-$49999	25.2	27.0	25.0	20.8	25.0	25.7
$50000-$74999	25.6	13.1	17.5	22.4	24.6	15.5
$75000+	26.8	9.8	8.7	29.7	24.2	18.5
Total	100.0	100.0	100.0	100.0	100.0	100.0

Table 14. Distribution of women 40–44 years old of certain race, ethnicity, nativity by educational attainment, United States, 1998

Educational attainment	White non-Hispanic	Hispanic	Black	Asian & Pacific Islanders	Native	Foreign Born
No high school	7.1	39.1	14.2	13.0	8.5	31.2
Complete high school	34.5	26.8	35.0	26.7	34.7	25.5
Some college, including Associate degree	29.0	21.7	33.5	19.7	30.0	18.1
Bachelors degree	20.4	9.5	13.3	27.8	18.8	17.6
Grad/Prof degree	9.2	2.9	4.1	12.8	8.1	7.6
Total	100.0	100.0	100.0	100.0	100.0	100.0

Table 15. Children ever born, women 40–44 years old of certain race, ethnicity, nativity by income category, United States, CPS 1998

Income category	Total	White	White non-Hispanic	Hispanic	Black	Other	Asian & Pacific Islander	Native	Foreign Born
Average number of children ever born									
Total	1.88	1.84	1.77	2.37	2.09	2.16	2.02	1.83	2.20
<$10000	2.33	2.18	1.97	2.71	2.62	3.03	2.00	2.27	2.51
$10000-$19000	2.10	2.09	1.73	2.96	2.00	2.35	2.59	1.91	2.71
$20000-$29999	1.85	1.76	1.66	2.20	2.03	3.88	2.77	1.72	2.48
$30000-49999	1.87	1.84	1.80	2.14	1.94	2.17	2.15	1.84	2.04
$50000-74999	1.82	1.81	1.77	2.44	1.99	1.47	1.71	1.79	2.05
$75000+	1.76	1.76	1.77	1.68	1.93	1.45	1.70	1.77	1.71
Children ever born (<$10,000 = 100)									
<$10000	100	100	100	100	100	100	100	100	100
$10000-$19000	90	96	88	109	76	78	130	84	108
$20000-$29999	79	81	84	81	77	128	139	76	99
$30000-$49999	80	84	91	79	74	72	108	81	81
$50000-$74999	78	83	90	90	76	49	86	79	82
$75000+	76	81	90	62	74	48	85	78	68
Children ever born ($75,000+ = 100)									
<$10000	132	124	111	161	136	209	118	128	147
$10000-$19000	119	119	98	176	104	162	152	108	158
$20000-$29999	105	100	94	131	105	268	163	97	145
$30000-$49999	106	105	102	127	101	150	126	104	119
$50000-$74999	103	103	100	145	103	101	101	101	120
$75000+	100	100	100	100	100	100	100	100	100

Table 16. Children ever born, women 40–44 years old of certain race, ethnicity, nativity by educational attainment, United States, CPS 1998

Educational attainment	Total	White	White non-Hispanic	Hispanic	Black	Other	Asian & Pacific Islander	Native	Foreign Born
Average number of children ever born									
Total	1.88	1.84	1.77	2.39	2.04	2.12	2.05	1.82	2.21
No high school	2.57	2.52	2.12	3.06	2.82	2.54	2.60	2.40	2.84
High school	1.97	1.92	1.90	2.21	2.13	2.31	2.31	1.94	2.20
Some college	1.81	1.76	1.74	1.98	1.90	2.34	2.37	1.79	1.96
Bachelors degree	1.67	1.67	1.68	1.43	1.69	1.06	1.67	1.64	1.84
Grad/Prof degree	1.32	1.33	1.33	1.32	1.12	0.96	1.44	1.31	1.36
Children ever born (No high school = 100)									
No high school	100	100	100	100	100	100	100	100	100
High school	77	76	90	72	76	91	89	81	77
Some college	70	70	82	65	67	92	91	75	69
Bachelors degree	65	66	79	47	60	42	64	68	65
Grad/Prof degree	51	53	63	43	40	38	55	55	48
Children ever born (Grad/Prof degree = 100)									
No high school	195	189	159	232	252	265	181	183	209
High school	149	144	143	167	190	241	160	148	162
Some college	137	132	131	150	170	244	165	137	144
Bachelors degree	127	126	126	108	151	110	116	125	135
Grad/Prof degree	100	100	100	100	100	100	100	100	100

Table 17. Children ever born, White non-Hispanic women 40–44 years old, by educational attainment and income, United States, CPS 1998

	Total	<$10000	$10000-$19999	$20000-$29999	$30000-$49999	$50000-$74999	$75000+
Total	1.77	1.97	1.73	1.66	1.80	1.77	1.77
No high school	2.19	2.25	2.10	1.98	2.32	1.96	3.11
High school	1.90	2.08	1.87	1.80	1.94	1.95	1.78
Some college	1.74	1.51	1.33	1.61	1.80	1.74	1.86
Bachelors degree	1.69	3.36	1.53	1.15	1.49	1.73	1.79
Grad/Prof degree	1.33	0.37	0.28	0.87	0.94	1.17	1.56

Table 18. Children ever born, White non-Hispanic women 40–44 years old, native and foreign-born, by educational attainment and income, United States, CPS 1998

	Native	Foreign-born	Foreign/Native		Native	Foreign-born	Foreign/Native
Total	2.08	2.57	124	Total	2.08	2.57	124
				<$10000	2.55	2.76	108
No high school	2.89	3.00	104	$10000-$19999	2.68	2.88	107
High school	2.07	2.43	117	$20000-$29999	1.81	2.49	138
Associate degree	2.06	2.04	99	$30000-$49999	1.99	2.33	117
Bachelors degree	1.19	1.69	142	$50000-$74999	2.24	2.58	115
Grad/Prof degree	1.50	1.40	93	$75000+	1.59	1.88	118
No high school	193	214		<$10000	160	147	
High school	138	174		$10000-$19999	169	153	
Associate degree	137	146		$20000-$29999	114	132	
Bachelors degree	79	121		$30000-$49999	125	124	
Grad/Prof degree	100	100		$50000-$74999	141	137	
				$75000+	100	100	

Table 19. Children ever born, White non-Hispanic women 40–44 years old, native and foreign-born, by educational attainment and income, United States, CPS 1998

	Native	Foreign-born	Foreign/ Native		Native	Foreign-born	Foreign/ Native
Total	1.78	1.65	93	Total	1.78	1.65	93
				<$10000	2.00	1.68	84
No high school	2.21	2.01	91	$10000-$19999	1.75	1.26	72
High school	1.90	1.81	95	$20000-$29999	1.65	1.80	109
Associate degree	1.76	1.52	86	$30000-$49999	1.81	1.63	90
Bachelors degree	1.69	1.73	102	$50000-$74999	1.76	2.05	116
Grad/Prof degree	1.34	1.19	89	$75000+	1.78	1.42	80
No high school	165	169		<$10000	112	118	
High school	142	152		$10000-$19999	98	89	
Associate degree	131	128		$20000-$29999	93	127	
Bachelors degree	126	145		$30000-$49999	102	115	
Grad/Prof degree	100	100		$50000-$74999	99	144	
				$75000+	100	100	

Table 20. Proportions childless, women 40–44 years old of certain race, ethnicity, nativity by income category, United States, 1998

Income Category	White non-Hispanic	Hispanic	Black	Asian & Pacific Islanders	Native	Foreign Born
Total	20.1	14.5	17.3	18.1	19.7	14.9
<$10000	22.7	11.9	9.9	22.9	17.4	13.3
$10000-$19999	19.3	6.4	11.8	18.6	15.0	10.8
$20000-$29999	25.4	11.1	18.8	9.6	23.7	10.9
$30000-$49999	20.7	23.8	21.6	20.9	21.0	21.4
$50000-$74999	18.5	9.0	20.5	18.1	18.9	12.7
$75000+	18.9	21.1	23.3	18.3	19.8	15.3

Table 21. Proportions childless, women 40–44 years old of certain race, ethnicity, nativity by educational attainment, United States, 1998

Educational attainment	White non-Hispanic	Hispanic	Black	Asian & Pacific Islanders	Native	Foreign Born
Total	20.1	14.5	17.3	18.1	19.7	14.9
No high school	16.1	11.2	11.1	22.4	12.5	14.4
High school	14.0	14.3	12.8	13.4	13.9	13.3
Some college	18.9	8.2	16.3	8.6	17.9	13.1
Bachelors degree	26.3	32.6	27.6	22.5	28.6	13.9
Grad/Prof degree	32.9	34.4	46.0	24.4	34.7	26.0

Table 22. Proportion childless, women 40–44 years old of certain race, ethnicity, nativity by educational attainment and income, United States, 1998

	Total	<$10000	$10000-$19999	$20000-$29999	$30000-$49999	$50000-$74999	$75000+
Total	20.1	22.7	19.3	25.4	20.7	18.5	18.9
No high school	16.1	18.0	14.3	29.3	13.2	7.7	-
High school	14.0	17.6	16.9	18.2	13.5	11.4	13.0
Some college	18.9	28.0	22.6	21.6	19.1	17.9	16.7
Bachelors degree	26.3	14.3	23.7	50.9	34.2	23.2	22.1
Grad/Prof degree	32.9	81.7	83.8	51.9	52.1	39.3	22.7

Table 23. Estimated number of pregnancies (excluding miscarriages), percentage distribution of pregnancies, by outcome and intention, and selected measures on unintended pregnancy, all by characteristic, 1994

Characteristic	Number of pregnancies	Percent distribution of pregnancies				Percent of births that were unintended	Percent of pregnancies that were unintended	Percent of unintended pregnancies that ended in abortion	General pregnancy rate*		
		Intended births	Unintended births	Abortions	Total				Total	Intended	Unintended
Total	5,383,800	50.8	23.0	26.6	100.0	30.8	49.2	54.0	90.8	46.1	44.7
Age at outcome											
<15 a)	25,100	18.3	33.2	48.5	100.0	64.5	81.7	59.4	13.7	2.5	11.2
15-19	781,900	22.0	42.7	35.3	100.0	66.0	78.0	45.3	91.1	20.0	71.1
15-17	306,100	17.3	46.5	36.2	100.0	72.9	82.7	43.8	59.0	10.2	48.8
18-19	475,800	25.0	40.2	34.8	100.0	61.7	75.0	46.4	140.3	35.1	105.2
20-24	1,479,500	41.5	26.2	32.3	100.0	38.7	58.5	55.2	164.1	68.1	96.0
25-29	1,405,200	60.3	17.2	22.5	100.0	22.2	39.7	56.7	147.0	88.7	58.4
30-34	1,111,400	66.9	14.6	18.4	100.0	18.0	33.1	55.7	100.0	66.9	33.1
35-39	482,400	59.2	17.9	23.0	100.0	23.2	40.8	56.3	43.7	25.9	17.8
40+ b)	98,300	49.3	17.9	32.8	100.0	26.7	50.7	64.7	9.9	4.9	5.0
Marital status at outcome											
Currently married c)	3,003,900	69.3	19.3	11.3	100.0	21.8	30.7	37.0	95.2	66.0	29.2
Formerly married	356,700	37.5	21.8	40.7	100.0	36.8	62.5	65.1	64.7	24.3	40.4
Never married	2,023,100	22.3	31.0	46.7	100.0	58.2	77.7	60.1	91.0	20.3	70.8
Poverty status**											
<100%	1,358,000	38.6	31.3	30.1	100.0	44.8	61.4	49.0	143.7	55.4	88.3
100-199%	1,292,500	46.8	27.7	25.4	100.0	37.2	53.2	47.9	115.2	53.9	61.2
200%+	2,733,200	58.8	15.9	25.4	100.0	21.3	41.2	61.5	70.8	41.6	29.2
Race											
White	3,981,700	57.1	21.2	21.6	100.0	27.1	42.9	50.4	82.7	47.3	35.5
Black	1,130,700	27.7	28.6	43.7	100.0	50.8	72.3	60.4	136.7	37.8	98.9
Other	271,400	50.0	22.0	28.0	100.0	30.5	50.0	56.0	93.9	46.9	46.9
Ethnicity											
Hispanic	900,200	51.4	22.4	26.1	100.0	30.4	48.6	53.8	143.0	73.5	69.4
Non-Hispanic	4,483,600	50.7	22.6	26.7	100.0	30.9	49.3	54.1	84.6	42.9	41.7

*Pregnancy rates for this category are expressed as per 1,000 women aged 15-44, except for rates for age groups. a=Denominator for rates is women aged 14. b=Numerator for rates is women aged 40 and older; denominator is women aged 40-44. c=Includes separated women. **Percentage of federal poverty level at time of interview. In 1994, the poverty level was $17,020 for a family of four. *Note :* Intention status of births in the five years before the 1995 interview.

Source: Henshaw, S.K., 1998, "Unintended pregnancy in the United States," Family Planning Perspectives 30 (1): 24-46.

Table 24. Estimated total pregnancy rates by poverty status and race/ethnic group, United States, 1994

Race/ethnic group	Poverty status			
	<100%	100-199%	>200%	Total
Total unintended pregnancy rate				
White non-Hispanic	1.56	1.46	0.74	0.97
Black non-Hispanic	3.23	3.00	2.88	3.05
Hispanic	2.82	1.81	1.51	2.02
Total intended pregnancy rate				
White non-Hispanic	1.09	1.47	1.24	1.28
Black non-Hispanic	1.17	1.24	1.09	1.16
Hispanic	2.62	2.03	1.57	2.08
Total pregnancy rate				
White non-Hispanic	2.64	2.93	1.97	2.26
Black non-Hispanic	4.40	4.24	3.96	4.20
Hispanic	5.44	3.84	3.08	4.10

Source: S.K. Henshaw, 2001.

Table 25. Estimated unintended and intended pregnancy rates by age, race/ethnicity and poverty status, 1994

Age	White non-Hispanic				Black non-Hispanic				Hispanic			
	<100%	100-199%	>200%	Total	<100%	100-199%	>200%	Total	<100%	100-199%	>200%	Total
Unintended pregnancy rates												
15-19	84.0	85.5	32.2	48.6	195.2	131.3	128.5	155.9	105.4	77.7	86.2	91.2
20-24	90.4	84.5	52.4	65.5	229.2	217.5	202.5	217.0	168.2	117.3	115.2	135.0
25-29	70.7	59.5	29.9	40.3	114.0	139.1	139.4	129.6	145.1	73.4	37.8	83.2
30-34	44.5	31.9	19.6	24.2	55.5	72.0	67.0	63.9	75.1	48.8	36.3	51.8
35-39	18.2	23.9	9.9	12.6	41.6	32.2	28.2	33.2	52.0	25.2	20.8	31.2
40+	3.3	6.7	3.3	3.7	9.6	7.8	10.0	9.5	17.5	19.4	6.6	12.5
15-44	56.6	50.2	21.7	30.3	119.1	100.5	86.3	101.2	103.2	63.9	47.4	70.3
Intended pregnancy rates												
15-19	28.3	28.6	6.6	13.4	35.4	26.0	13.3	25.3	76.7	71.9	23.7	59.4
20-24	76.3	80.1	45.4	58.0	71.8	78.1	54.2	67.0	166.9	135.9	80.8	128.9
25-29	62.4	108.0	85.5	86.7	60.5	73.4	68.0	66.4	141.5	96.8	119.3	119.6
30-34	33.6	48.5	76.9	67.6	40.1	54.4	57.9	50.4	85.6	69.8	65.6	72.9
35-39	12.4	21.2	28.7	26.1	15.8	16.6	22.9	19.1	53.9	12.6	24.6	29.6
40+	4.7	7.5	4.0	4.4	11.3	0.0	0.9	3.0	0.0	18.5	0.0	4.6
15-44	40.0	52.4	40.7	42.5	42.6	42.3	34.1	39.0	96.7	72.5	53.6	73.1

Source : S.K. Henshaw, 2001.

Table 26. Estimated total unintended, mistimed and unwanted pregnancy rates by poverty status and race/ethnic group, United States, 1994

Race/ethnic group	Poverty status			
	<100%	100-199%	>200%	Total
Total unintended pregnancy rate				
White non-Hispanic	1.56	1.46	0.74	0.97
Black non-Hispanic	3.23	3.00	2.88	3.05
Hispanic	2.82	1.81	1.51	2.02
Total mistimed pregnancy rate				
White non-Hispanic	1.03	0.99	0.56	0.71
Black non-Hispanic	1.72	1.91	1.87	1.82
Hispanic	1.59	1.13	1.10	1.27
Total unwanted pregnancy rate				
White non-Hispanic	0.53	0.47	0.17	0.27
Black non-Hispanic	1.51	1.09	1.01	1.22
Hispanic	1.23	0.68	0.41	0.76
Total mistimed pregnancy rate as % of unintended rate				
White non-Hispanic	66	68	77	72
Black non-Hispanic	53	64	65	60
Hispanic	56	62	73	63
Total unwanted pregnancy rate as % of unintended rate				
White non-Hispanic	34	32	23	28
Black non-Hispanic	47	36	35	40
Hispanic	44	38	27	37

Source: Henshaw, 2001.

Table 27. Percentages of mistimed and unwanted pregnancies ending in abortions, by poverty status and race/ethnic group, United States, 1994

Race/ethnic group	Poverty status			
	<100%	100-199%	>200%	Total
Percent of mistimed pregnancies ending in abortions				
White non-Hispanic	40	38	53	46
Black non-Hispanic	49	60	78	62
Hispanic	57	48	61	56
Percent of unwanted pregnancies ending in abortions				
White non-Hispanic	50	43	62	54
Black non-Hispanic	46	64	75	59
Hispanic	44	44	62	48

Source: Henshaw, 2001.

Table 28. Estimated total unintended, mistimed and unwanted fertility rates by poverty status and race/ethnic group, United States, 1994

Race/ethnic group	Poverty status			
	<100%	100-199%	>200%	Total
Total unintended fertility rate				
White non-Hispanic	0.88	0.89	0.33	0.50
Black non-Hispanic	1.70	1.16	0.66	1.19
Hispanic	1.35	0.97	0.59	0.96
Total mistimed fertility rate				
White non-Hispanic	0.62	0.62	0.26	0.38
Black non-Hispanic	0.88	0.77	0.41	0.69
Hispanic	0.69	0.59	0.43	0.56
Total unwanted fertility rate				
White non-Hispanic	0.26	0.27	0.06	0.13
Black non-Hispanic	0.81	0.39	0.25	0.51
Hispanic	0.66	0.38	0.16	0.40
Total mistimed fertility rate as % of unintended rate				
White non-Hispanic	70	69	80	75
Black non-Hispanic	52	66	62	58
Hispanic	51	61	73	59
Total unwanted fertility rate as % of unintended rate				
White non-Hispanic	30	31	20	25
Black non-Hispanic	48	34	38	42
Hispanic	49	39	27	41

Sources: Henshaw, 2001.

Table 29. Estimated rates of unintended pregnancies, unintended births and abortions per 1,000 women, and percentage of unintended pregnancies ended in abortion, by age and marital status, 1981, 1987, and 1994

Characteristic	Unintended pregnancy			Unintended birth			Abortion			Percent ended by abortion		
	1981	1987	1994	1981	1987	1994	1981	1987	1994	1981	1987	1994
Total	54.2	53.5	44.7	25.0	26.6	20.9	29.2	26.9	24.1	53.9	50.3	54.0
Age at outcome												
15-19	78.1	79.3	71.1	35.2	37.1	38.9	42.9	42.2	32.2	54.9	53.2	45.3
20-24	93.6	102.7	96.0	42.3	50.2	43.0	51.4	52.5	53.0	54.8	51.1	55.2
25-29	60.6	66.1	58.4	29.3	35.4	25.3	31.3	30.8	33.1	51.6	46.5	56.7
30-34	37.0	37.3	33.1	19.3	19.3	14.6	17.7	17.9	18.4	47.8	48.2	55.7
35-39	15.0	18.8	17.8	5.5	9.0	7.8	9.5	9.8	10.0	63.5	52.2	56.3
40+*	4.3	5.3	5.0	0.9	2.4	1.8	3.4	2.9	3.2	78.2	54.3	64.7
Marital status at outcome												
Currently married	u	41.5	29.2	u	29.8	18.4	u	11.7	10.8	u	28.2	37.0
Formerly married	u	54.6	40.4	u	19.0	14.1	u	35.7	26.3	u	65.3	65.1
Never married	u	71.5	70.8	u	23.2	28.2	u	48.2	42.5	u	67.5	60.1

*Numerator for rates is women aged 40 and older; denominator is women aged 40-44. a=Includes separated women. Notes: All measures exclude miscarriages. The intention status of births is based on births in five year age groups in the five years before the interviews in 1988 and 1995 and in the four years before the 1982 interview. u=unavailable.

Source: Henshaw, S.K., 1998, "Unintended pregnancy in the United States," *Family Planning Perspectives* 30 (1): 24-46.

Table 30. Total planned and unplanned pregnancy rates, total pregnancy rate, and ratio of total unplanned to total pregnancy rates, 1976-1983

Country	Year following fertility survey	Total planned pregnancy rate	Total unplanned pregnancy rate	Total pregnancy rate	Ratio of total unplanned pregnancy rate to total pregnancy rate
Belgium *	1977	1.17	0.83	2.00	41.5
Canada**	1985	1.24	0.79	2.03	38.9
Denmark	1976	1.32	1.18	2.50	47.2
Finland	1978	1.03	1.06	2.09	50.7
France	1979	1.18	1.35	2.53	53.4
Netherlands	1983	1.37	0.28	1.65	17.0
Sweden	1982	1.39	0.80	2.19	36.5
Great Britain	1977	1.35	0.63	1.98	31.8
United States	**1983**	**1.25**	**1.31**	**2.56**	**51.2**

* Abortion estimate is for 1979; ** Abortion estimate is for 1982
Source : Jones, E.F. et al., 1989, *Pregnancy, Contraception, and Family Planning Services in Industrialized Countries,* Yale University Press, New Haven and London, p. 12.

Table 31. Total fertility and pregnancy rates by planning status: Ontario, Quebec, the Netherlands, Great Britain and United States, 1976 and 1982

Rate	Ontario* 1984	Quebec** 1984	Netherlands*** 1982	Great Britain 1976	**United States 1982**
Total fertility rate	1.7	1.5	1.5	1.7	**1.8**
Total planned pregnancy rate	1.2	1.1	1.4	1.3	**1.1**
Total unplanned fertility rate	0.5	0.3	0.1	0.4	**0.7**
Total pregnancy rate	2.1	1.7	1.7	2.0	**2.6**
Total planned pregnancy rate	1.2	1.1	1.4	1.3	**1.1**
Total unplanned pregnancy rate	0.9	0.6	0.3	0.7	**1.4**
Total abortion rate	0.4	0.3	0.2	0.3	**0.8**

*Fertility and abortion data for 1982 rather than 1983; **Abortion data for 1982 rather than 1985;
***Abortion data for 1982 rather than 1983

Source : Jones, E.F. et al., 1989, *Pregnancy, Contraception, and Family Planning Services in Industrialized Countries,* Yale University Press, New Haven and London, p. 65.

Figure 1. Total period fertility rates, United States and averages for selected low-fertility countries, 1970-1999

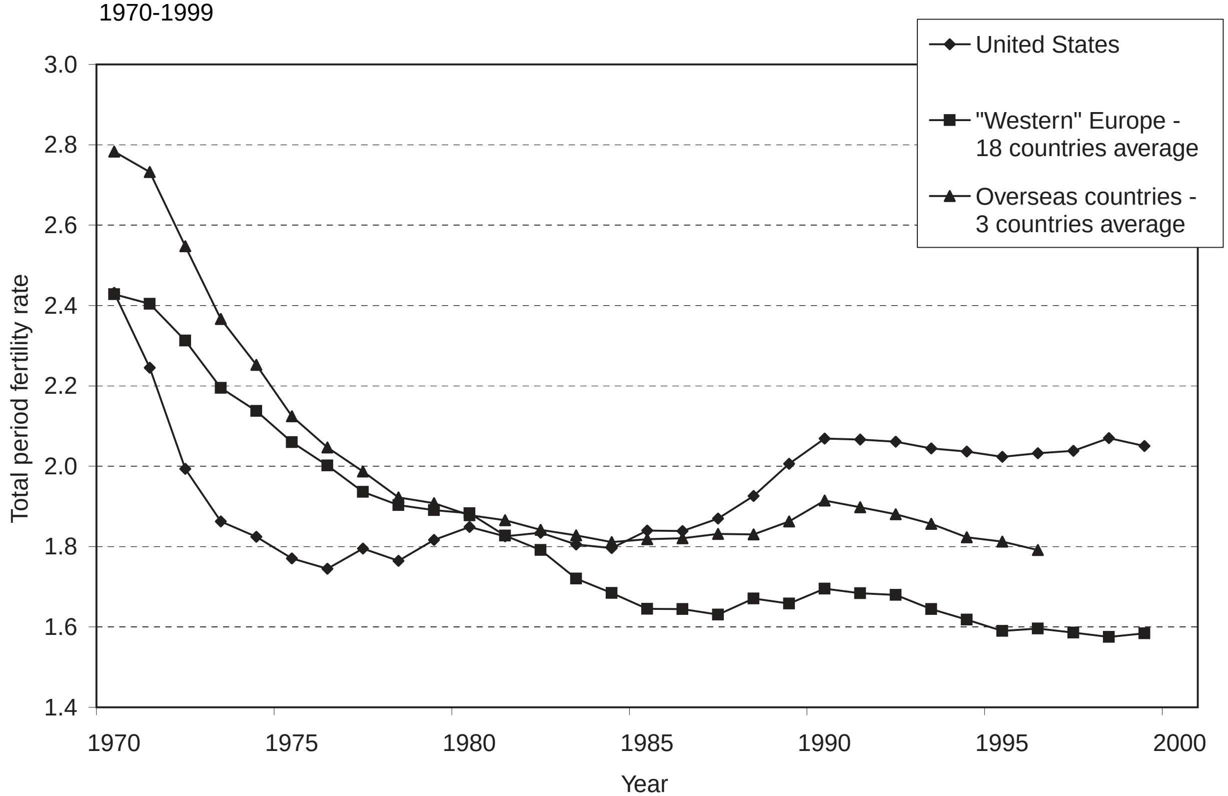
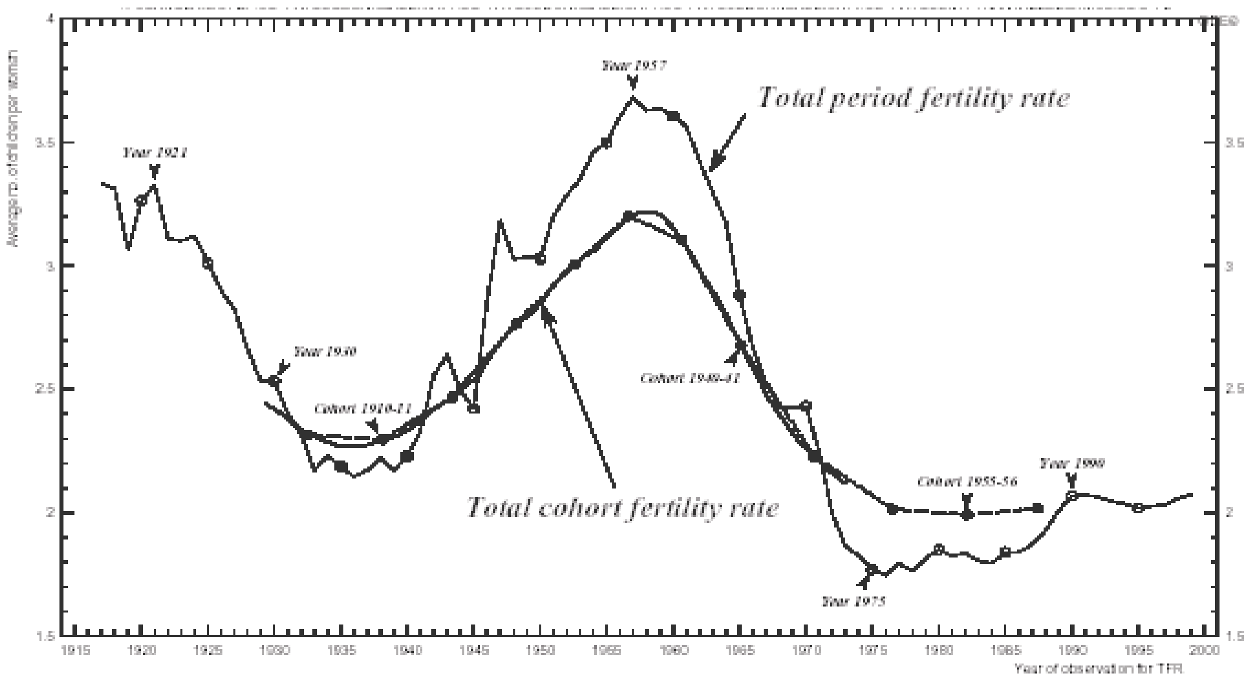
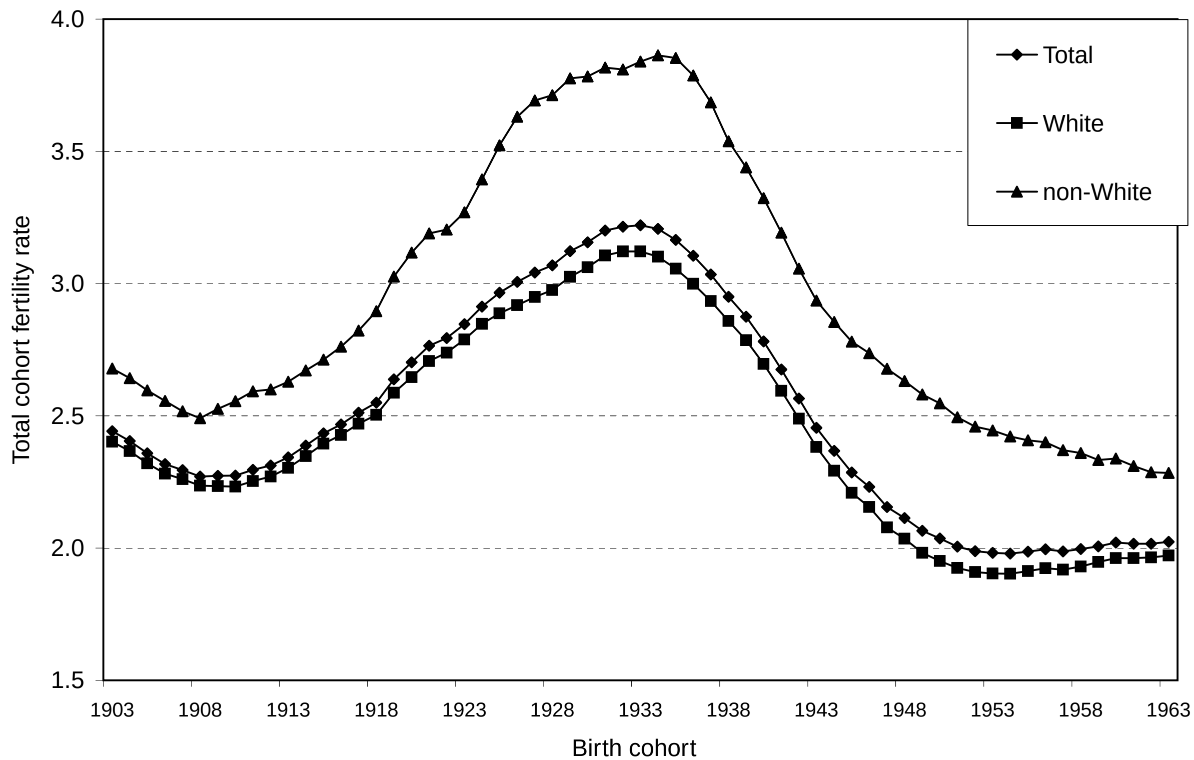


Figure 2. Total period fertility rate, United States, 1917-97, and total cohort fertility rate in United States birth cohorts 1902-60 lagged by the average age in childbearing

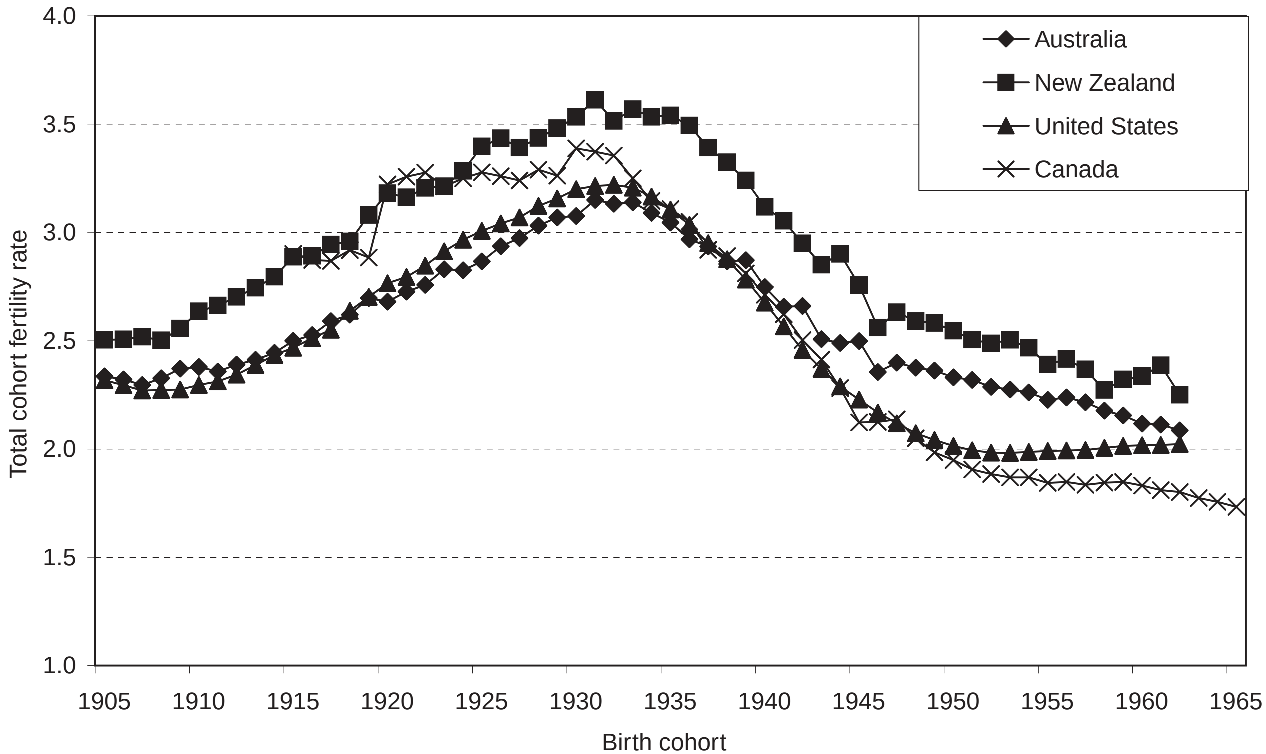


Note: Dotted line indicates that a minor portion of the total cohort fertility rate has been estimated

Figure 3. Total cohort fertility rate, White and non-White population, United States, birth cohorts 1903-1963



Fiugre 4. Total cohort fertility rates, United States and selected overseas countries, birth cohorts 1905–1965



Fiigure 5. Total cohort fertility rate, United States and selected West European countries, birth cohorts 1905-1965

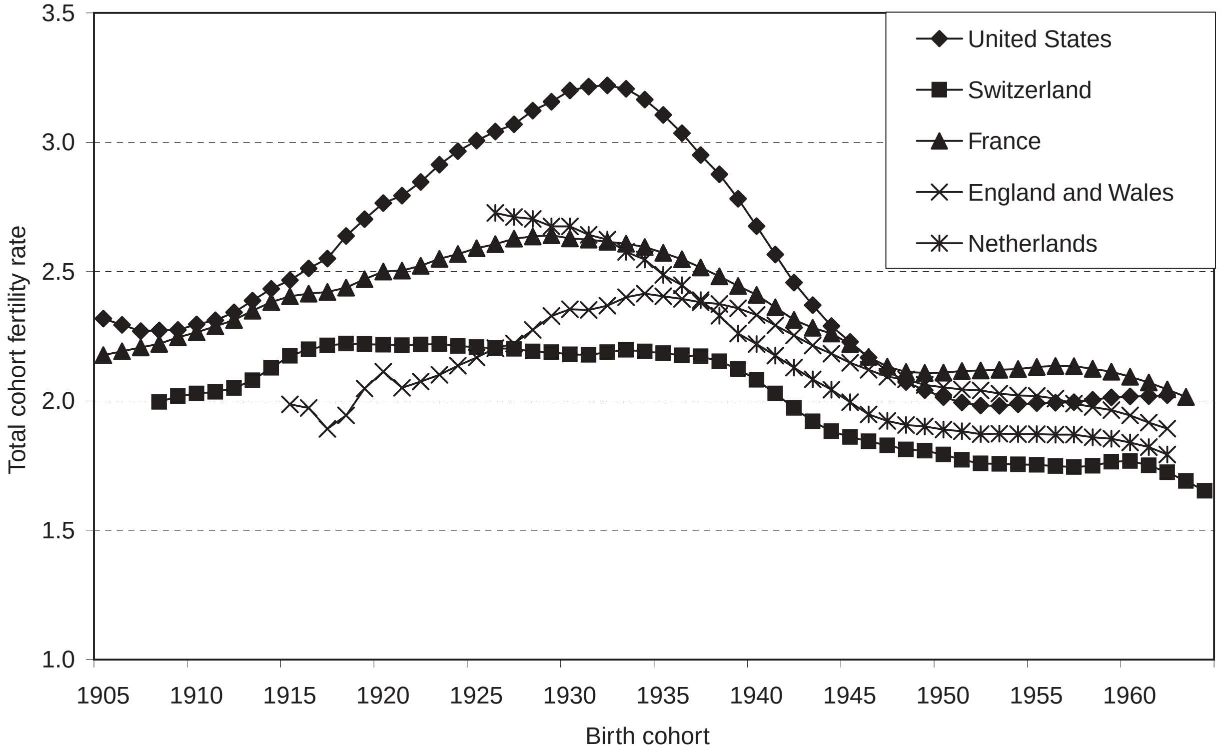


Figure 6. Total cohort fertility rate, United States and selected South European countries, birth cohorts 1905–1965

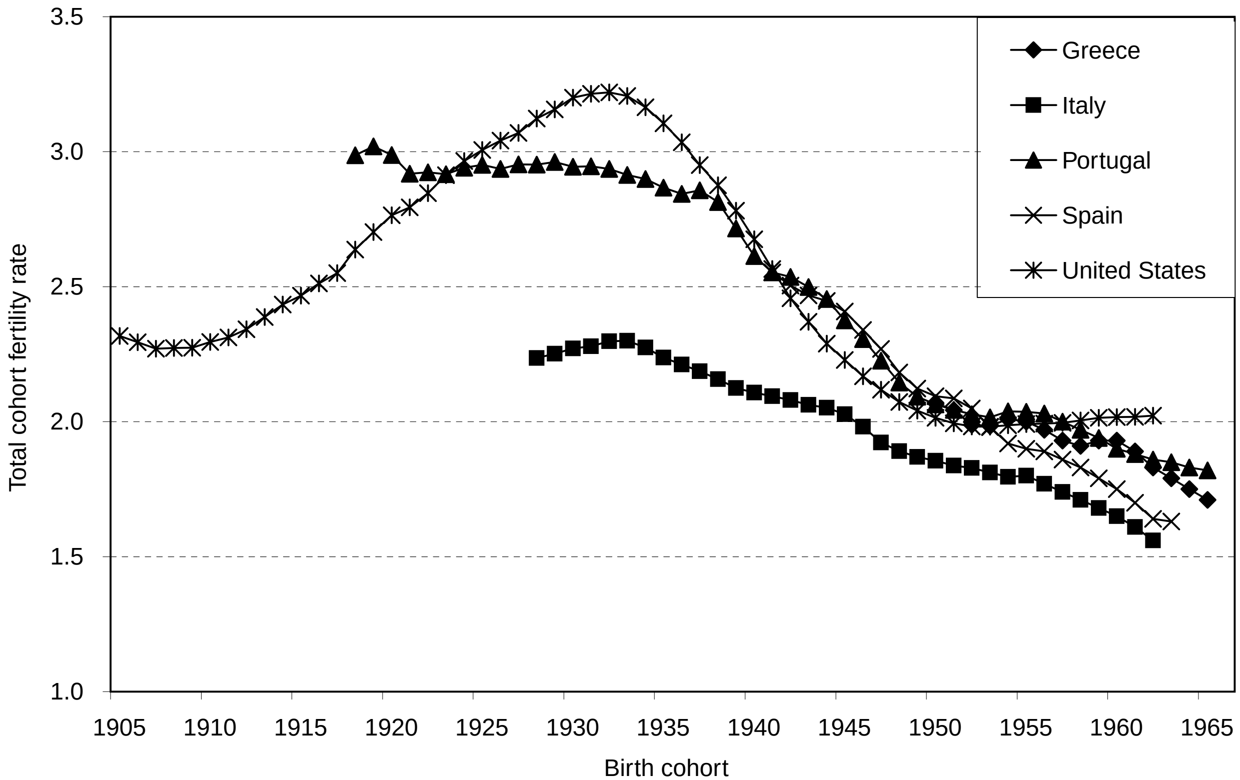


Figure 7. Total cohort fertility rate, United States and selected North European counties, birth cohorts 1905–1965

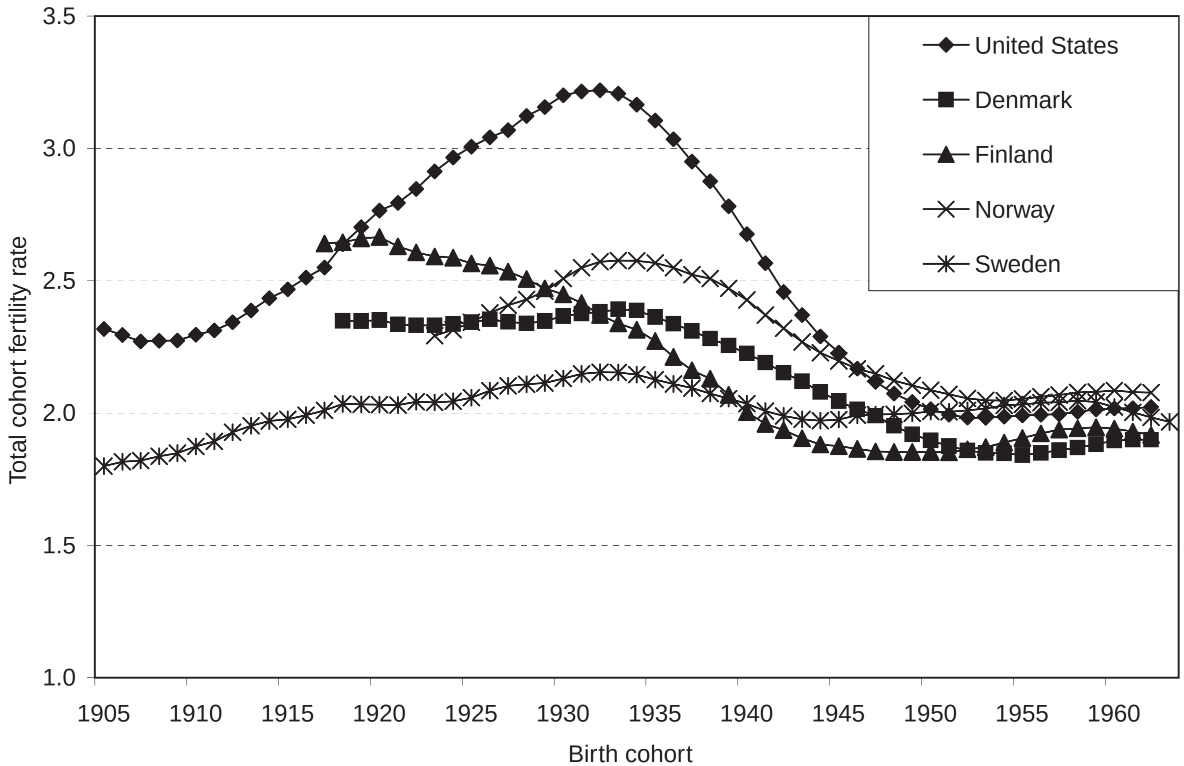


Figure 8. Age-specific fertility rates, United States, birth cohorts 1930–31, 1940–41, 1950–51 and 1960–61

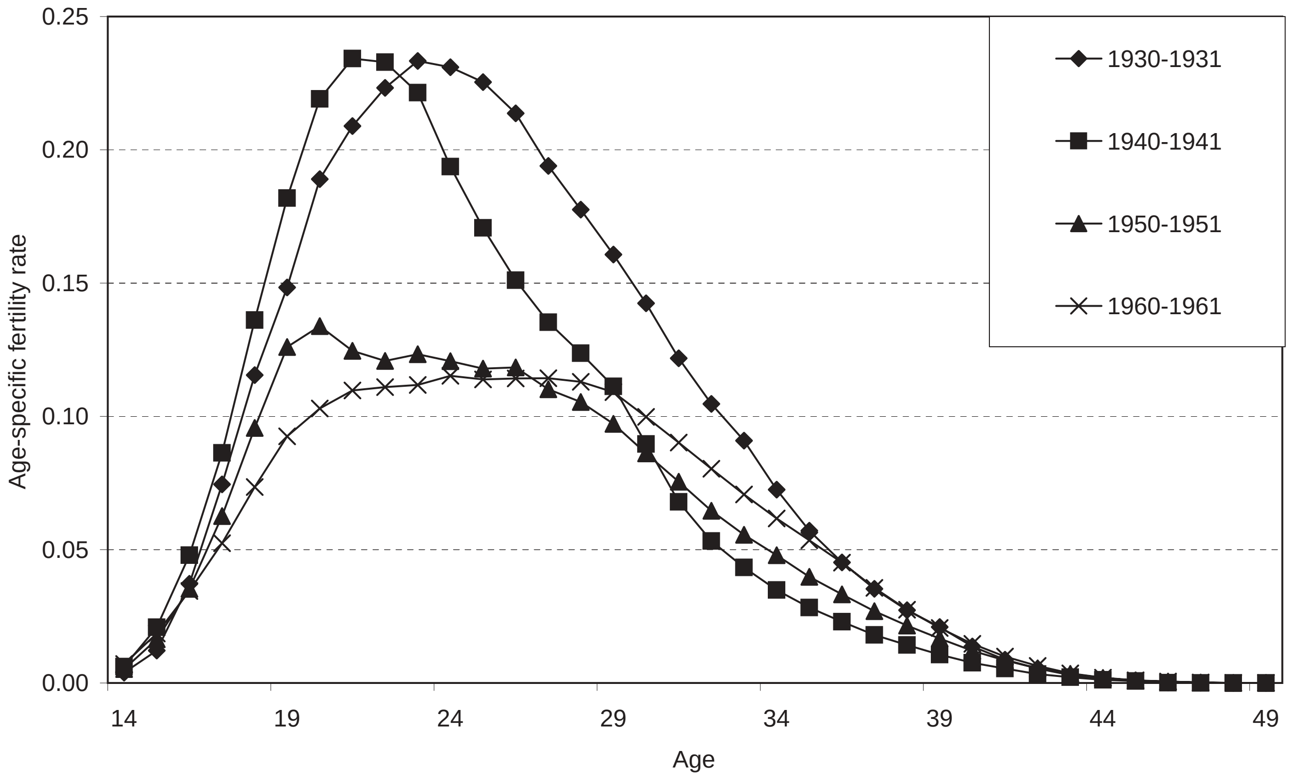
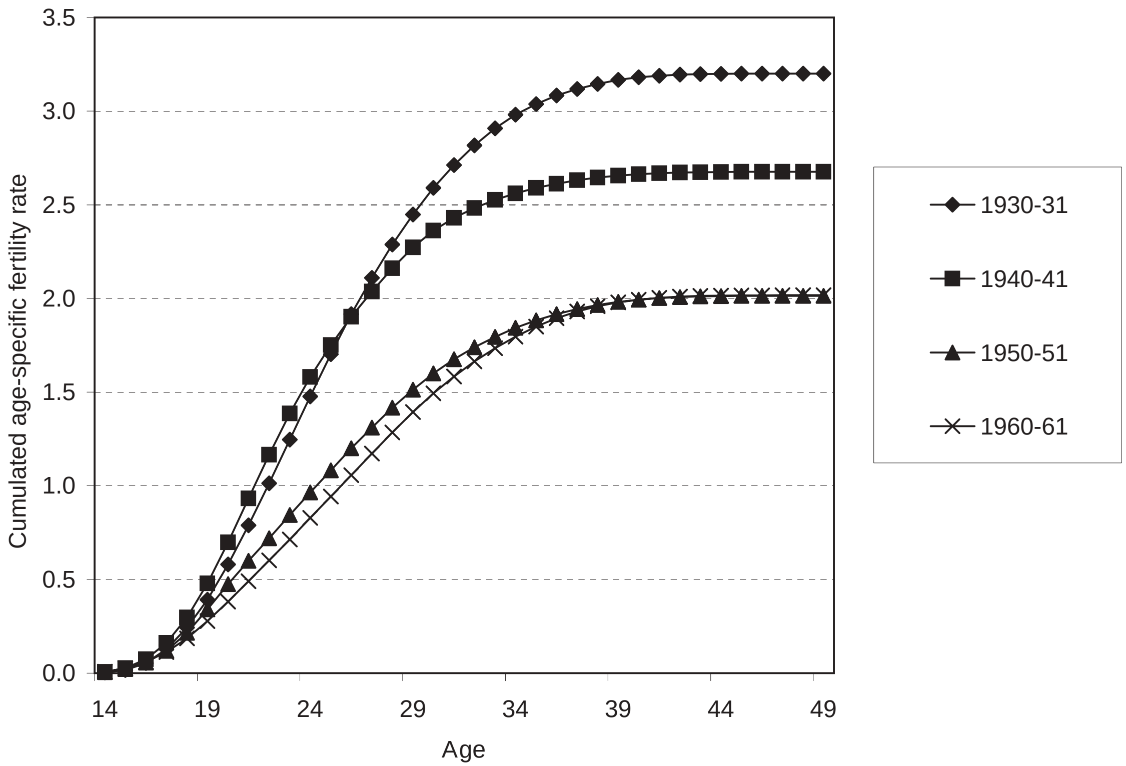
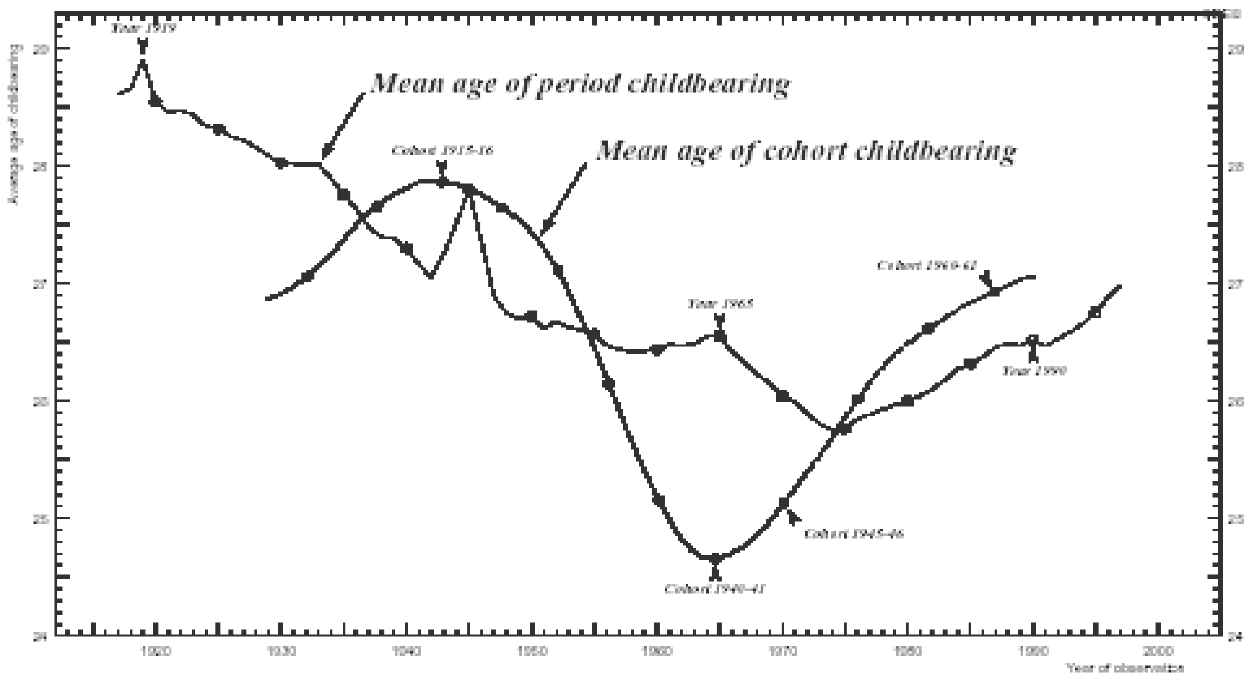


Figure 9. Cumulated age-specific fertility rates, United States, birth cohorts 1930–31, 1940–41, 1950–51 and 1960–61



Firgure 10. Mean age of period childbearing, United States, 1917–97, mean age of cohort childbearing, birth cohorts 1903–63



Note: Dotted line indicates that a minor portion of the mean cohort age of childbearing was estimated

Figure 11. Age-specific fertility rates, United States, birth cohorts 1960–61, 1965–66, 1970–71 and 1975–76

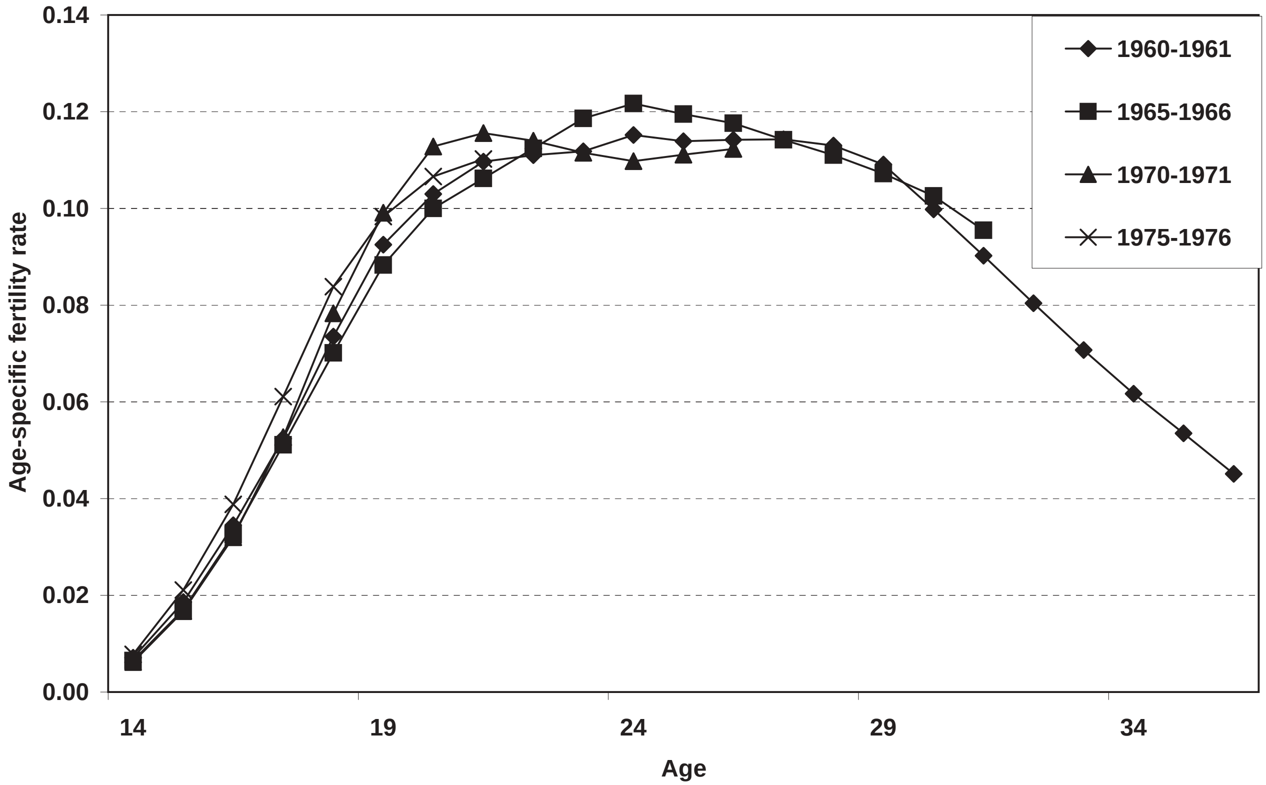


Figure 12. Differences in cumulative age-specific fertility rates between base and subsequent cohorts, United States, women born in 1950–51, 1960–61, 1965–66, 1970–71 and 1975–76

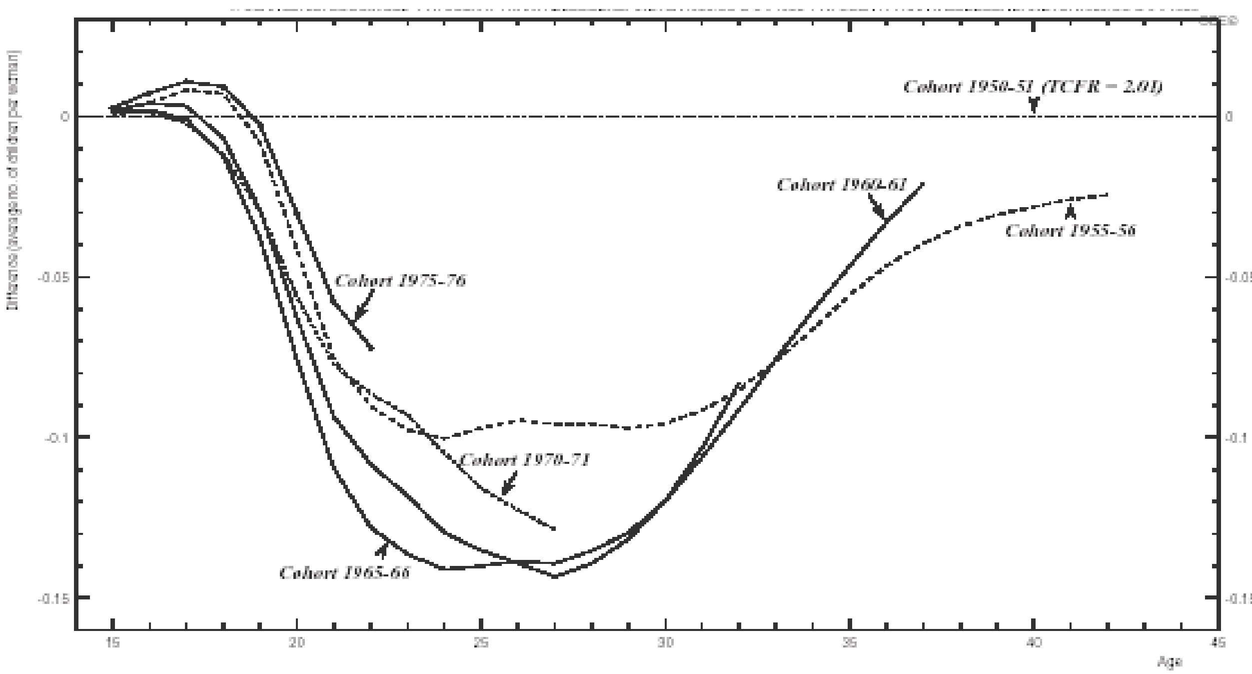


Figure 13. The proportion of childbearing realized by exact age 27, United States and selected "Western" countries, birth cohorts 1930–31, 1940–41, 1950–51 and 1960–61

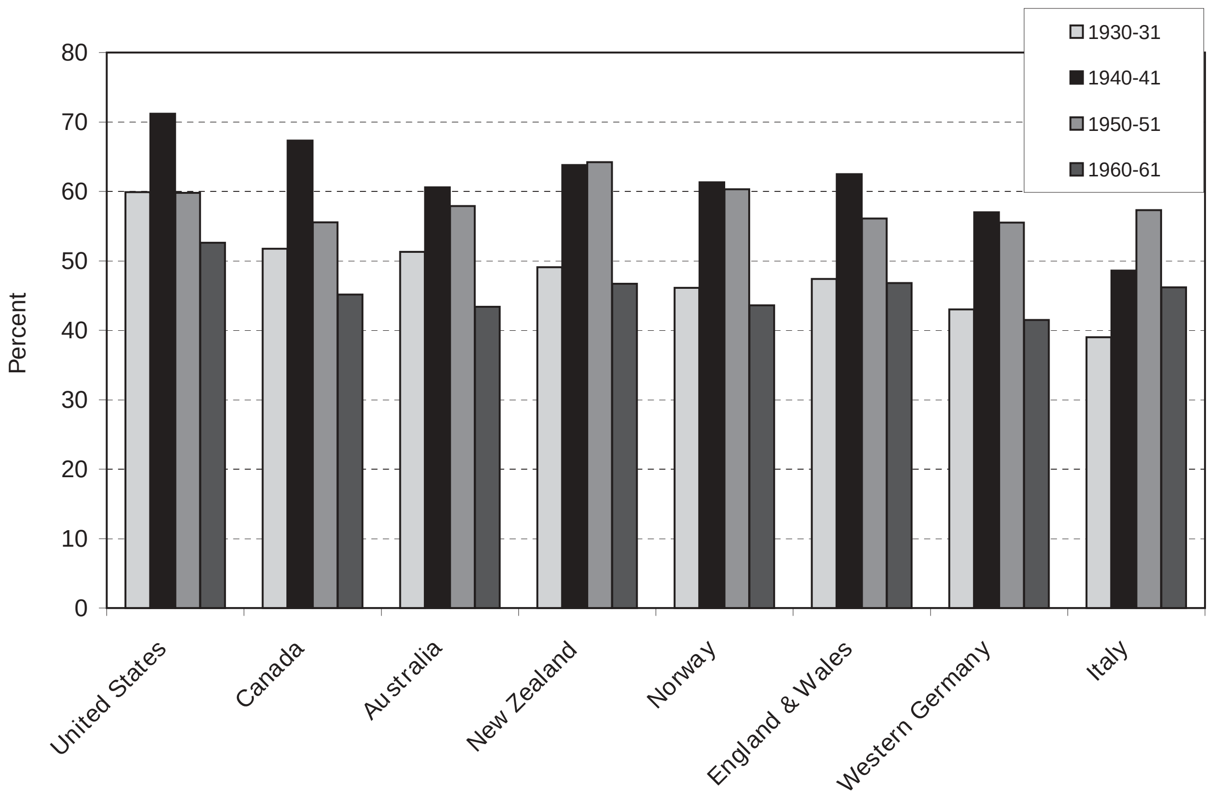


Figure 14. Parity distribution at age 40, United States, birth cohorts 1930–1960

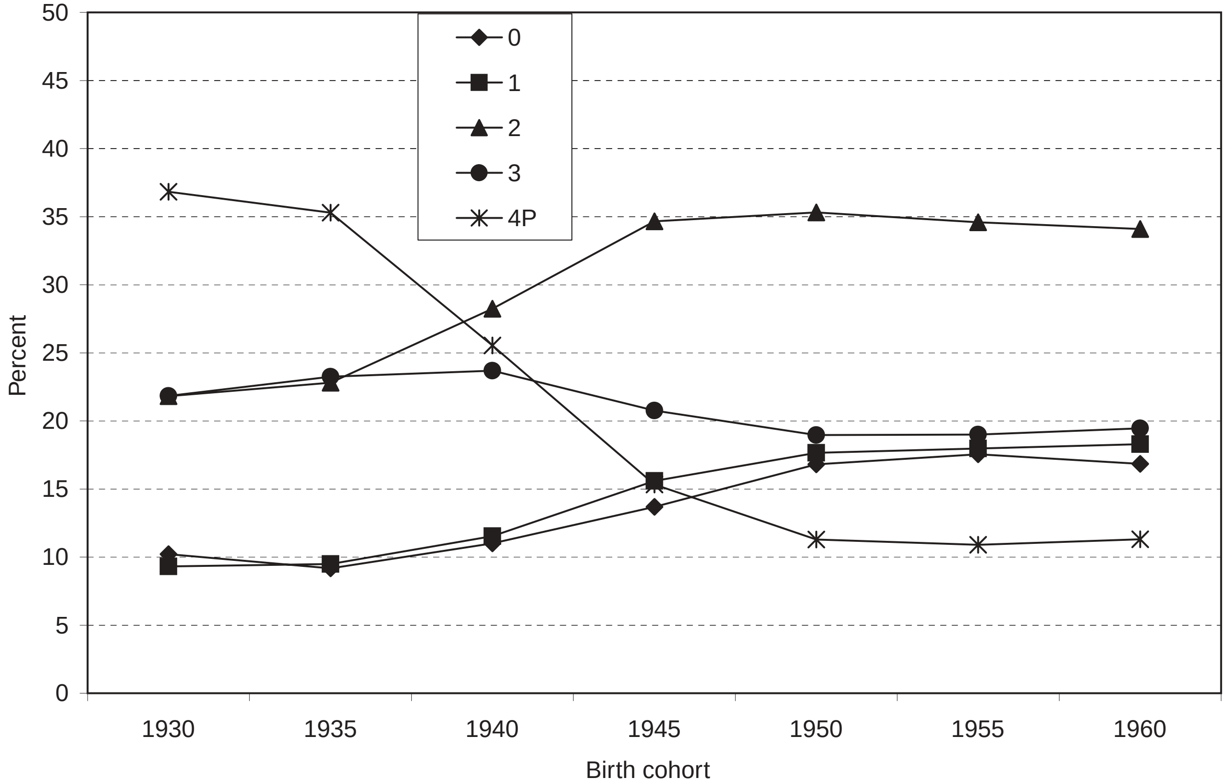


Figure 15. Parity distribution at age 40, Norway, birth cohorts 1935–1960

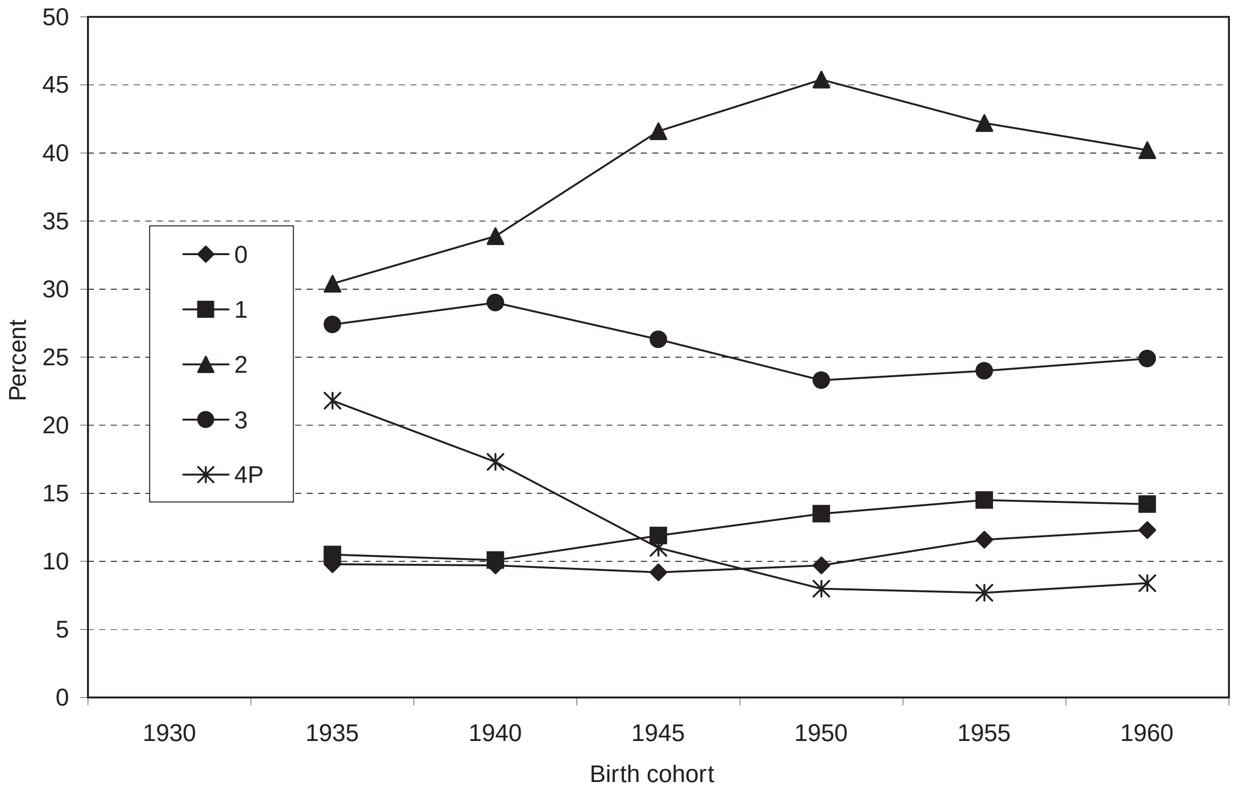


Figure 16. Parity distribution at age 50, West Germany, birth cohorts 1940–1960

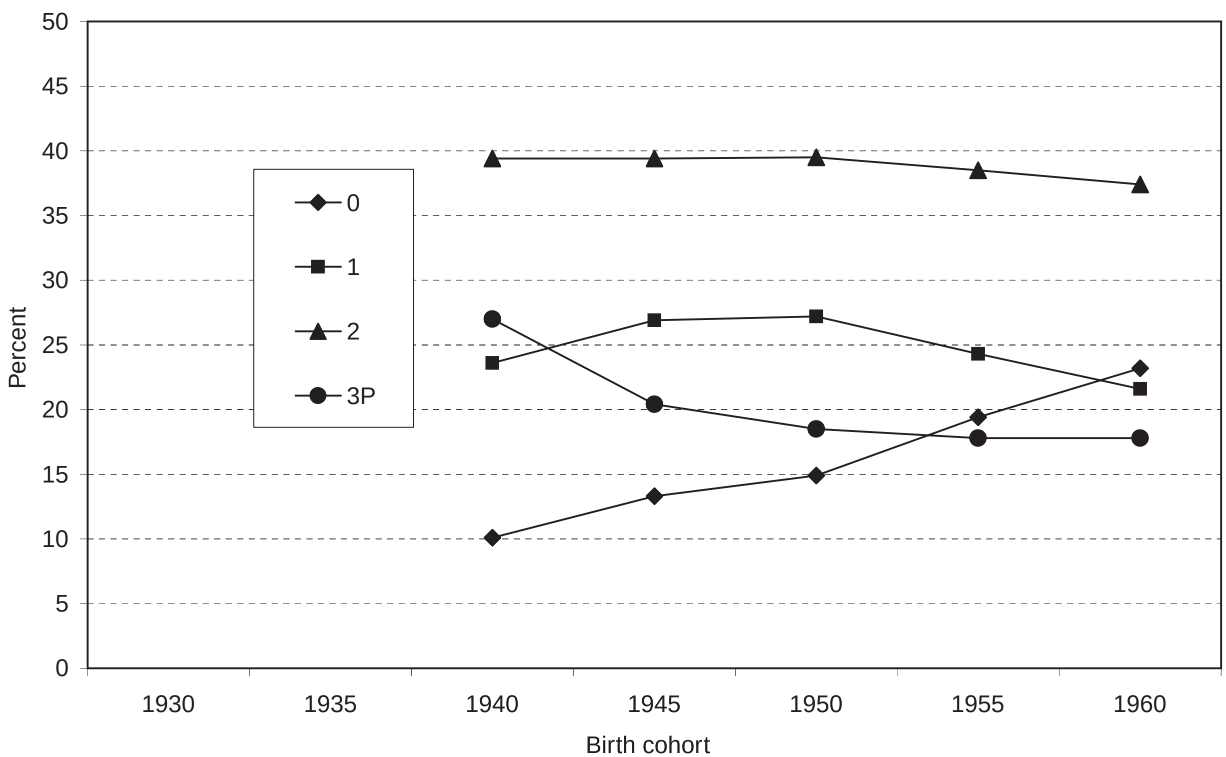


Figure 17. Parity distribution at age 40, Italy, birth cohorts 1930–1960

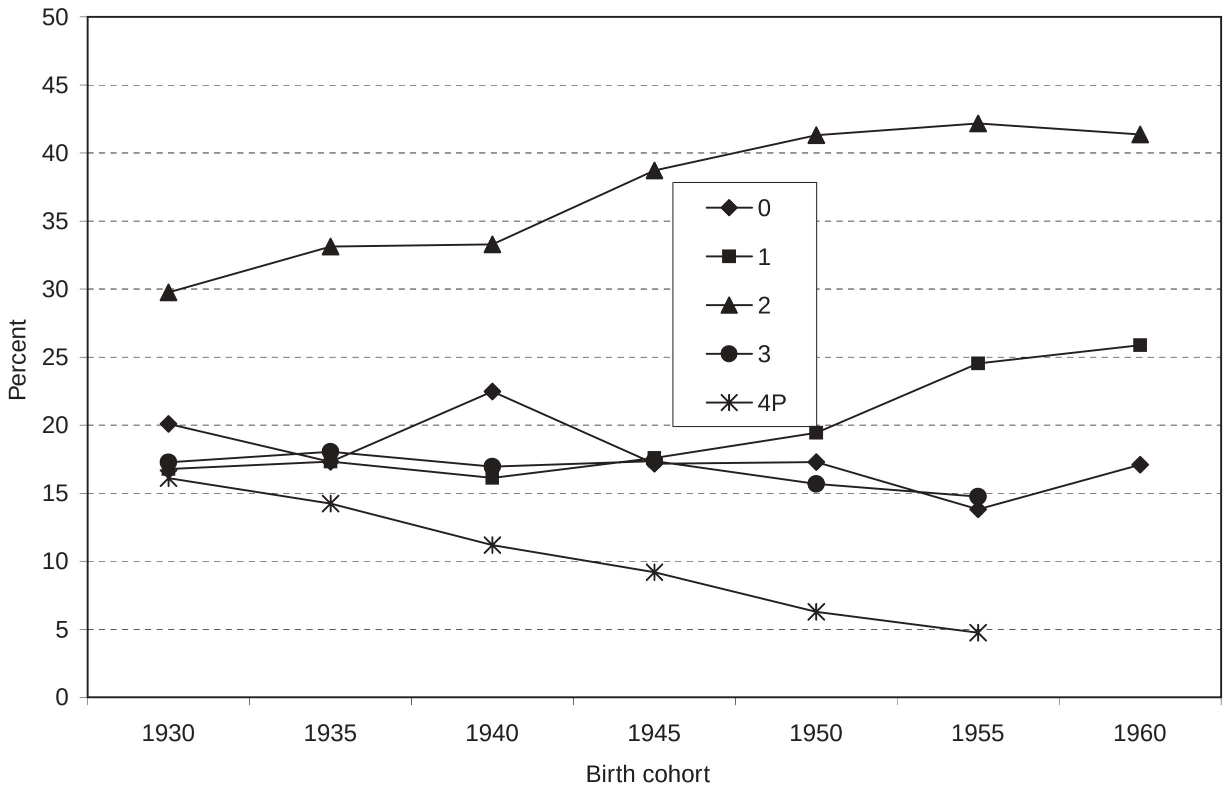


Figure 18. Parity distribution at age 40, Netherlands, birth cohorts 1930–1960

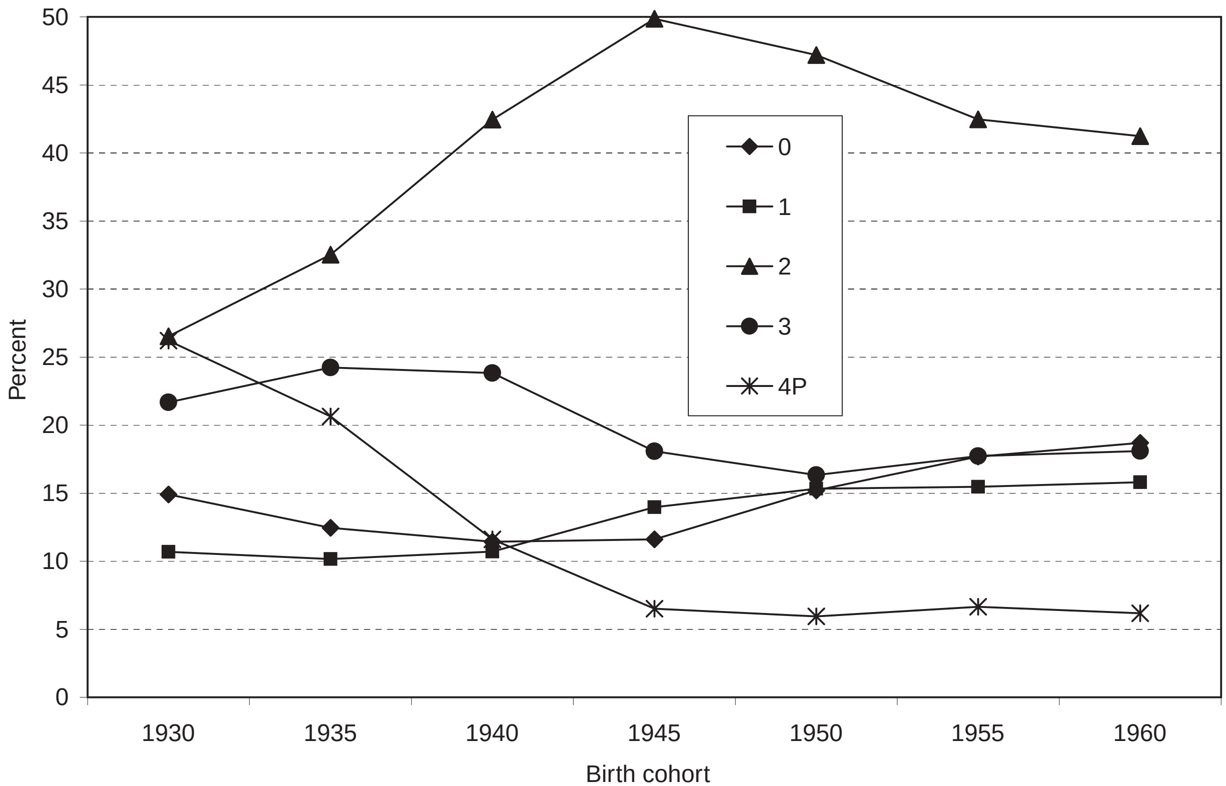


Figure 19. Parity distribution at age 40, Finland, birth cohorts 1930–1960

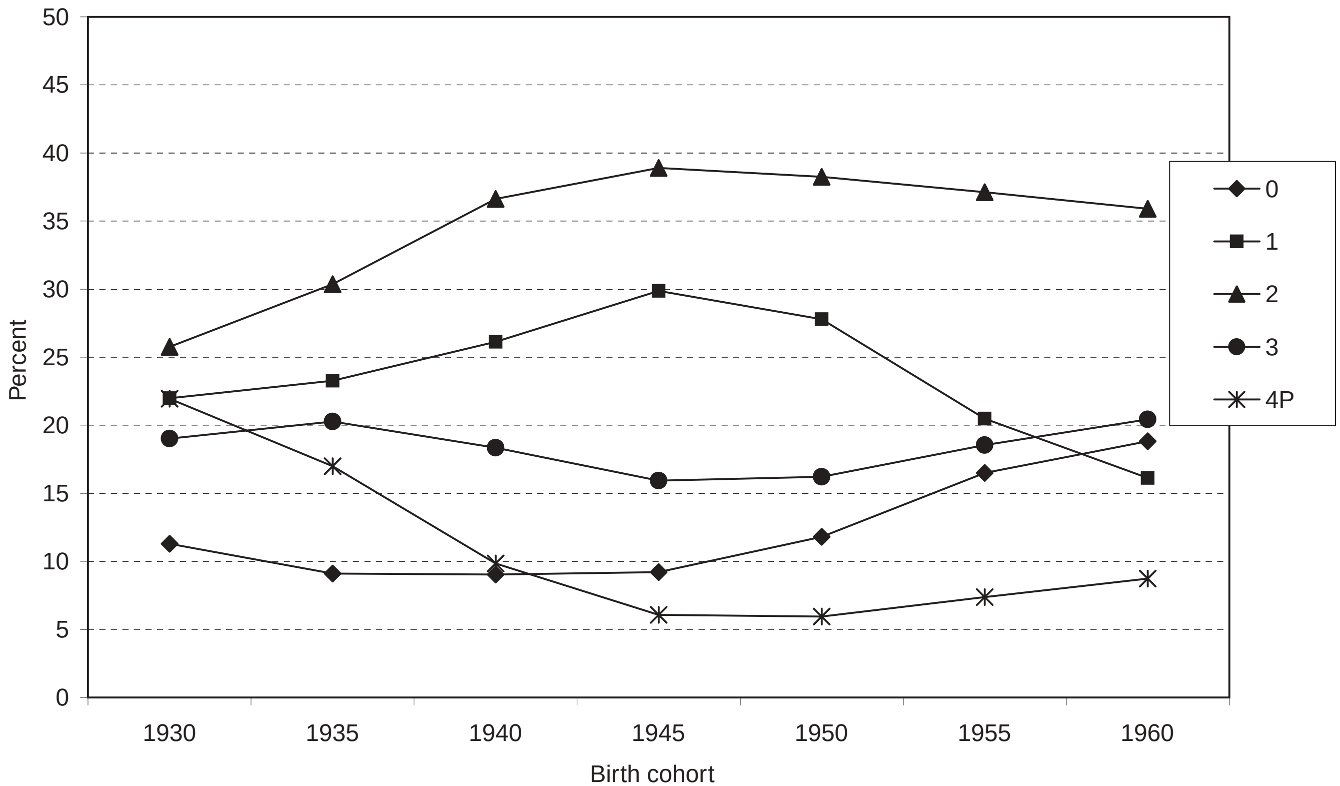


Figure 20. Parity progression ratios, United States, birth cohorts 1930–1966

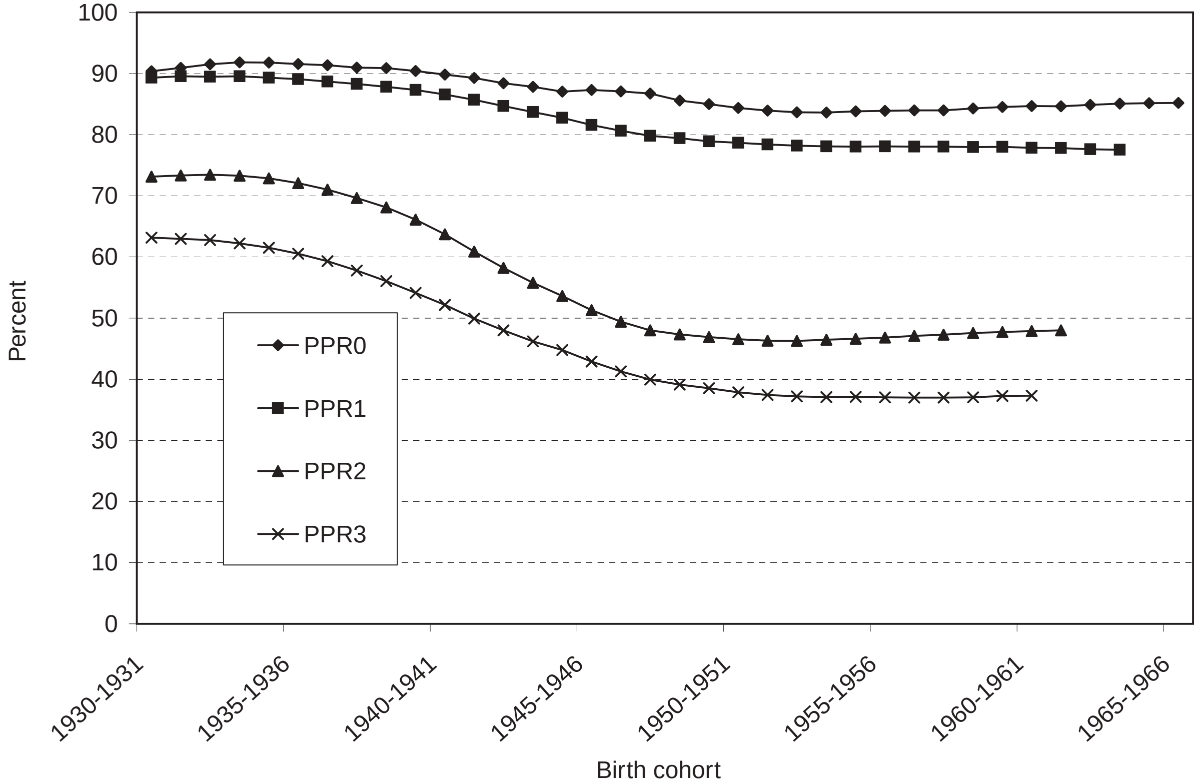


Figure 21. Parity progression ratios, Finland, birth cohorts 1925–26 to 1964–65

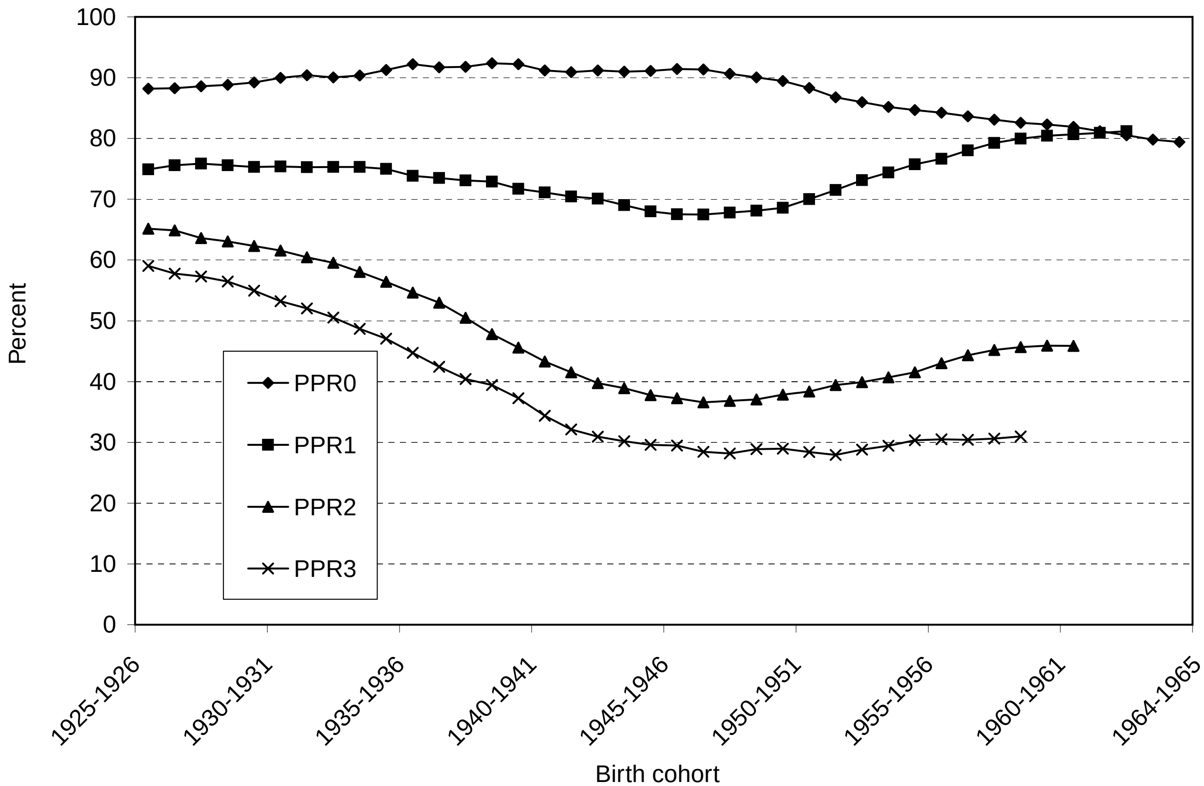


Figure 22. Parity progression ratios, Netherlands, birth cohorts 1935–1966

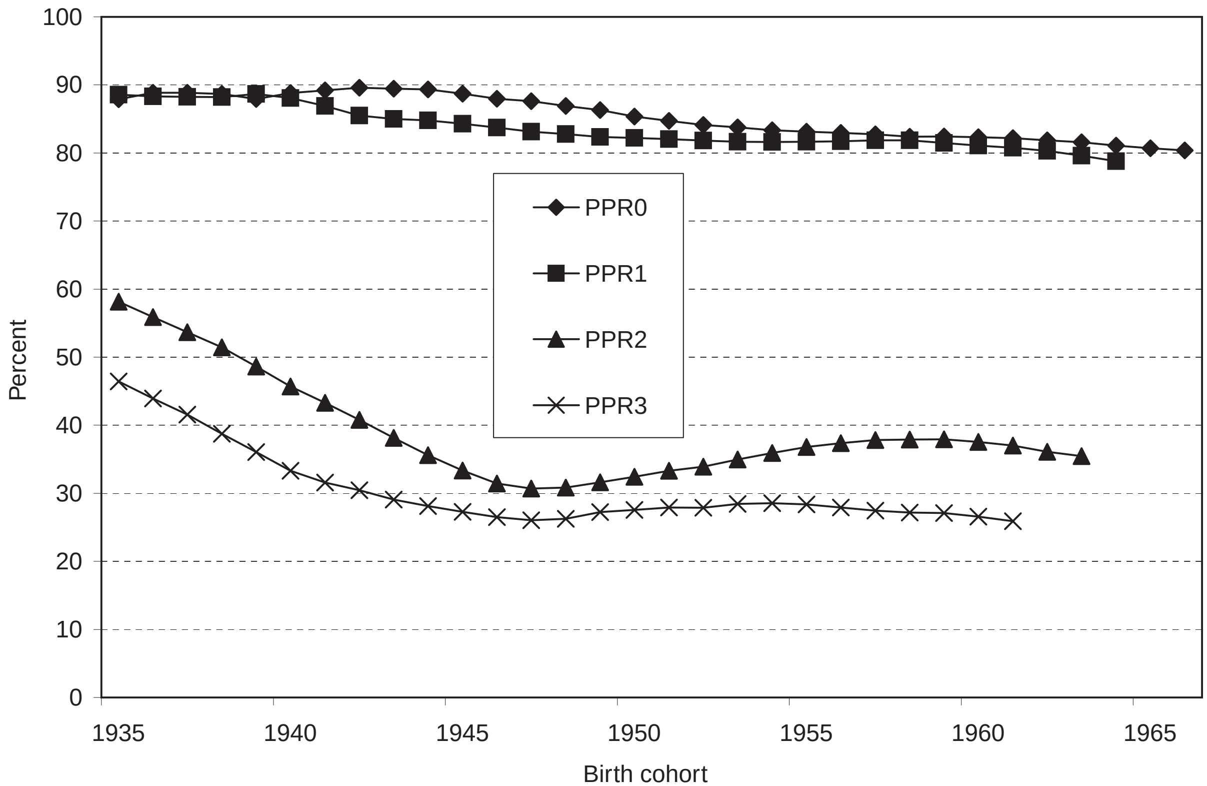


Figure 23. Age-specific fertility rates, United States, Whites, birth cohorts 1931, 1941, 1951 and 1961

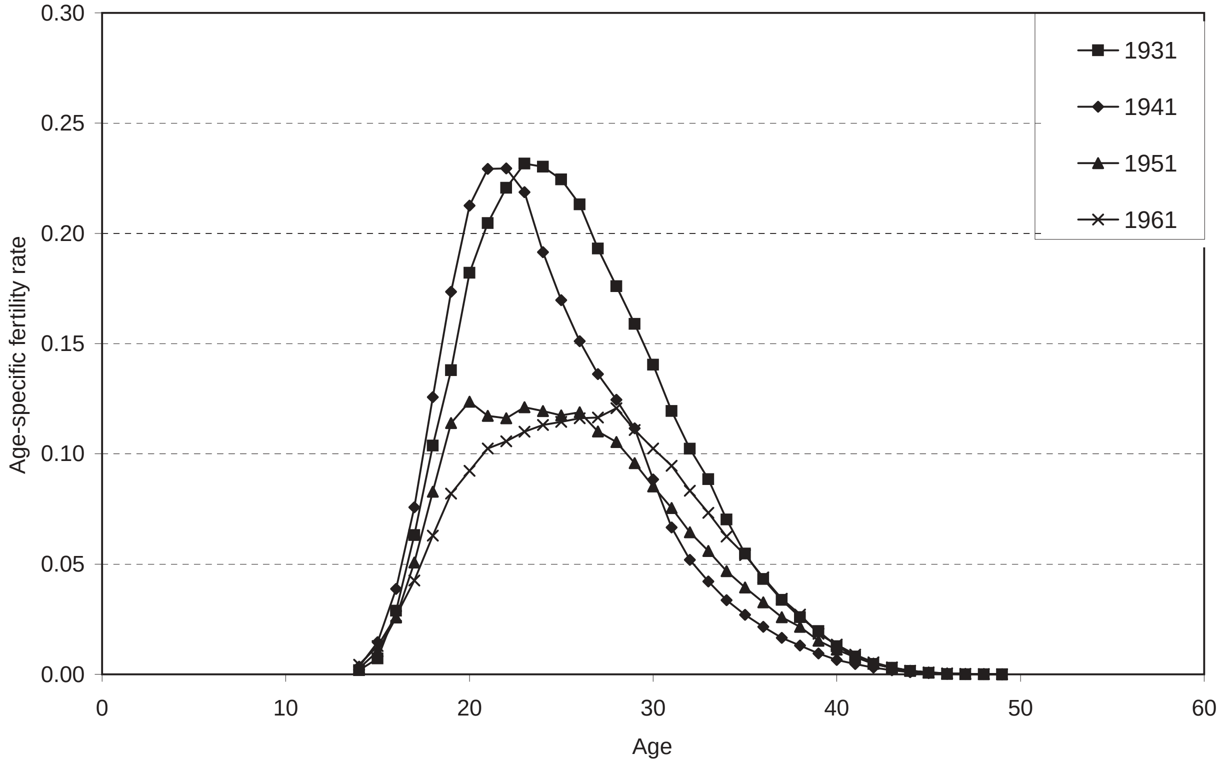


Figure 24. Age specific fertility rates, United States non-Whites, birth cohorts 1931, 1941, 1951 and 1961

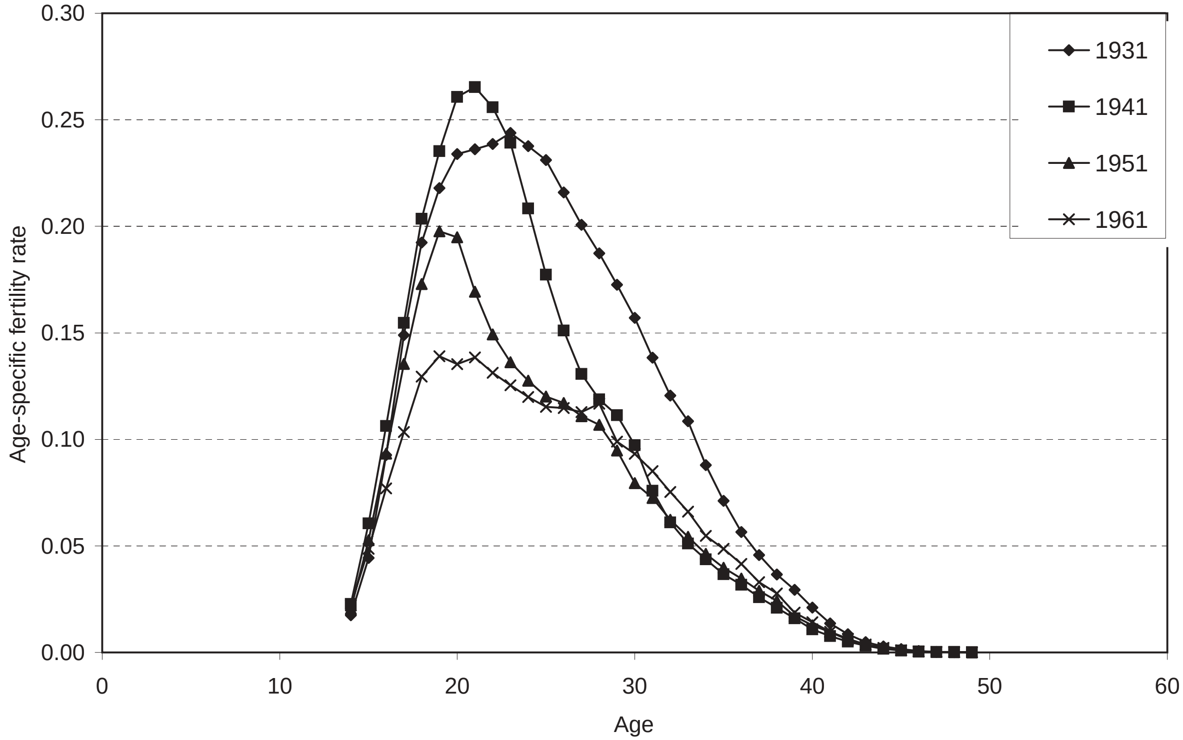


Figure 25. Mean age of cohort childbearing, by race, United States, birth cohorts, 1903–1964

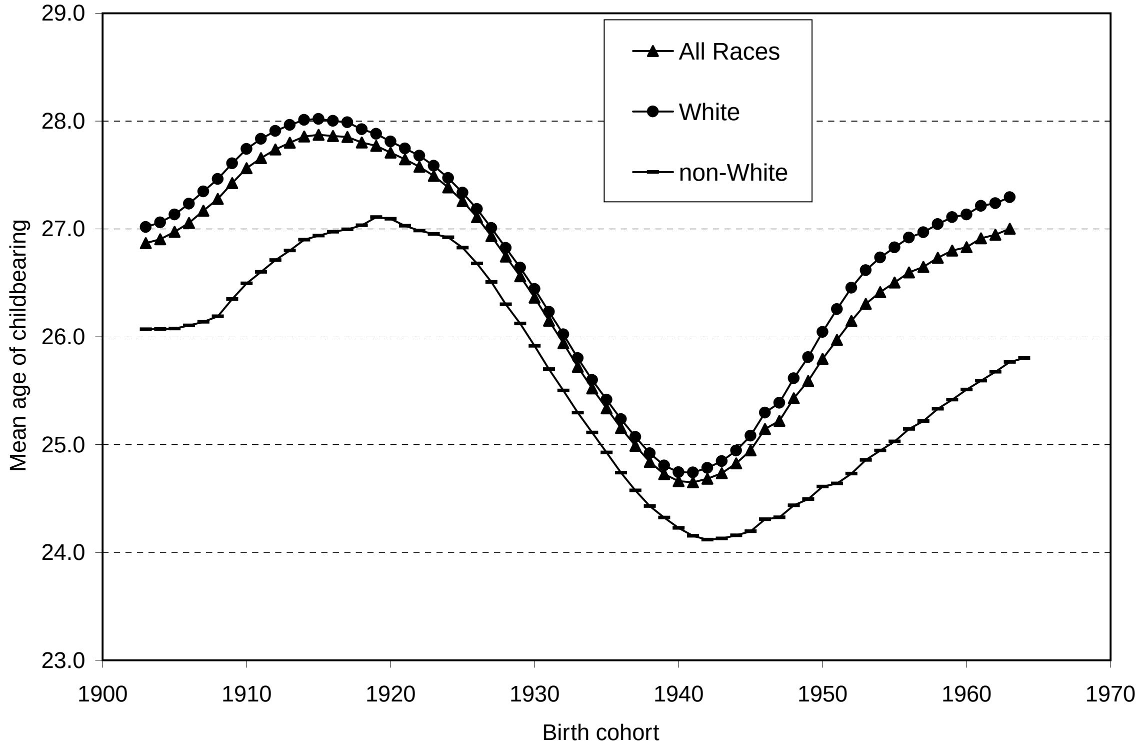


Figure 26. Age-specific fertility rates, United States, Whites, birth cohorts 1961, 1966, 1971 and 1976

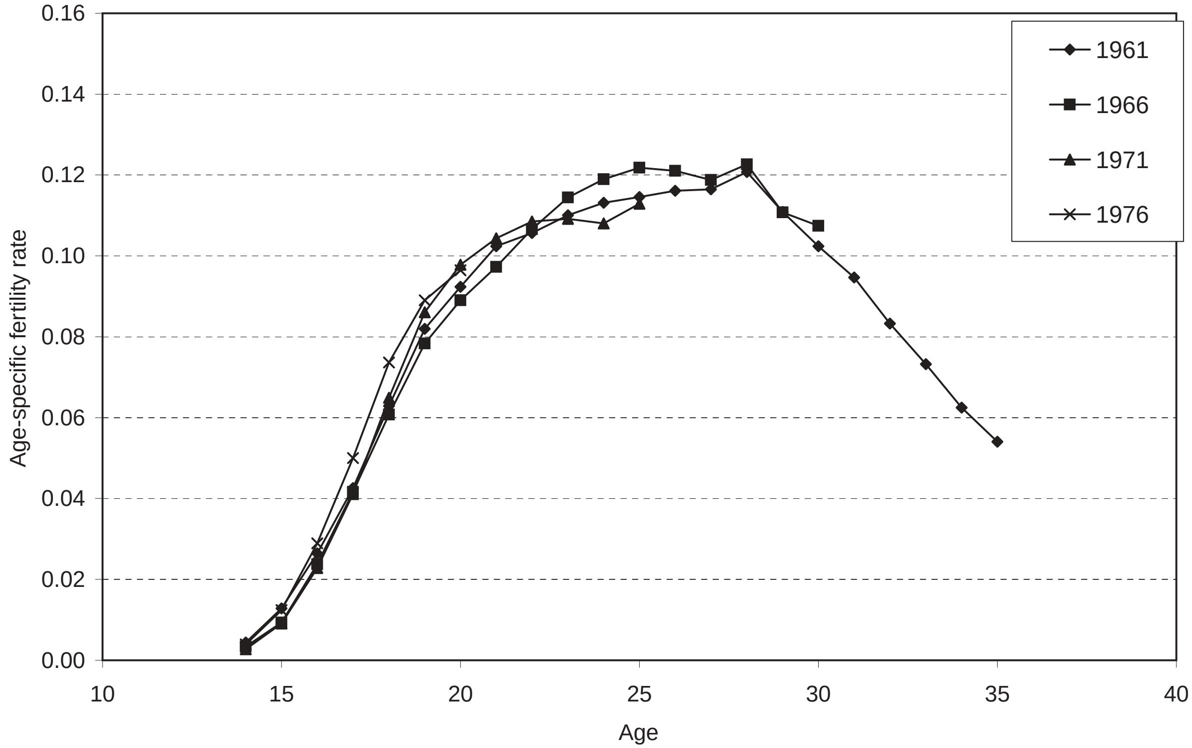


Figure 27. Age-specific fertility rates, United States, non-Whites, birth cohorts 1961, 1966, 1971 and 1976

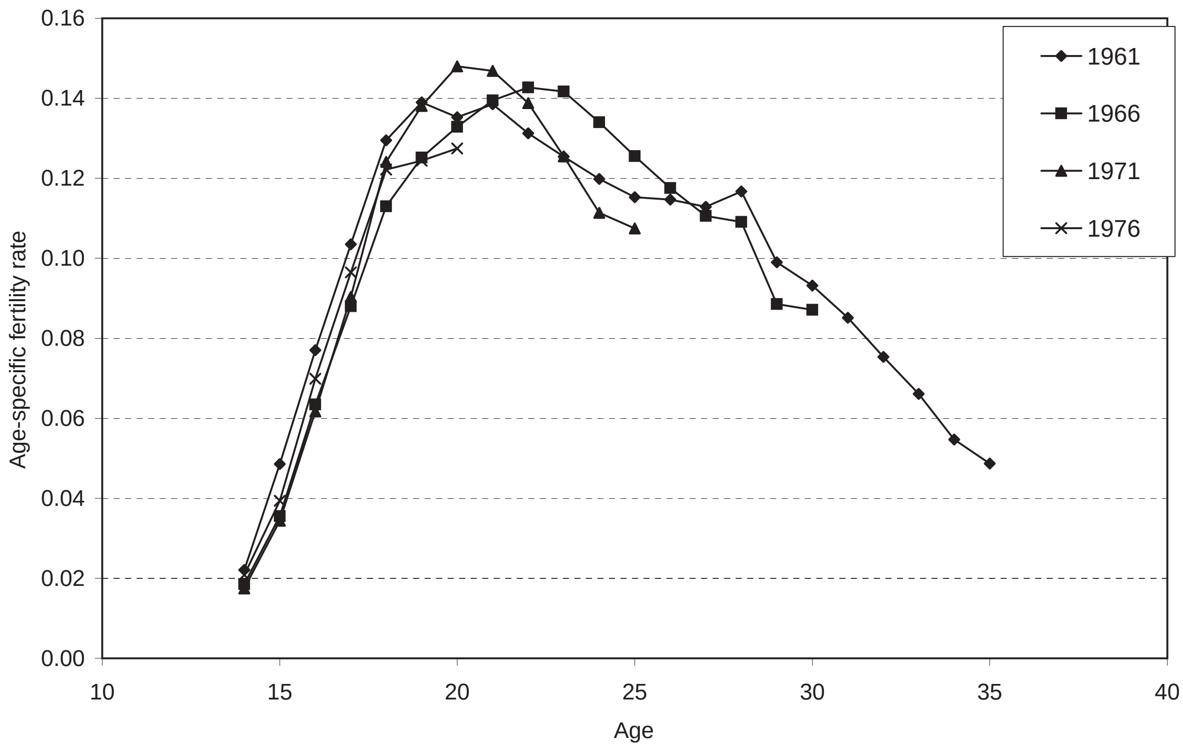


Figure 28. Differences in cumulative age-specific fertility rates between base and subsequent cohorts, United States, White women born in 1951 (base) and 1956, 1961, 1966, 1971 and 1976

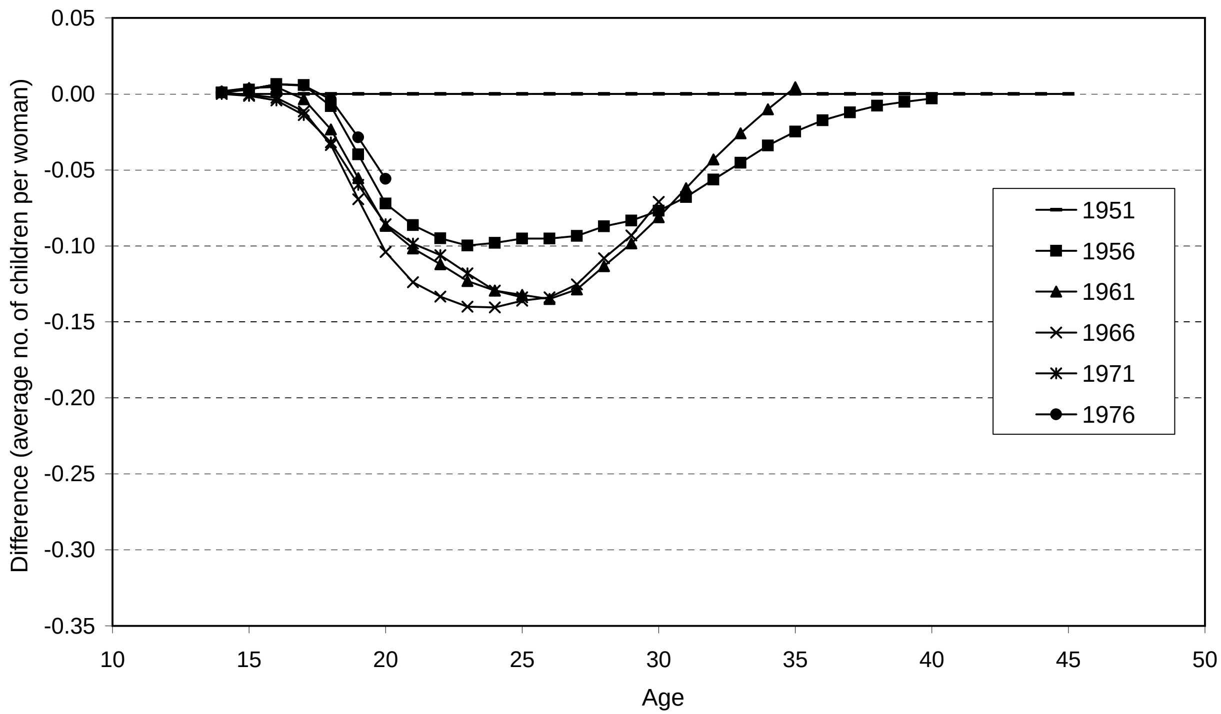


Figure 29. Differences in cumulative age-specific fertility rates between base and subsequent cohorts, United States, non-White women born in 1951 (base) and 1956, 1961, 1966, 1971 and 1976

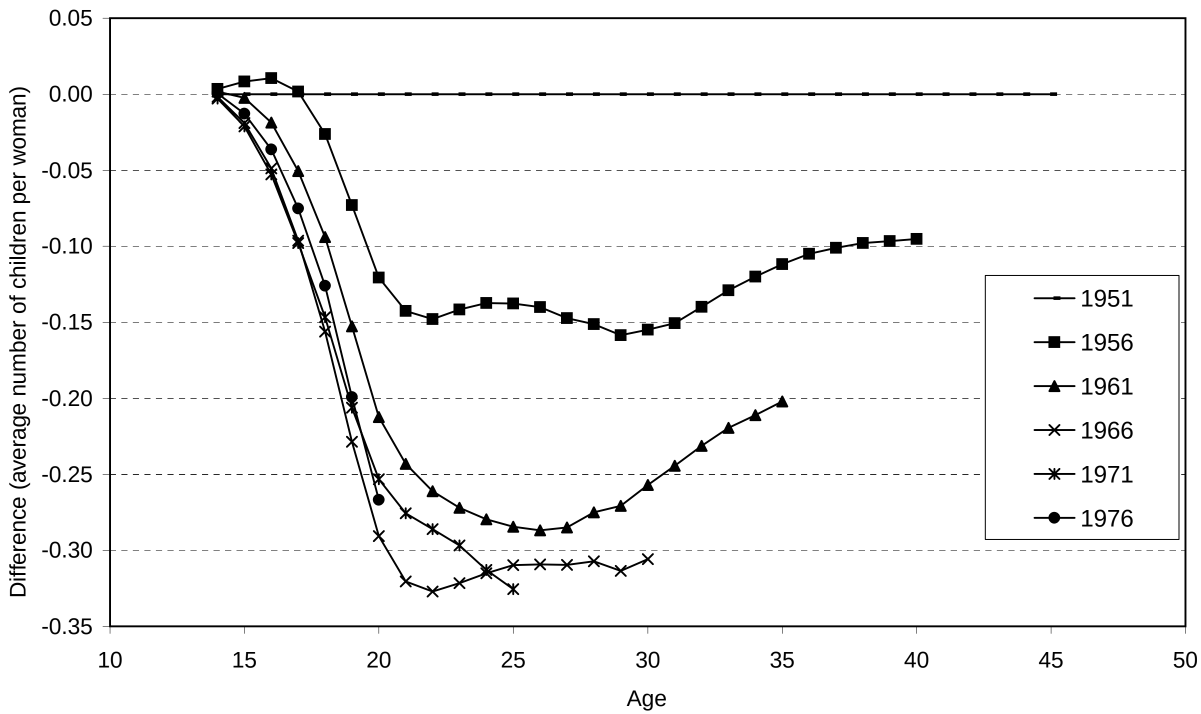


Figure 30. Parity distribution at age 40, United States, White women, birth cohorts 1903–1965

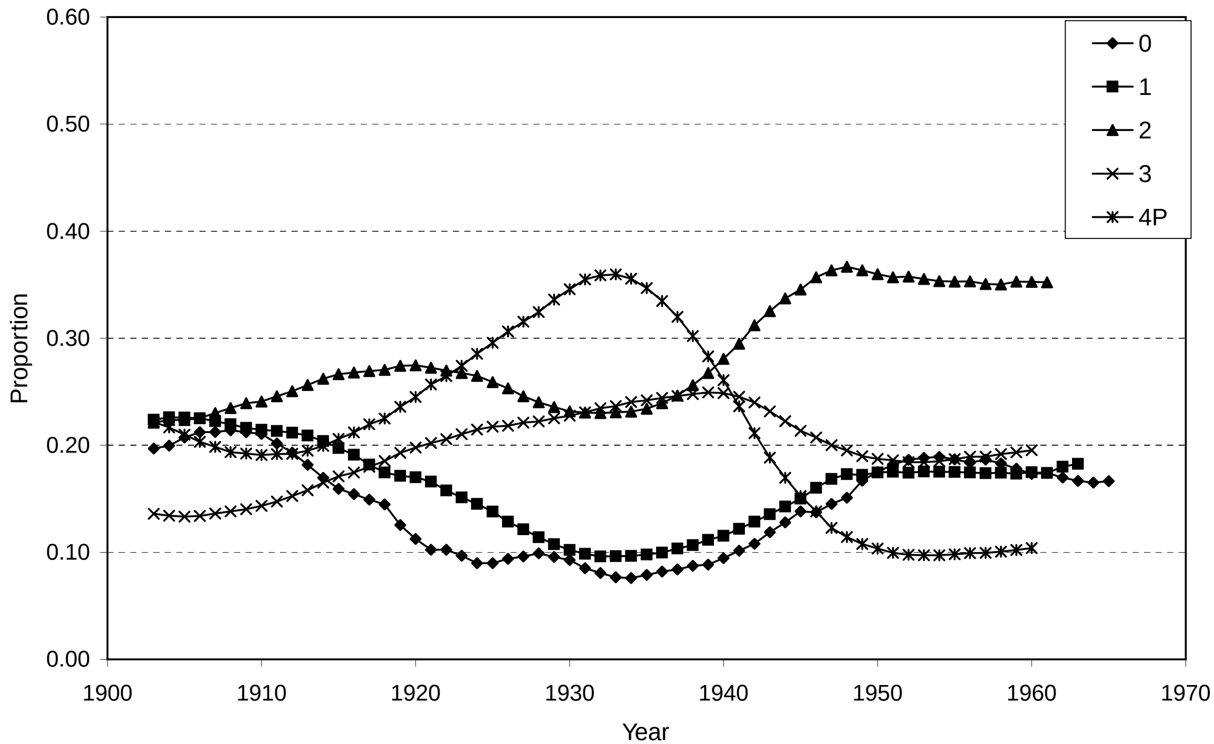


Figure 31. Parity distribution at age 40, United States, non-White women, birth cohorts 1903–1966

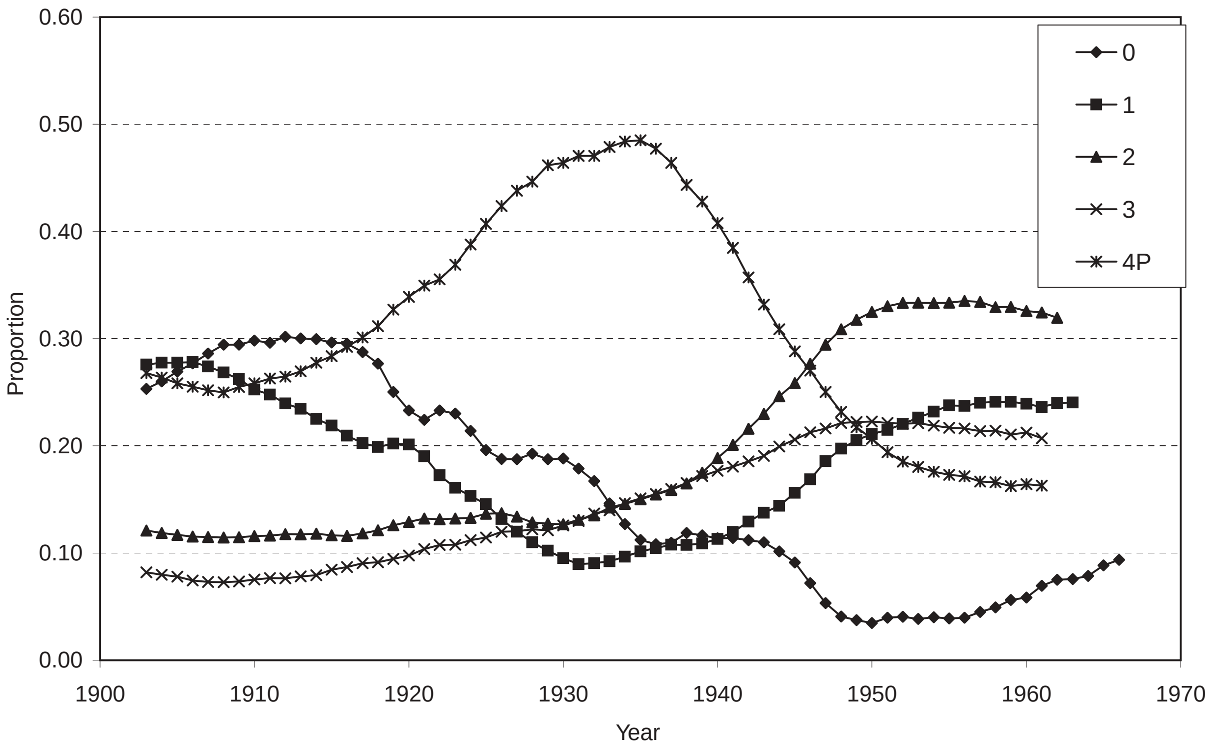


Figure 32. Parity progression ratios, United States, White women, birth cohorts 1903–1965

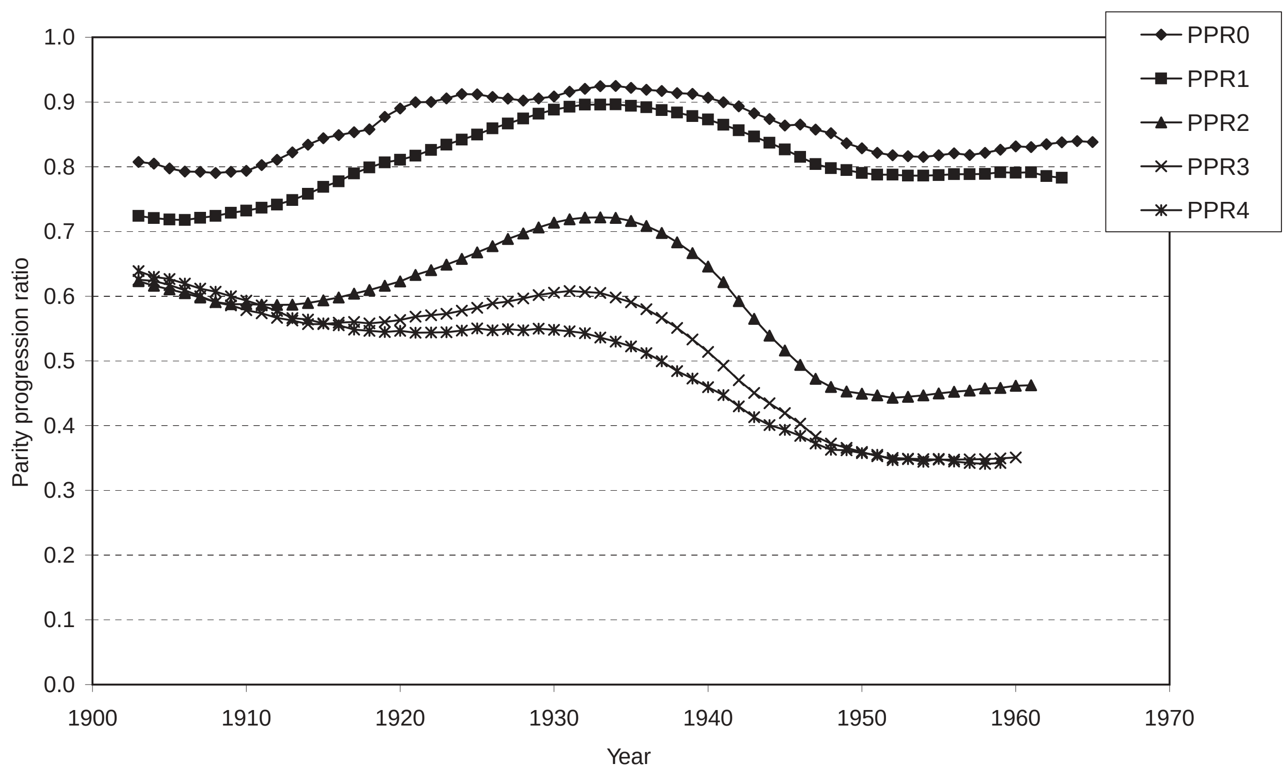


Figure 33. Parity progression ratios, United States, non-White women, birth cohorts 1903-1966

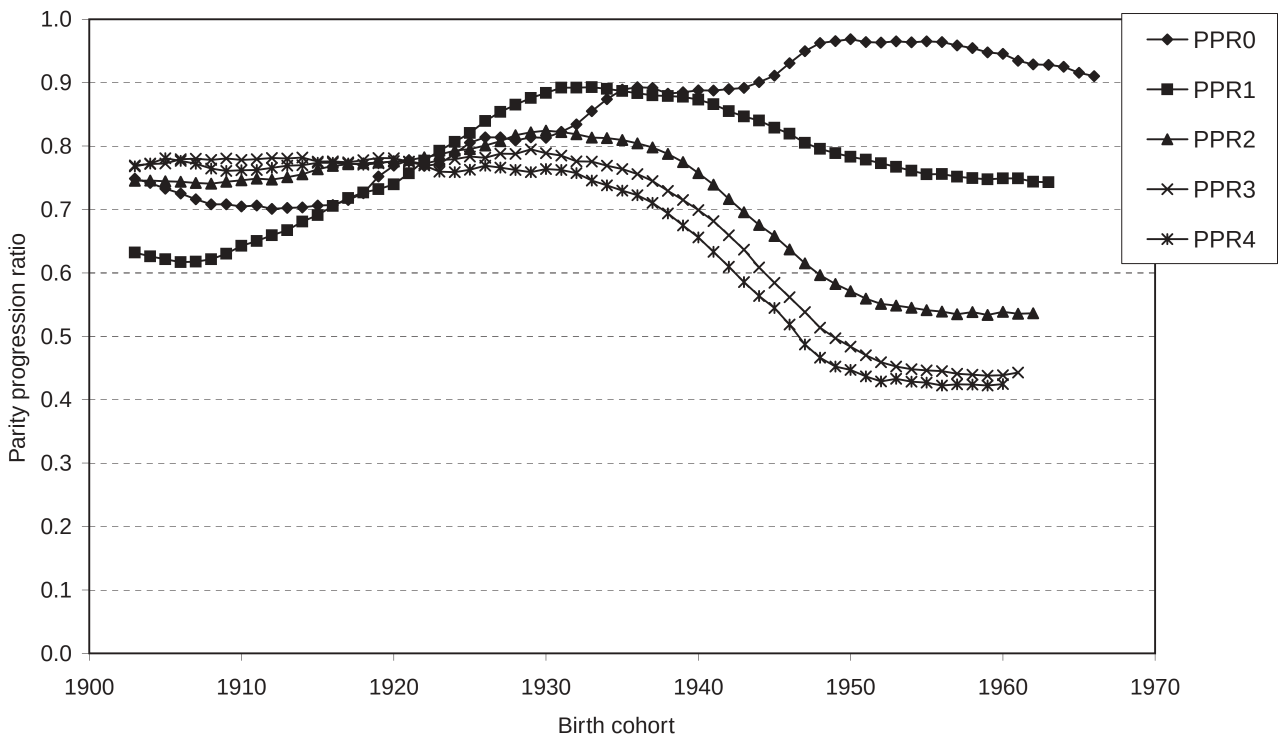


Figure 34. Women 40-44 years old of certain race, ethnicity or nativity by income category, United States, CPS 1998

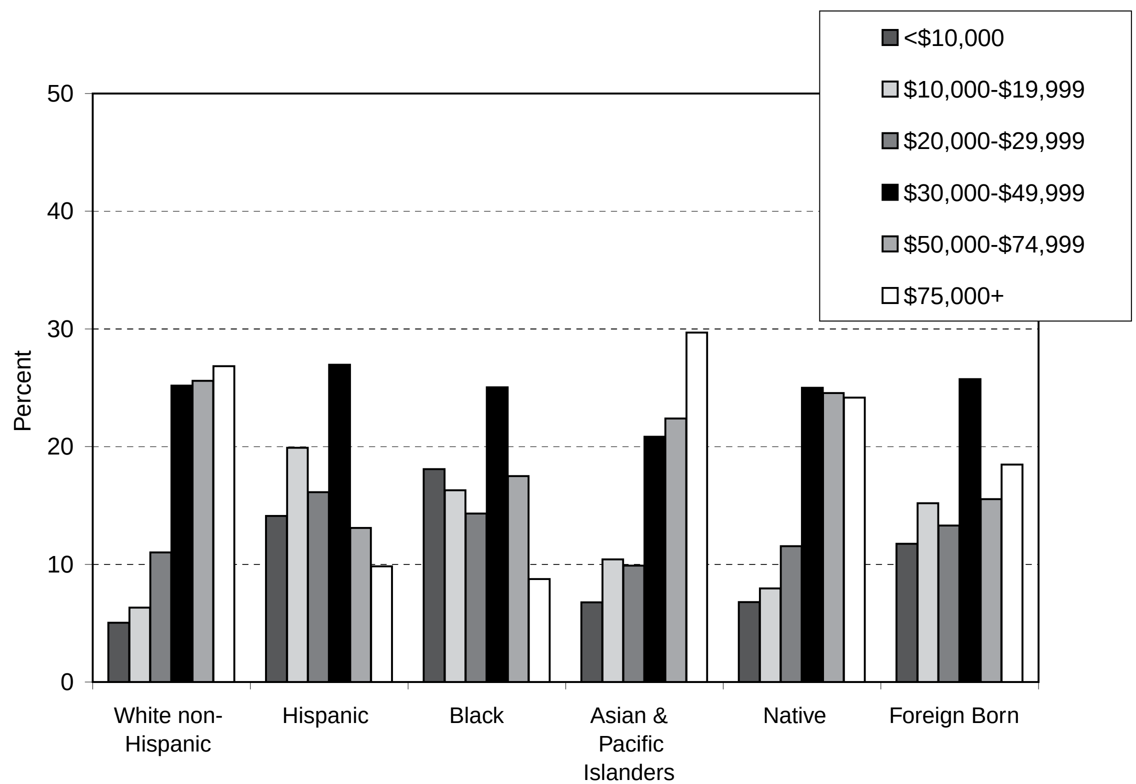


Figure 35. Women 40-44 years old of certain race, ethnicity, and nativity, by educational attainment, United States, CPS 1998

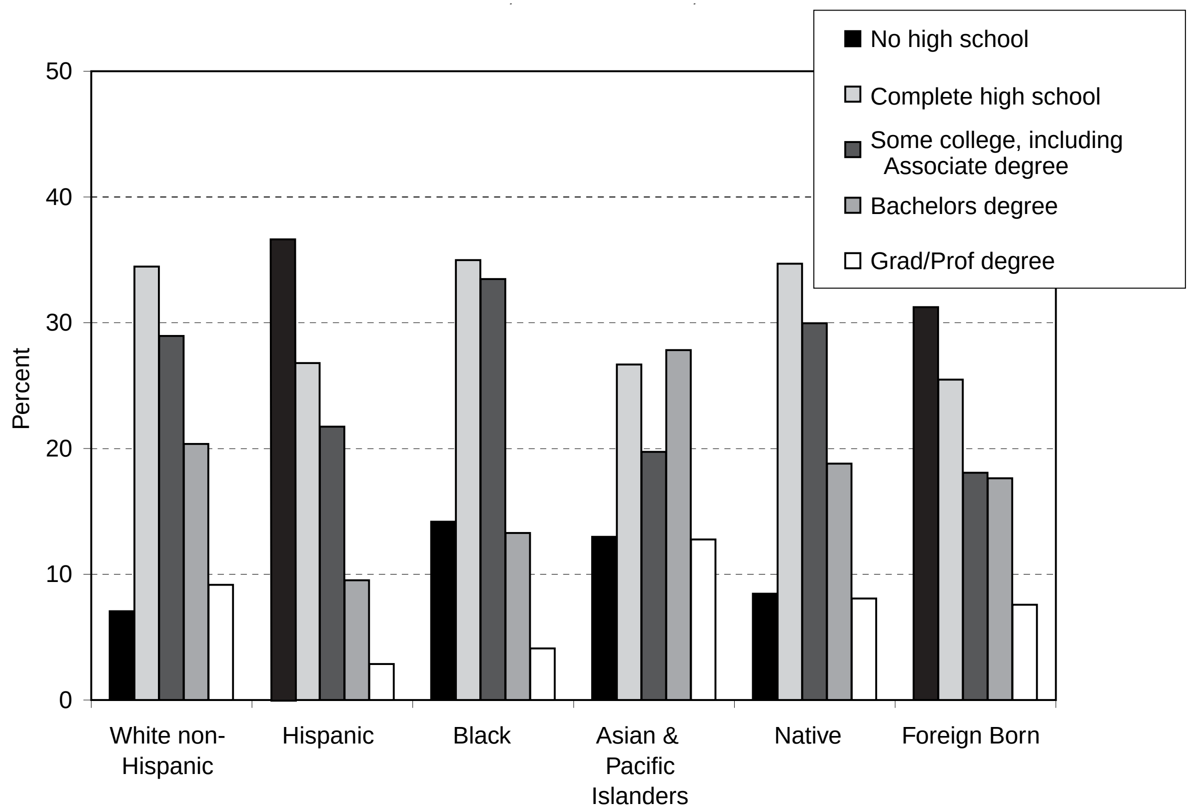


Figure 36. Children ever born, women 40–44 years old of certain race, ethnicity, nativity by income category (<$10,000=100). United States, CPS 1998.

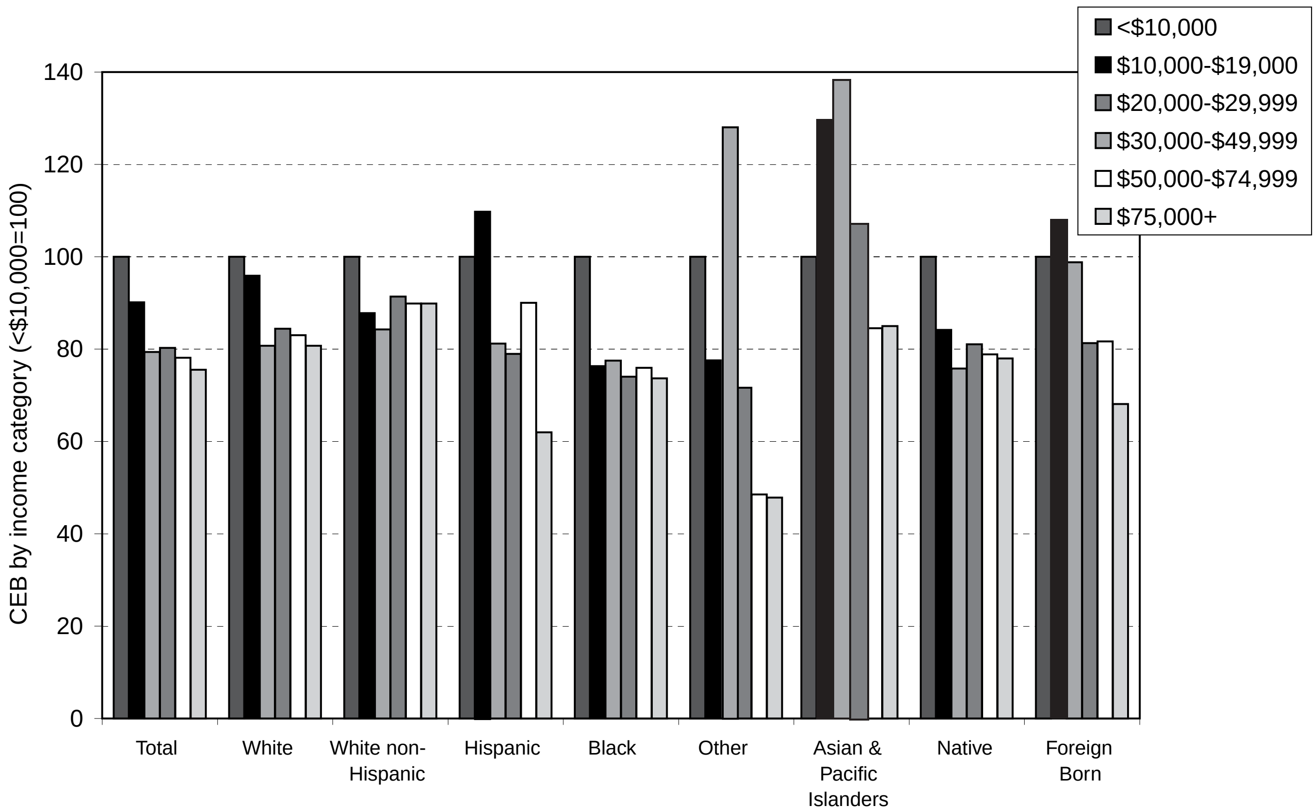


Figure 37. Children ever born, women 40–44 years old of certain race, ethnicity, nativity by educational achievement (No High School=100), United States, CPS 1998

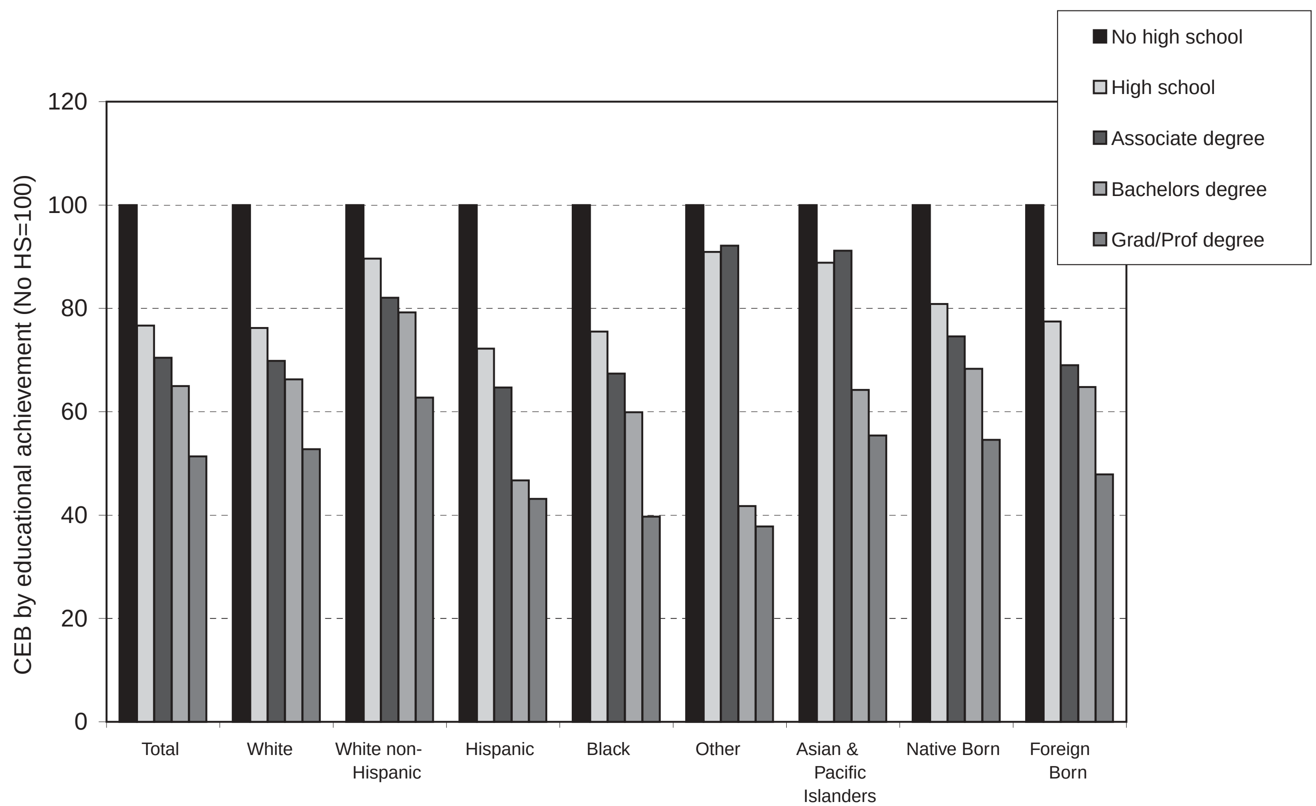


Figure 38. Children ever born, White non-Hispanic women 40–44 years old, by educational attainment and income, United States, CPS 1998

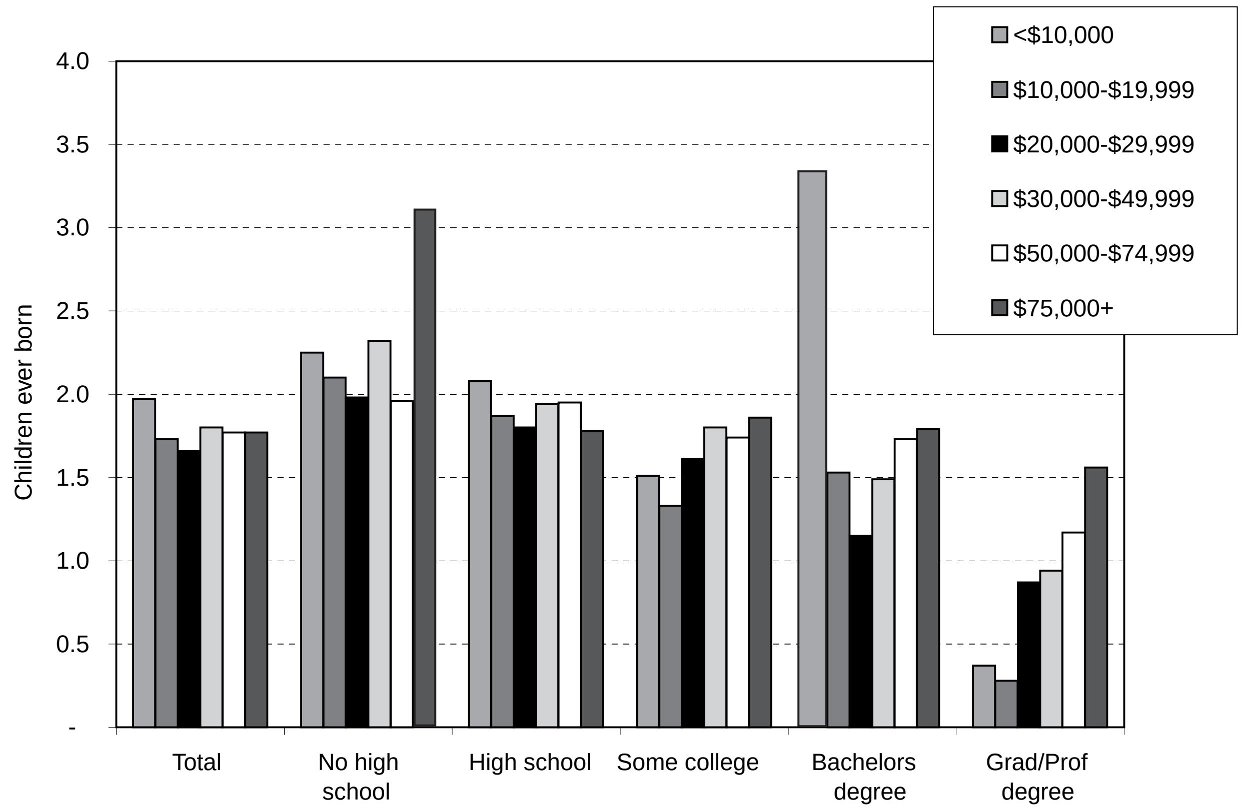


Figure 39. Proportions childless, women 40–44 years old, of certain race, ethnicity, nativity by income category, United States, CPS 1998

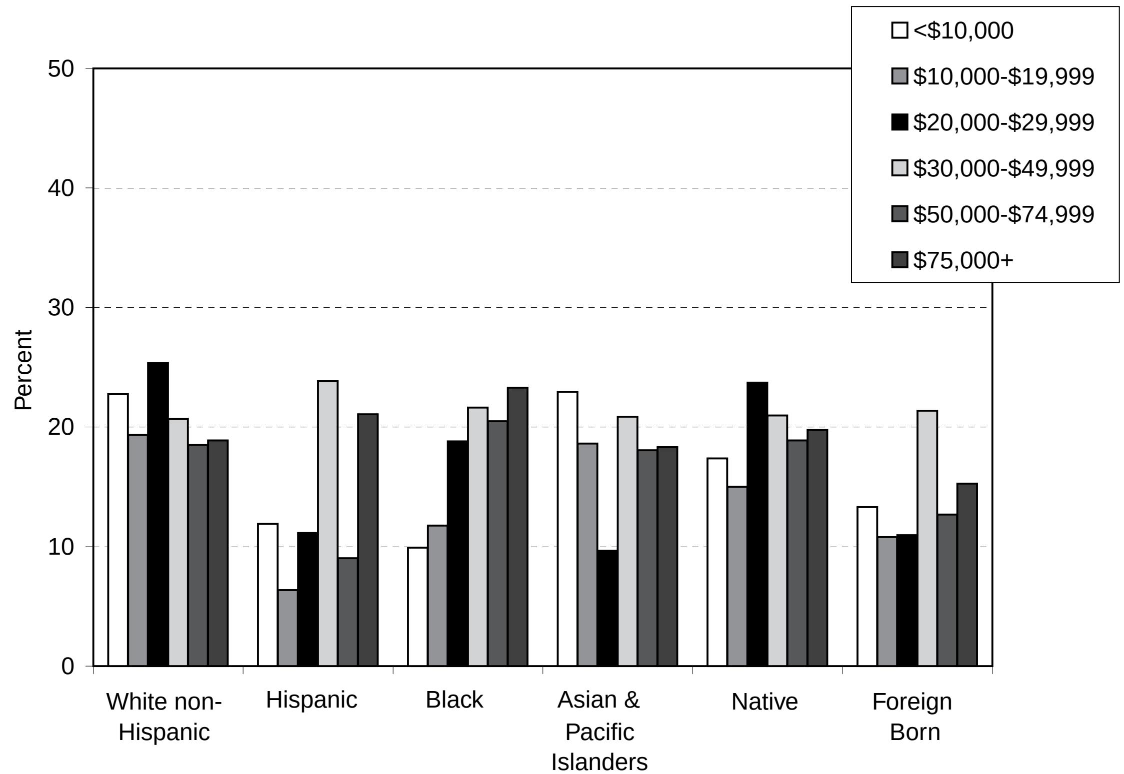


Figure 40. Proportions childless, women 40–44 years old of certain race, ethnicity, nativity by educational attainment, United States, CPS 1998

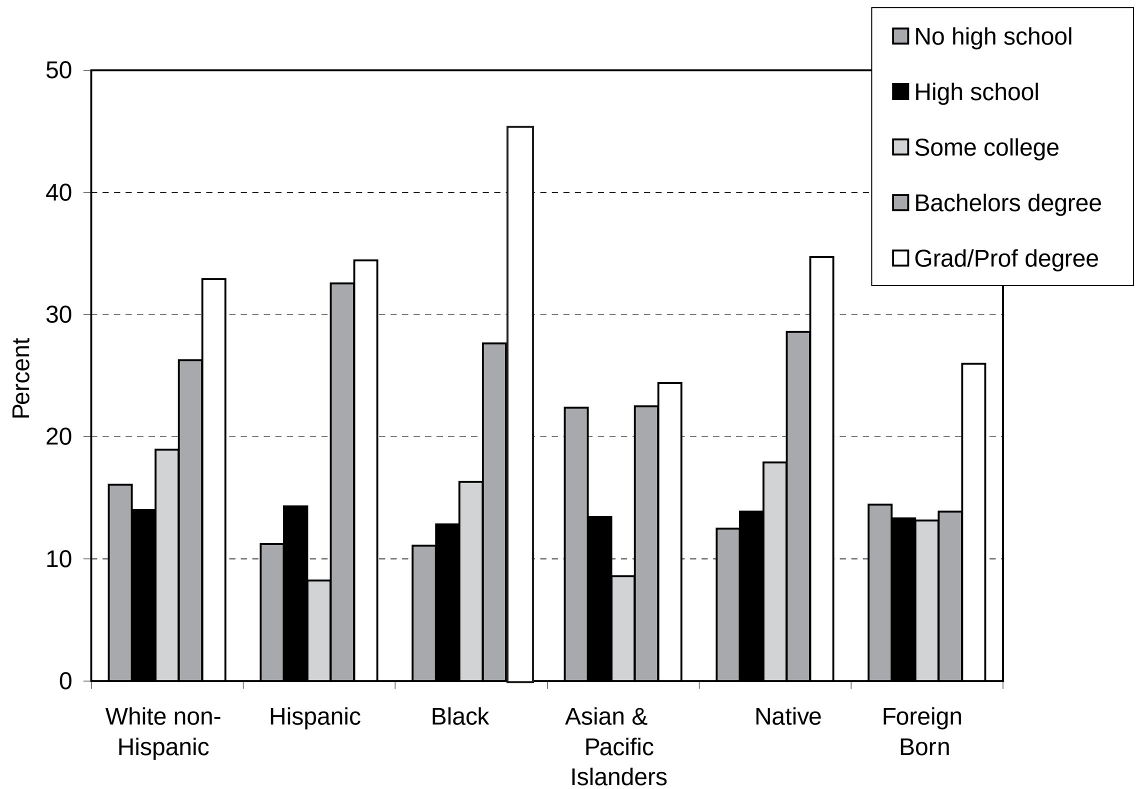
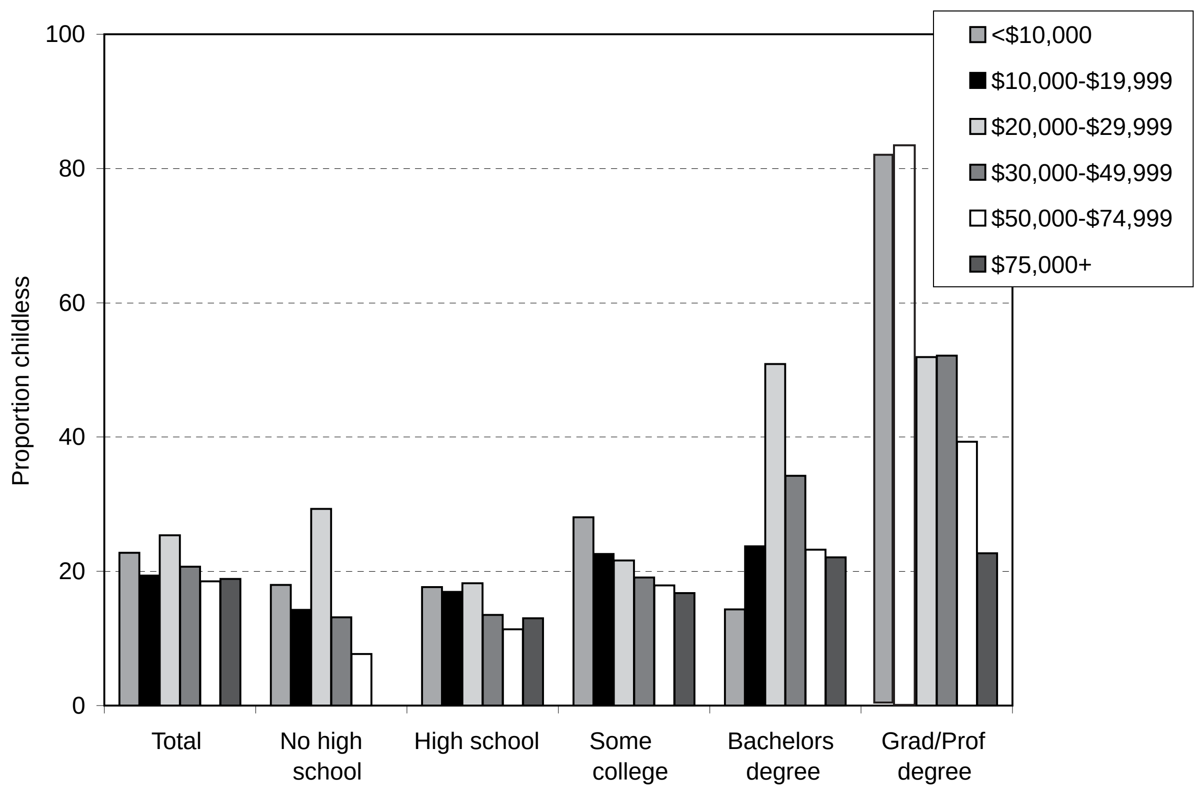


Figure 41. Proportions childless, White non-Hispanic women 40–44 years old, by educational attainment and income, United States, CPS 1998



Discussion

Afternoon Session Welcome
Nancy Gordon, *Associate Director for Demographic Programs*

I am happy to add my welcome to all of you and look forward to the really important interaction among all those present. I used to do projections at the Congressional Budget Office, so I have a lot of experience at developing models and assumptions–-as well as trying to fish information out of data that is never quite what you would wish. I just want to say welcome and thank you for your help.

Robert Schoen, *Discussant*

The Frejka-Kingkade paper, especially the first half of the paper, shows that the authors have amassed a wealth of data that bear on the magnitude of fertility in the United States. There is much to be learned from that data, and reading their extensive documentation (41 figures, 31 tables) was very informative. I should say as well that I tend to agree with their basic prediction, that U.S. fertility will be stable or slightly declining in the years ahead. Nonetheless, I must take issue with them on a number of points, which I will discuss under the headings (1) what is not in the paper that I would have liked to see; and (2) what is in the paper that I am sorry to see. I will conclude with two brief observations on the important question the paper addresses, and the sort of answer we should be seeking.

To begin, what is missing? There are, for better or worse, a number of theories that try to explain contemporary fertility. The work of Judith Blake, Richard Easterlin, Gary Becker, Valerie Oppenheimer, Ron Lesthaeghe, and others have identified processes, relationships, and variables important in determining fertility in developed societies. In the paper, there are passing references to Becker, Blake, and Easterlin, but little effort to bring this extensive body of thought to bear in an organized fashion. Indeed, while the paper talks at considerable length about fertility intentions, it never addresses the question of why people want to have children and what meaning children have to their parents. There is no family context; marriage, cohabitation, and nonmarital fertility are barely touched on. There is no real consideration of the changing—or more accurately—changed role of women. Data related to women's labor force participation are nowhere to be found. The paper, in effect, largely decontextualizes fertility behavior.

The effects of migration are also missing. The paper notes this omission in passing, but offers no explanation for it. Given the high levels of documented and undocumented migration to the United States, and the paper's own data that migrants, especially Hispanic migrants, have high levels of fertility compared to the native born, that omission is extraordinary. The effects of migration are susceptible to demographic analysis, and data are available to perform such analyses, but the paper simply ignores the matter. That issue should not be ignored. Using data for the U.S. available from the National Center for Health Statistics, Stefan Jonsson of Penn State examined the nativity of women giving birth to daughters. In 1978, 90% of the mothers were born in the United States, 3% in Mexico. Twenty years later, in 1998, only 80% of mothers were U.S. born, while 8% were born in Mexico. This is a dramatic change over the very period where the need, according to the paper, is to explain why U.S. fertility did not decline. Now other Western countries have had substantial migration from developing countries with relatively high fertility, so both the relative and absolute effects of migration need to be analyzed in some depth. Still it is reasonable to say that the question is not whether migration had an effect, but how much of an effect did it have.

Now what is in the paper that I was sorry to see? Quite a bit. I do think that fertility intentions can contribute a good deal to an understanding of contemporary fertility, but they need to be used with care. As defined in the National Surveys of Family Growth, a birth is considered unintended if it is either unwanted or *mistimed*. Thus a woman who wants a child but gives birth a year before she planned to do so would be characterized as having an unintended birth. Thus fertility timing, not level, can be crucial. As the paper notes, timing failures, not unwanted children characterize most unintended fertility behavior at young ages, and it is precisely at those ages that the U.S. has many unintended births. As the paper shows, a distinguishing feature of U.S. fertility is that, compared to other Western countries, births occur earlier in the life cycle. That fact, and that fact alone, may account for why the U.S. appears as an outlier with respect to intendedness. The possibility was never considered in the paper, nor were data on *unwanted* births examined.

Second, the paper makes considerable use of data on unintended pregnancies and on their termination by abortion. While not without some value, survey data on pregnancies and their outcomes must be used with great caution because of reporting problems. The underreporting of abortions in U.S. survey data is considered such a problem that Requests For Proposals issued by NICHD, the leading U.S. agency supporting fertility research, generally tell researchers not to propose employing such data because reviewers find them to be unusable for careful analysis. The paper makes no mention of this quality issue.

Third, the paper notes that the U.S. population is among the best educated in the world. However, it appraises the quality of U.S. education very critically. In the absence of any demonstrated effects or quantitative measures, section 4.2.2 of the paper states that lower levels of adult functional literacy in the U.S. "may offer one explanation of why rates of unintended pregnancies and births are higher in the United States than in Western Europe." That view is reiterated in the Summary in section 4.3. There is no discussion of how this educational deficit relates to the paper's finding that fertility is strongly and inversely associated with educational level.

Fourth, the paper gives a double message. Initially, it argues that U.S. fertility is "high" because so much of it is unintended. Numerous inadequacies in the provision of contraceptive services are noted in section 4.1.1 (and elsewhere). Section 4.2.6 states that "Physical abuse of women is widespread in the United States," and cites a reference that such abuse leads to unintended pregnancies. Section 4.2.3 states that "The poor are continuously preoccupied with numerous existential needs, and contraception gets to be neglected." There is virtually a litany of horrors describing how Americans, living in ignorance, fear, and misery, have been forced to endure excessive fertility. Yet in section 4.2.4, the paper advances a series of arguments to the effect that American fertility is "high" because Americans want more children, and that a number of features in American society support childbearing and parenting. Wealthier parents can easily obtain "nannies". The low cost of many goods and services makes childrearing more affordable. The Earned Income Tax Credit subsidizes poor parents. Shift work is available to accommodate two-earner parents. How is the reader to evaluate this? No quantitative estimates of the effect of any of these factors is offered.

The paper does say what it considers most important. Reviewing the evidence in section 4.3, it finds two "mechanisms" of relevance to fertility projections. The first is the health and family planning system, "from the point of view of delivering contraceptives", and the second is the educational system, "from the point of view of raising levels of functional literacy, because that provides part of the explanation for rates of contraceptive failure." Somewhere, somehow, the paper went off track, and an explanation of U.S. fertility became a platform for advocating family planning. As I have argued above, such conclusions are based on misleading and exaggerated claims, not on quantitative analyses, calculations, or comparative data. They only trivialize the social and economic determinants of American fertility.

Let me offer two final observations. First, the size of the fertility differences between the U.S. and the rest of the developed world may be largely explainable by the high fertility of migrants to the U.S., who now bear one-fifth of all American children. That may not be the whole story, particularly after migrant fertility in other developed countries is taken into account, but the question is amenable to straightforward demographic analysis, and such calculations should be made.

Second, we should remember that we are talking about rather small differences between the U.S. and other developed countries. Fluctuations in fertility, here and abroad, are easily that large. Rather than focus on why a country's TFR is 1.6 rather than 1.9, it would be much better to ask why people have *any* children at all. The fundamental issue underlying the projection of fertility is understanding what shapes fertility intentions, what motivates contraception, and what drives fertility behavior.

Josh Goldstein, *Discussant*

His first thought is that the real question here is will U.S. fertility ever be as low as in some European countries?

The Frejka paper makes an important contribution by updating the Heuser data set. He hopes this is made public. It's remarkable how stable fertility rates have been for the last 30 years. It would be better to have some distinction other than White and non-White. What we really need is something with Hispanic or nativity elements.

He also found it interesting that fertility is highest at both ends of socioeconomic spectrum(according to their analysis of a recent Current Population Survey). This actually shows what economists would predict—at a certain point you are so rich that the quantity/quality trade off does not apply.

Goldstein disagrees with the general cohort approach used in the paper. To study the 1990s we have to use period data. He thought this paper was going to be about adjusting current period rates to see just how high American fertility really is [Bongaarts-Feeney adjustments]. It isn't about that, but someone needs to do that paper too.

He believes that U.S.-born women do have European fertility levels. If so, then the question becomes how much immigration will there be and what will be the level of immigrant fertility. He presented the Census Bureau's current projection of Hispanic fertility which shows that Hispanic fertility within the U.S. is higher than our IPC projections for Hispanics in the sending countries. This may be because the immigrants feel wealthier in the U.S. In any event, if we really believe this, we need to document it better.

He doesn't think U.S. fertility will reach European levels. In Europe we have seen huge postponements of fertility, tremendous unemployment of the young, continued rigidity of sex roles, and strong inter-generational links. None of these exist in the U.S. We also have more living space for unanticipated births. We are much more religious. Our fertility will not go so low unless some of those factors arise in U.S. Conversely, it might rise with breakthroughs in reproductive technology or the renewal of family values or the end of delayed childbearing.

Floor Discussion

McDonald

He thinks we should look at geography to see the sources of the fertility—states, urban-rural, etc.

Henshaw

Schoen needs to know that the abortion data are from actual national counts, not surveys. The unintended birth statistics are from the National Survey of Family Growth. These results were added to the abortion data. Henshaw doesn't think there is an undercount of abortions. He agrees with Frejka that contraceptive failure rates are higher in the U.S. Americans want more children than do Europeans, but have them before they intended to—so they are unintended, but not actually unwanted. He also believes the religiosity of Americans is a contributing factor. He suggests that expansion of Medicare coverage to almost half the women giving birth may have raised the number of births. Future welfare reform may actually drop our fertility.

Day

He was happy to see a straight-forward presentation not full of high-powered statistics. Until recently a well-motivated high school student could understand most demography. Economists are threatening to ruin the field. He thinks this was a well done paper and that the comments were marvelous.

Coleman

He agrees with Schoen about the importance of other factors such as female labor force participation. But he thinks that their already intolerably long paper would be completely unmanageable. Perhaps we need a fourth paper to make such comparisons. In the U.S. low fertility may be the result of gender equity, low taxes, availability of cheap childcare labor. The White non-Hispanic TFR in the U.S. is about 1.85—similar to the level in parts of

Norway, France, and Ireland. He suggests that in parts of Northern Europe the same levels of fertility were achieved through different policies such as state action, public pressures, paternity leave, etc.

Bean

Most of the attention in the U.S. is related to the fertility of the foreign-born. He believes we should remember some interesting patterns among the native -born. His own work suggests that native-born Hispanics have higher fertility than immigrants. He therefore agrees with the Bureau's assumptions for Hispanics and not with Goldstein.

Murphy

The U.S. is like the very highest-fertility of the Europeans. Why? We know U.S. is a lot more religious. We also know they start their childbearing a lot earlier and have larger families. The baby boom was bigger in the U.S. Maybe that experience is carrying over and Americans are just used to bigger families. But that has to be fading out rapidly which implies that fertility will decline.

Ahlburg

He disagrees with Day about complexity. He says we need more complexity to truly understand the relationship of fertility and income. Dennis also argues that Frejka's income data are flawed because he did not use the March CPS data (which is properly weighted). The problem may be that this is "transitory" income when what is needed is "permanent" income. He thinks important decisions such as childbearing are related to permanent income. We need to look at a number of CPS's to see how robust these findings are. We also need to extend multi-variate analyses because the question is complicated. He concludes that analyses are needed that a high-school kid can't understand.

Bongaarts

He thinks women in U.S. are achieving their desired family size. Some unwanted balances out some who can't have how many they want. To him the question is: why aren't Europeans achieving their fertility goals? A good part of the difference is due to the rising mean age at childbearing(which he asserts is no longer rising in the U.S.). He suggests the shortfall in Europe is due to unemployment, weak economic growth, etc.

Frejka

He agrees with Coleman that we need another couple of papers. It is next to impossible to include everything that theoretically should be in it. He agrees with Macdonald on the need for geographic analysis. Several people talked about need to disentangle mistimed fertility from unwanted fertility. Research should be done on how to best separate them and get at only the unwanted. The Heuser data are released by the National Center for Health Statistics, but we have assembled it for use. Tomas does not agree with Goldstein about the need for period data to study recent trends. He argues that the cohort patterns and trends do show a lot. He argues that his work on "functional illiteracy" was not based on just one table. He says he made a "comprehensive analysis of functional literacy" which shows that the quality of education in the U.S. is much lower. He claims that many surveys have shown this.

Van de Kaa

The Frejka paper was very informative. But he wonders about the value of comparing Western Europe with the U.S. Fertility is the outcome of a process and a decision sequence. People want to have sex. If they can't cohabit and must have children within marriage–this is like Southern Europe where they have a child, but stop quickly. In Northern Europe, cohabitation is possible so people do not marry and tend to postpone childbearing. The comparison just in terms of actual fertility misses all that complexity. The fertility differences "can only be explained by a different reaction of people who want to achieve the same thing but in different cultural settings and they find different solutions for what they want to achieve." Dirk agreed with Schoen about the need to use theory–he mentioned Ronald Inglehart of the University of Chicago who tries to distinguish between materialism and post-materialism or even "post modernization. The differences in the ways which by values diffuse in countries may explain a lot.

Haub

First, the TFR of White non-Hispanic native-born women is not 1.65 as Goldstein posited. Haub argues that in a very broad-brush way fertility has been related to economic circumstances beginning with the 30s and continuing through the economic changes that took place in the 50s and 70s. We knew we were heading for a recession before September 11th. He argues that we can see the current economic downturn and should anticipate a short-run fertility decline.

Spar

Perhaps women have had to enter the labor-force whether they wanted to or not—because they needed to eat. After all, real household income hasn't gone up in the last 20 years or so. What would American fertility levels be if women had not entered the labor market?

He thinks the issues being discussed today have become more attitudinal and less statistical.

Ventura

NCHS is in the process of developing historical fertility tables for Blacks and for non-Hispanic Whites.

Rao

A decomposition of the American TFR to see how much is contributed by different racial or nativity groups might be informative.

McDonald

We should not overlook the parity heterogeneity in U.S. Quite high proportions have no children or only 1, but many also have 3 and 4 children. He thinks as a consequence it is inadequate to interpret phenomena by saying "Americans do this" or "Americans do that."

Chamie

The basic conclusion of Frejka is that American fertility will be stable or decline slightly. Does anyone disagree?

Goldstein

Bean is more qualified than am I, but it does not seem plausible to me that the Census Bureau's middle forecast for Hispanics is so different from Hispanic fertility in the immigrant-source countries. I can believe that U.S. would be a little higher than source countries, but only a little.

Schoen

I think people's arguments with Frejka are not with the factual parts, but with the theoretical part as well as the significance of the intendedness data. He was struck by Bongaarts comment that if you ask U.S. women how many kids they want—they say 2. And if you look at how many they have—they say 2. Schoen agrees that this is not a coincidence. He thinks this may eventually happen in Europe, citing a paper by Frejka and Calot in Population and Development Review.

Frejka

In the 1998 CPS the native-born White fertility is 1.8 and the foreign-born White fertility is lower.

He agrees with McDonald about the heterogeneity of the U.S. population. He says we need to look at the trends in the race/ethnic differentials and at critical factors like religiosity, family values, gender issues, etc. Also the change in values mentioned by Van de Kaa should be studied, plus geographic variations, what the intendedness data mean and so on.

Goldstein

The Census Bureau is projecting a rise in fertility over the next 50 years and I tend to disagree. It is quite possible, but I would not predict a rise as a result of more immigration. When I see a straight line. I am going to predict its continuation.

Van de Kaa

He predicts a substantial downturn in American fertility. The main difference between the U.S. and other countries is that fertility starts very early. Just as elsewhere, young people in the U.S. will come to see the advantages of remaining independent from children for awhile. This shift in tempo will result in a decline in fertility for awhile. We will see more people not achieve the "2" they want.

Ahlburg

Frejka and Kingkade disagree with their stated conclusions because they wrote early in the paper, "As a matter of fact, there is the possibility of a fertility increase as the cumulative fertility rates up to the known ages for the cohorts of the mid '60s and mid-70s were consistently on a moderately increasing trend." He was surprised that they concluded flat or down trend when what they wrote suggested flat or up. Since he believes in Easterlin, he must conclude that fertility will rise.

George

With regard to the relationship between immigration and fertility, in Canada we found that immigrants from high-fertility countries will initially have high fertility. But when you control for duration of stay, we found that immigrants eventually approximated the native level. He assumes everyone ends up at a 1.5 TFR eventually.

Alcantara

She expects U.S. fertility to remain high, particularly if migration is heavily Hispanic. The Phillippines exemplify a country where the culture is pro-natalist, the religion is pro-natalist, and if you can afford children, you have as many as you can afford. She sees the possibility of a fertility increase, particularly where the economy is good or immigrants are concentrated.

Bean

He attempts to answer Goldstein about the future of the Mexico TFR by saying that 25 years ago he would never have expected it to be as low as it is now. But now there are still areas of Mexico with high TFRs so there is lots of room for further decline. He asks Bongaarts what the tempo-adjusted TFR would be for Europe during the 1990s—he assumes it would be higher than 1.5, but not clear just how high.

Bongaarts

Roughly speaking the European TFRs go up 0.2 to 0.4 when you remove the tempo effect. He thinks the "core TFR" for American Whites is very close to the "core fertility" of Europeans, about 1.8. The differences are quite small in his opinion.

McDonald

The U.S. has relatively lower housing costs than Europe. The tax benefits, the support for education, and the direct financial support for families with children are all a bit higher in the U.S. than we used to think. The indirect costs such as child care are relatively cheap. He doesn't see much impetus toward lower American fertility. There may be issues about the quality of child-care, but that is a different issue. The U.S. seems to be a relatively child-friendly society without signs saying "no dogs or children allowed." The U.S. has a lot of gender equity in the labor force and within the house. What might bring down U.S. fertility is uncertainty. For example, in Europe there is a great deal of uncertainty about unemployment levels. We might check this by age at marriage, but there is no evidence of uncertainty yet. He agrees with Goldstein about future Hispanic fertility which is another reason to expect declining fertility. The one big difference is teenage fertility which is the real source of the cumulative fertility difference with Europe. Will teenage U.S. fertility fall? That is the question.

Coleman

He is struck by the different ways people have of "massaging" the TFR. He says the difference between an implied TFR of 1.6 and 1.9 for a society is important. Coleman says the implied shortfall in Southern Europe is still significant. He says that Hispanic situation in U.S. not so odd—that mexican immigrants to the U.S. actually have higher fertility than do women in Mexico. In England, the TFR of those from South Asia or Turkey is considerably higher than the TFR in the sending countries.

He agrees with Frejka on the importance of good contraception as an explanation of low fertility. French demographers used to use same reasoning to explain relatively high British fertility. It has also been argued by Simon Szretes that Anglo-Saxons have an immature approach to sex which may lead to higher levels of unintended pregnancy in the English-speaking world.

Keilman

The main discriminant in immigrant fertility is not whether or not they are foreign-born—it is their degree of socialization to the receiving country. In Norway those who come after age 18 have much higher fertility than those who arrived at younger ages.

Day

Isn't one of the unanticipated consequences of the introduction of the "pill" that it has become the woman's job not to get pregnant—and teen males take no responsibility?

Frejka

With regard to the Van de Kaa prediction that U.S. fertility may fall. He is not so sure because he thinks there really are many more poor and uneducated in the U.S. than in Europe. Also, he never meant to draw the conclusion Ahlburg made—that was just mentioned as a possibility.

O'Connell

He worked on the first CPS survey to ask unmarried women about babies. Most of his research work then was oriented to justifying the value of the question. Later, when expectations were very stable, people said they were meaningless and just stuck on 2 kids. Now the data have been reduced to only intermittent collection.

Tuesday Afternoon

Session III

Should Fertility Intentions Inform Fertility Forecasts?

S. Philip Morgan
(Duke University)

Should Fertility Intentions Inform Fertility Forecasts?

S. Philip Morgan
(Duke University)

Abstract

If women's fertility intentions at age 20 or so were highly predictive of their subsequent behavior, then their completed cohort fertility could be projected from these stated intentions. Theoretical and empirical bases for skepticism exist, but some recent data suggest impressive correspondence between intentions and subsequent behavior. Based on the theoretical concerns and on empirical regularities, we propose an analytic model that adjusts estimates of aggregate intent. The model's parameters reflect key assumptions underlying the projections and thus can indicate which assumptions are least plausible and which ones are primarily responsible for projection errors. However even given this innovation, the role for fertility intentions in fertility forecasting is modest—even if the proposed analytic model performs well, its suggested role in forecasting fertility is secondary to period estimates of quantum [as defined and estimated by (Bongaarts and Feeney 1998)] and is dependent upon accurate assessment of whether the future will be substantially different than the present.

Introduction

Can fertility intentions/expectations (hereafter intentions) stated at a young age be used to predict completed fertility? This question has received substantial attention in the demographic literature (See Table 1). We examine this literature and conclude that intentions may be of some use for population forecasting. We discuss the reasons for our guarded conclusion. Consistent with this guarded assessment, we suggest a procedure that incorporates some corrections for the limitations of intention data.

We use a simple equation to organize the following discussion. Specifically, aggregate completed fertility (at age x, F_x) can be decomposed into three parts as shown in the equation below.

$$F_x = F_{x-y} + \left(I_{x-y} \bullet \frac{{}_yB_{x-y}}{I_{x-y}}\right) \tag{1}$$

Where F_{x-y} is mean fertility y years earlier,

I_{x-y} = mean fertility intentions y years earlier,

${}_yB_{x-y}$ = mean behavior (or births added) over the y year period $x-y$ to x.

Current parity and intentions at age $x-y$ can be observed/measured directly in many surveys (including the CPS). The difficult task is predicting fertility y years into the future, a task equivalent to estimating the final term of equation (1): *the proportion of fertility intended at age x-y that will be realized at age x* (i.e., the *predictive validity of fertility intentions*). While the formula above is general and can be applied when y is small, the potential value of intentions data (at least for forecasting purposes) increases if $y>10$, and especially when x = age 40 and $y>15$ (i.e., when intentions in young adulthood, age <25, can be used to predict completed fertility, age 40+).

Terminology

Before discussing the theoretical and empirical difficulties of estimating completed fertility from intentions data, we pause to justify our use of the term *fertility intentions*. Fertility intentions refer to questions that ask how many additional children women *intend/expect*. Demographers (like others) acknowledge the conceptual distinction between intentions and expectations. Intentions refer to planned actions toward a particular goal or "a determination to act in a certain way"; expectations incorporate an assessment of what might happen independent of one's intention and acknowledge factors beyond one's control (e.g., "to anticipate or look forward to the coming or occurrence").[1] In the context of expected fertility, respondents are invited to consider possible contraceptive failures and/or the inability to have all the children that they might intend (due to sub-

[1] Webster's New Collegiate Dictionary. 1981. G. & C. Merriam Co.

fecundity, for instance). Despite an explicit prefatory statement to the expectation question (see Ryder and Westoff 1965:20), respondents in the 1965 NFS gave answers to intention and expectation questions that were "virtually indistinguishable" (Ryder and Westoff 1965). A likely explanation is that respondents are unable to anticipate their future contraceptive failures or subfecundity and thus cannot incorporate them into their response. Thus in practice, stated intentions are nearly identical to stated expectations. But we stress that the concept being measured is almost certainly the respondent's stated *fertility intentions* regardless of the term used in the questionnaire item.[2]

The Fertility Intentions Concept and Its Measurement

At their core, fertility intentions are linked to a *fixed target* model, e.g., individuals or couples "formulate a desired completed family size and pursue this relative constant target throughout their reproductive life (Lee 1980:205). Such a model was central to early versions of the "new home economics" e.g., (Becker 1960). Given such a *fixed target* model, the standard survey-research operationalization has face validity. Specifically, respondents are asked if they intend another child. If yes, they are queried as to "how many more?" Demographers have raised serious concerns about the concept/operationalization of such a "fixed target" model.

Perhaps the most important substantive criticism challenges the one-time nature of the *fixed target*. Instead, strong substantive arguments suggest a series of sequential decisions. Children are generally born one at a time not in lots (Namboodiri 1972). This reproductive structure imposes a set of birth intervals. In turn, these intervals allow for a set of sequential decisions or at least for reassessments of earlier decisions. An extreme statement of this view is Ryder's statement that the relevant fertility decision is "whether to let the next ovulation come to fruition"—suggesting a sequence of month-by-month decisions. Key to this sequential perspective are the claims: i) that experience with prior births may affect the decision to have subsequent births and ii) that some births may be normative (first and second) and others less so (third births). This distinction (between fixed target and sequential models) is important because it challenges whether a response of 2 vs. 3 additional children has any behavioral consequence. One could argue that since intending 1, 2 or 3 more children has the same parity-specific intent (to have another child), no other information valuable for prediction is contained. Thus, for instance, if spouses disagree on the number of additional children intended this might spark discussion or conflict, much like different political views might, but this difference requires no immediate resolution because the parity specific intent—the decision with current behavioral implications (i.e., do we have another child)—does not vary between spouses (Morgan 1985:131).

Another way to express this concern is to argue that the number of additional births intended should not be treated as an interval level variable. In the standard procedure for calculating intended parity one adds current parity (F_{x-y}) to additional intended births (I_{x-y}). This calculation weighs an additional intended births of "3" (from woman z) equally with intentions from three women (v, w, x) who intend one additional child. If parity-specific intentions are more meaningful, then they should be "weighted/anticipated" more heavily/strongly.

A second major concern is related to the above. Specifically, the standard concept/operationalization of fertility intentions has no time referent. Consider, a pair of twenty-five-year old women. Both intend a first child. But one may intend a child soon (in the next year or so) and the other may intend to postpone the first birth for at least five years. Rindfuss and colleagues (Rindfuss, Morgan et al. 1988: Chapter 9) show that such women behave very differently in the short run (i.e., in a subsequent 3 year period). One suspects that the intent for a child in the more distant future allows for more intervening events that might alter one's intention. Alternatively, delay might be a highly rational plan for women with high educational aspirations. Thus again, the decision more clearly linked to fertility may be one that is *parity and time dependent*.

A third weakness of the fertility intent question is the absence of contextual referents. We assume that the unstated referent is "if things stay the same or if they work out as I expect."[3] There are a large number of idiosyncratic events that could lead to revised intentions. A marital disruption, for instance, may change a woman's economic situation and family environment such that intentions for future births are revised down-

[2] Ideal family size and desired fertility are related concepts. Ideal family size usually alters the referent for the question to "an average American family" or to a "family like yours" and removes it from the context of what the respondent is likely to do. Desired family size refers to the number a woman would like to have. Empirically, answers to this question are similar to answers to the intended/expected questions. The desired question might be expected to incorporate children desired, but not intended, because of constraints (i.e., subfecundity or shortage of income). It would not include an estimate of "unwanted fertility."

[3] It is commonly assumed that respondents make the invalid assumption that the future will resemble the past, as Westoff and Ryder (1977) state, "the same kind of forecasting error that demographers have often made." Also see Hendershot and Placek (1980).

ward or become uncertain. Alternatively, a woman not intending children might develop a relationship that makes childbearing more attractive, leading to the intent to have a child.[4] In the aggregate, such idiosyncratic events might cancel. Thus, aggregate intentions would approximate aggregate fertility. But much evidence suggests that cumulative cohort fertility is influenced by period factors that alter the timing of childbearing and the number of additional children couples intend (Lee 1980). In addition, decision-makers might anticipate other factors but standard intention questions do not incorporate them. For instance, suppose one has a son but desires a daughter. Thus, a report of one more child intended assumes that the next child will be a son. If not, then an additional child may be desired. Coombs' [see (Coombs 1974); (Coombs 1979a); (Coombs 1979b); (Coombs and Fernandez 1978)] work is illustrative of attempts to acknowledge that intentions reflect preferences for the number *and* sex composition of children not well represented by a single numerical response.[5] Another example of the import of contextual information is found in Barber's recent work (Barber 2001). She shows that attitudes toward alternatives that compete with parenthood and normative support for parenthood (i.e., the normative support for marital childbearing) alter the predictive validity of intentions for a first birth.

Fourth, the standard intention question does not reflect possible conflict or disagreement among decision-makers.[6] Should the intention question ask about the woman's intent or the couple's intent? Does the female partner know and incorporate the spouse's intent in her own report? Which spouse's report has greater predictive validity? Substantial evidence indicates that predictive power is enhanced by including the intent and characteristics of both partners (see Table 2). But supporting the argument that a *couple* response can be obtained from *one* respondent, Morgan (Morgan 1985) presents evidence that the wives' and husbands' reports of the couple's intent reflects their combined preferences.[7]

While clearly related to some of the above, ignoring the level of certainty constitutes a final weakness in past/current conceptualizations/measurement of fertility intentions (I_{x-y}) [see (Morgan 1981) (Morgan 1982)]. Several studies have shown that adding a dimension of certainty improves dramatically the reliability [(Thomson and Brandreth 1995) (Wu and Wang 1998)] and predictive validity of fertility intentions [(Rindfuss, Morgan et al. 1988: Chapter 8); (Thomson, McDonald et al. 1990); (Remez 2000)]. We suspect this is because certainty acts as an intervening (or proximate) variable mediating many concerns, like the ones mentioned above. In addition, uncertainty may reflect acceptability of a range of family sizes. If a woman reports her intent as "two more children" and reports her second and third family size preferences as "three" and "four" respectively, then one can interpret this as greater certainty that she will have at least two children. In contrast, others may give "two" as a first preference but this number is the maximum number they would consider. Second and third choices might be "one" and "none."[8] Scoring schemes have been suggested that identify and operationalize such ranges.

This set of substantive concerns challenges the usefulness of fertility intentions of young women for predicting their cumulative fertility. Sequential models suggest that intentions might be better characterized as a "moving target" heavily influenced by period and life-course factors (Lee 1980).

The Link of Intentions to Behavior

Even if intentions were formed early and did not change, they might not be strongly associated with behavior. We posit two illustrative models in Figures 1 and 2. Figure 1 shows a dominant influence of intentions on behavior. Intentions link social, economic and psychological variables to fertility. Note that factors affect intentions directly. In turn, intentions mediate these more distal effects on fertility. Under what circumstances will such a strong model describe the links between intentions and behavior? Ajzen and Fishbein [(Ajzen and Fishbein 1977):889] argue that the link between attitude/intention and behavior will be greatest when there is congruence on four aspects of attitude/intention and behavior: 1) when the relevant *action* is unambiguous and when the 2) *target*, 3) *context* and

[4] See Schoen et.al. (1999) for an excellent empirical study of longitudinal data showing the import of marital status for predictive validity.

[5] Pollard and Morgan (2001) present evidence that the sex of previous children has had a modest but consistent effect on U.S. fertility for 50 years. However, this effect seems to have disappeared in the last decade of the 20th century. Thus, this factor is likely of little importance for forecasts in the 21st century.

[6] We discuss only the influence of partners here but evidence of others' influence exists (e.g., Barber, J. S. and W. G. Axin 1998; Barber 2000; Barber 2001).

[7] Morgan (1985) shows that husbands and wives factor in their spouses' preferences. But they weigh own preferences more than their spouses'. These results pertain only to white, married couples.

[8] Blake, J. (1974) argues that concern about population growth led to propaganda induced decline in intentions within the range of acceptable choices. In effect, she argued that the 1960s and '70s intentions data contained a period-specific reporting bias.

4) *time* of action are specified. In the case of fertility, the *action* and *target* are apparent although not explicitly stated. One has unprotected intercourse (the *action*) in an attempt to become pregnant (the *target*). Or, one avoids intercourse or uses appropriate birth control methods to avoid having a child. But there are sources of considerable inconsistency here, because the action does not guarantee that the target will be achieved. In fact, the target of having a birth requires the successful completion of a series of events—having intercourse, not using contraception, conceiving, and not having a spontaneous or induced abortion. Realizing intent for multiple children requires a much longer series of actions. Inconsistency between intent and outcome is more likely when the outcome rests on a series of behaviors/decisions that span considerable amounts of time (Davidson and Jaccard 1979:1365).

Perhaps more problematic for the link between fertility intent and behavior is the absence of an explicit contextual referent (*context*). We discussed this problem above at some length. Moreover, also noted above, the standard intent question does not have a *time* referent (respondents are asked to estimate births over the next two decades). Note that the long time frame for realizing fertility intentions allows substantial time for the socioeconomic decision context to change.

If the strong model in Figure 1 holds, then these issues are of little practical concern. Fortunately, a substantial body of research exists on which to assess this model. This empirical research indicates a clear and strong link of intentions to behavior. A large set of studies uniformly show that intentions are related to actual fertility (see Table 1). But these same studies clearly indicates a more complicated process, like that shown in Figure 2.[9] Specifically, some subgroups have higher fertility than others net of intentions, i.e., a *direct effect* indicated by solid lines that bypass the proximate variable, fertility intentions [e.g., see (Schoen, Astone et al. 1999); (Rindfuss, Morgan et al. 1998); also see Table 3]. Such direct effects might reflect sub-group pronatalist pressures unanticipated by young women that increase the likelihood of births regardless of stated intention. Also some groups are better able to predict their future behavior, i.e., an *interaction* of intent with covariates, indicated in the figure by multiple dashed lines between intent and fertility [e.g., see (Wilson and Bumpass 1973); (Williams, Abma et al. 1999); (Schoen, Astone et al. 1999) also see Table 3]. By definition, group differences in unwanted fertility (frequently due to contraceptive failure) imply differences in predictive validity. We conclude that inconsistency between intent and behavior is patterned and thus should not be ignored. Evidence is strong for the more complicated models linking intent and behavior—models like model 2 (Morgan 1982).

In empirical importance, period factors stand out as factors that have direct effects on fertility. NiBhrolchain (NiBhrolchain 1992) argues that period effects on fertility dominate cohort ones in developed countries. From this period perspective, cumulated aggregate fertility is seen as the summed consequence of period factors. Cohort factors (including intentions at an early age) are seen as a relatively unimportant influence on the age pattern or quantity of fertility. This perspective is captured in Bongaarts and Feeney's (Bongaarts and Feeney 1998) recent adjustment of the total fertility rate, an important innovation to which we return in a subsequent section.

Age effects/interactions may result because of the flexible timing norms surrounding fertility in the U.S. Fertility postponement is an accepted and long-standing strategy to life-course exigencies (see (Morgan 1991) (Rindfuss, Morgan et al. 1988)). But there are limits to this flexibility that are imposed by the "biological clock" (e.g., real and/or perceived changes in fecundity) and by social norms regarding ages that are "too old" for motherhood (Rindfuss and Bumpass 1978). Since exigencies may be represented for the aggregate as period factors, there may be stronger age effects in some periods than in others. Rindfuss and colleagues (Rindfuss, Morgan et al. 1988) report that at young ages (i.e., 22–25) and in the period 1976–79, intended first births were 18 times more likely to be intended but not realized than not intended but born (i.e., births were much more likely to be postponed than advanced). They argue that this finding was almost certainly accentuated by strong antinatalist period factors of the late 1970s. In a similar vein, Namboodiri (1981) argues that the 20–25 year age group is hypersensitive to period effects compared to other age groups. This hypersensitivity would change the predictive validity of intentions differentially for different age groups.

Marital status plays a major role in a number of models linking fertility intentions to behavior (e.g., (Barber 2001); (O'Connell and Rogers 1983) (Schoen, Astone et al. 1999). In a number of empirical studies, the intentions of married women are shown to be more accurate than those for unmarried women. The presumed process posits that the intentions of unmarried women are more conditional and less certain. This could result because of the widespread belief that

[9] See Barber's (2001) figure 1 for an even more complicated, but more realistic model. In addition, Morgan has argued that behavior may change due to period factors and that intentions are subsequently revised downward.

women should be married before having children. A related normative belief is that parents should jointly decide if and when to have children. Women not married at time x-y cannot be assured that they will marry. If they do marry, they cannot be assured that their partner will desire children when and if they do. Thus, nonmarriage provides a proxy for some of the uncertainty in intentions discussed above (especially in terms of a time referent and aspects of a contextual referent).

Race/ethnicity differences in predictive validity are of special interest given the Census Bureau's task of forecasting population change for race/ethnic subgroups. The literature consistently shows that unintended pregnancies and unintended births are more common for African Americans and Hispanics (e.g., Trussell and Vaughan 1989; Henshaw 1998). This suggests that aggregate intent for these groups would underestimate their completed fertility. However, there is some evidence that subfecundity might be more common, at least for African Americans (Mosher 1982). Obviously, these differences are potentially offsetting. Also, the literature shows consistently that race/ethnic differences are sharply attenuated if characteristics such as marital status and poverty status are controlled. These results suggest that such factors mediate or explain the effects of race, e.g., African Americans are more likely to have unplanned pregnancies and unwanted births because they are less likely to be married and more likely to be poor.

The literature cited above suggests a strong but a variable link between fertility intentions and behavior. Table 3 summarizes some of this evidence. The implication of these findings is that any attempt at estimating intentions needs to disaggregate the population and estimate predictive validity separately for subpopulations. Such a strategy would allow one to separate aggregate change due to changes in predictive validity from those caused by changing population composition.

Previous Attempts to Predict Fertility from Intentions Data

Three general strategies have been used to predict completed fertility from intentions data. The first two adopt a "fixed target" model (e.g., couples "formulate a desired completed family size and pursue this relative constant target throughout their reproductive life"). The third strategy allows for a "moving target" as cohorts adjust to contemporary period factors [(Lee 1980):205].

The first strategy simply projects the sum of current fertility and stated intentions. If using *married women's intentions*, this would translate into something like the following: if intentions remain constant and "barring infecundity, extra-marital births and contraceptive failure, and assuming all women marry sufficiently young to achieve their (intended) family size" then intentions would equal behavior [(Lee 1980):209]. No one recommends this approach for population forecasting because such strong assumptions cannot be defended. But this model has been the basis for many descriptive studies of the correspondence between intentions and behavior. The classic summary of this literature is the pessimistic statement of Westoff and Ryder (Westoff and Ryder 1977), i.e., fertility intentions suffer the same flaws as other measures of period fertility (also see Long and Wetrogan 1981). Intention data did not anticipate the sharp downturn in fertility in the 1970s. For this reason, a recent NAS review (Bongaarts and Bulatao 2000) did not suggest a major role for intentions in forecasts of fertility.

A second strategy maintains the "fixed target" perspective and attempts to "correct" for intentions that tend to overstate/understate actual fertility. Specifically, the targets remain fixed but obstacles to meeting them result in achieved fertility substantially below (or above) intended. If proxy variables for these obstacles can be identified, then they can be used to adjust the degree of predictive validity. For instance in settings with strong norms against nonmarital childbearing, one can view unmarried women as having an additional obstacle to childbearing, i.e., getting married. Likewise, younger persons' intentions may prove overly optimistic because they do not anticipate the difficulty of combining children with labor force participation. Work experience accumulated at later ages makes obstacles more obvious. Finally, the demographic adage, in the aggregate fertility delayed is fertility foregone, is partly based on the inevitable increase in subfecundity and infecundity with age [see (Bongaarts and Potter 1983): Chapter 7]. Thus, rationales exist for expecting that some groups' fertility will not meet expectations: younger women, older women, unmarried women and working women. Assuming that these proxy variables will provide proper weights for deflating fertility intentions in the future, one can use them to "correct" intention data. In addition, one could anticipate the effects of likely changing population composition in the future. For instance, declining proportions married would suggest greater deflation of stated aggregate intentions, all else equal.

The "moving target" model acknowledges that individual and aggregate intentions can change over the reproductive life of the cohort. At its extreme, relaxing the "fixed target" assumption implies that fertility intentions are of no use for prediction. However, if the pattern of change in intentions over time can be predicted with great confidence then adjustment might be straightforward. If, for instance, we could build a model predicting fertility intentions and could project level of intentions into the future, then these projected intentions could be used to predict fertility in a second stage projection. Lee (Lee 1980) suggests such a strategy in his influential article. But neither the empirical or theoretical basis for such a model exists nor is it foreseen. de Beer (de Beer 1991) suggests a theoretically related but highly constrained version of the "moving target" model. His strategy allows intentions to change with age as in a previous cohort. While this provides a technical solution, it is not appealing substantively since the behavioral theory underlying it is not specified. Insufficient evidence is accumulated to suggest that such cohort parameters are structural. If they are not structural but change with period, then parameters from a previous period would be applied incorrectly to a current one. Such a model implies unidirectional change and thus is especially problematic for long-term forecasts.[10]

Recent Evidence of Predictive Validity

Given the theoretical/empirical issues above, we find it surprising that there is substantial evidence that mean intended parity is *relatively* stable and *frequently* provides good/useful estimates of mean completed parity. Of course, aggregate consistency between intentions and fertility has been noted for the pre-1970 period (see Westoff 1981) But the striking correspondence we document below is for recent cohorts. This new empirical evidence is important and we review it in some detail. Perhaps the predictive validity of reproductive intentions is greater now than in earlier periods. One reason could be that women intend fewer additional births than in the past and may be more certain about this smaller number. In addition, Barber (Barber 2001) suggests that American society is more accommodating of personal preference (and less constrained by social norms) and that such change may increase the importance of women's individual attitudes and intentions, including fertility preferences.

While no one recommends taking intentions at face value (strategy 1 above), recent data suggests that errors in doing so might be small *for contemporary U.S. cohorts*. For example, in Table 4 we show published data from selected CPS surveys conducted since 1976.[11] Note that the structure of these data does not allow great precision in identifying birth cohorts.[12] Nevertheless, expected parity (current parity + additional expected births) is slightly above two for all cohorts at each age. In Figure 3 we show current and intended parity for two cohorts (c1953-57 and c63-67). With increasing age (from 18–24 to 30–34) observed parity rises toward expected parity. The gap remaining for those 30–34 will close almost entirely if these cohorts bear children over the next decade at rates comparable to those of contemporary women. Thus, the predictive validity of aggregate CPS intentions seems high.

A second example comes from calculations based on the 1979 National Longitudinal Survey of Youth and shown in Figure 4. This survey has multiple panels covering much of the 1957-65 cohorts' reproductive experience. We show here data only for the youngest (c1965) and oldest (c1957) cohort. We show data separately by race. While young white women (less than 23) seem to overstate intentions, by the early twenties aggregate intentions approximate 2.2 and 2.0 children per woman for the two cohorts and show great stability. Fertility is nearly complete for the c1957 and intentions approximate current parity. Note that current parity at age 41 (i.e., completed fertility) could have been accurately predicted from intention data over 15 years earlier. Again one is struck by the predictive validity of aggregate intentions. The later cohort (i.e., c1965) shows a pattern of declining intentions with age, but this tendency is slight and this cohort has not completed its fertility. Results for blacks are slightly different. If there is a tendency for aggregate intentions to change it is upward, as opposed to the downward drift suggested by the data for whites.

Some published work uses a version of the second approach, i.e., adopts the fixed target model but deflates group intentions of certain groups based on proxies for their likelihood of achieving intentions. Using a set of CPS surveys spanning the 1971–81 period, O'Connell and Rogers (O'Connell and Rogers 1983) report that married women's birth intentions were realized. They link the failure of intentions to predict the 1970s fertility decline to the increase in unmarried women and a shift toward later ages at childbearing. Thus, for married women, aggregate intentions were realized. Since (in this period at least) unmarried women have intentions similar to married women (O'Connell and Moore 1977), accurate prediction for the full cohort would require significant deflation of the intentions of unmarried women. Van de Giessen (Van de Giessen 1992) is the best illustration of this second approach. He estimates predictive validity for population subgroups using data from the fertility surveys in the Netherlands (see his Table 12.6). For

[10] An attractive feature of the model is that the data requirements are minimal. A repeated cross-sectional survey provides the necessary data to implement the de Beer "partial adjustment" approach.

[11] Table posted at: *http://www.census.gov/population/socdemo/fertility/tabH8.txt*

[12] We have matched cross-sectional age groups to approximate a birth cohort. Most problematic is the earliest age group is 18–24, not 20–24.

instance, young women without partners are expected to realize only .90 of their stated intentions, those living with husbands are expected to realize 1.0 of their stated intentions; unpartnered older (i.e., age 33–37) women are expected to realize only .50 of intended fertility. He shows such adjustment factors for roughly 50 (not mutually exclusive) subgroups. A research program could be instituted to estimate, predict and understand variation in predictive validity. Given reliable/valid estimates of predictive validity one can construct reliable/valid aggregate estimates of completed cohort fertility.

A Proposal for Adjusting Fertility Intentions for Projection Purposes

Our proposed strategy returns to the equation at the outset and uses basic demographic decomposition and standardization procedures. Specifically, equation (1) can be rewritten as a weighted sum of subgroup means.

$$F_{xp} = \sum_{c=1}^{\max}(F_{x-y,p-y,c} + ((I_{x-y,p-y,c} \cdot (\frac{{}_{y}B_{x-y,p-y,c}}{I_{x-y,p-y,c}}))) \cdot (\frac{n_c}{N_{xp}}) \qquad (2)$$

The population is composed of *C* groups (*c*=1 to *C*) and *n* is the number of women in subgroup *c* and N_{xp} the total population of women age *x* in year *p*. Each group *c* has an average parity (F) at age x–y in period p–y. Each group also has a mean number of additional children intended (I), again at age x–y in period p–y. The proportion of intended fertility realized by age x can vary across subgroups (*C*) and across periods or cohorts (*P*).

The subgroups can be defined in many ways, e.g., race/ethnicity and/or marital status subgroups.[13] Changing distributions can be observed, modeled or projected. Variation in *F* and *I* can be observed from data or can be estimated from models.[14]

Our proposed strategy has a number of advantages.

1) The proposed strategy is simple and transparent. It is a straightforward application of fundamental demographic concepts of decomposition and standardization.

2) As an identity, the model holds for any and all subgroups. Moreover, aggregate estimates of intended parity can be exactly represented as the weighted sum of subgroup means. Thus, total and subgroup estimates can be entirely consistent. This feature is important given the Census Bureau's interest in projecting both the total population and race/ethnic subgroups.

3) The key parameter, the proportion of intended fertility realized, is a straightforward measure of predictive validity. A substantial literature on this topic exists in demography vis-a-vis fertility and in the social sciences more generally.

4) The proposed strategy builds directly on the work of others. Our proposal is entirely consistent with suggestions of O'Connell and Rogers (O'Connell and Rogers 1983) for U.S. data and is very similar to the strategy (i.e., *method of limiting factors*) proposed by van de Giessen (Van de Giessen 1992): 232) and applied to data for the Netherlands.

5) The data requirements for implementation and for assessment of prediction are minimal. **Specifically, if the CPS continues to collect data on intentions and current parity and if the series of reproductive histories were reinstated, then a powerful tool for prediction and for assessment would be in place.** We develop this argument in Appendix II. But in short, the June CPS fertility intention supplement asks the relevant questions for measurement of F_{x-y} and I_{x-y}. Subsequent CPS surveys can provide an estimate of

$$({}_{y}B_{x-y} / I_{x-y})$$

for the entire population. But in order to estimate predictive validity for subgroups defined by parity or marital status at time *p–y* then retrospective marriage and fertility histories will be necessary. Specifically, CPS retrospective data would allow one to identify (at time x) a subpopulation that was a given marital status and parity at time *x–y*.

[13] Complications arise if the denominator in equation 2, intended additional births (*I*), equals 0.0. This could happen if the population of women were disaggregated by their intent for more children at time *x-y* (i.e., one subgroup could be women intending no additional children; thus, *I* equals *0.0*). Such a circumstance can be handled easily and need not be discussed at length here. In brief, equation 2 can be applied to the portion of the population intending children and equation 3 below to the subset intending no more children. The intended parity for the full population would be the weighted mean of these two estimates.

$$F_{xp} = \sum_{c=1}^{\max}(F_{x-y,p-y,c} + ((1.0 \cdot (\frac{{}_{y}B_{x-y,p-y,c}}{1.0}))) \cdot (\frac{n_c}{N_{xp}}) \qquad (3)$$

[14] One promising modeling strategy is the joint estimation of F and I in a bivariate probit model (as suggested by Calhoun, C. A. and J. de Beer (1991). In this model, current parity is estimated in a selection equation and intention is the conditional choice variable. Predicted fertility intentions (I_{x-y}) can then be modeled as a function of characteristics at x-y or those anticipated at *x+t* (where t is some forecast period into the future).

(See Appendix II for examples of such estimation with fixed covariates).

A number of important details need to be resolved as part of an agenda to estimate, understand and project predictive validity. We discuss these issues in Appendix II so as not to postpone our general argument and obscure our general strategy.

The Use of Intended Parity in Forecasting Fertility

At best, the above procedure will provide an estimate of cohort completed fertility years ahead of observed completed fertility. For instance, we could (if the procedure worked perfectly) project now the completed fertility for the 1975 birth cohort (those 25 in 2000). A series of such estimates prior to 1975 coupled with observed completed parity for earlier cohorts would provide us with a valuable time series. Note that we are still not provided with any surefire way to predict completed fertility for subsequent cohorts, those born in 1990, 2000 or 2020. However, stability in completed fertility or a secular trend might provide some hints of the likely future course of completed fertility.

A second problem is that the cohort component projection methods (the dominant method of projecting populations) are period, not cohort based. Age-specific, period fertility rates (summed to create the total fertility rate, TFR) are the needed input for these projections. We suggest that intention data be used in conjunction with, and as a secondary source to period fertility data. Specifically, using the Bongaarts and Feeney (Bongaarts and Feeney 1998) framework and terminology, for long run projection one needs to forecast the underlying quantum component of period fertility.[15] The related concept for cohorts can be directly observed as completed fertility or projected using methods discussed above.

The Bongaarts and Feeney (Bongaarts and Feeney 1998) procedures allow estimates of current levels and trends in **quantum** (by parity). In our view, these quantum estimates should be the primary basis for future fertility forecasts (and the basis for period based projections). That is, demographers should examine trends in underlying period quantum parity by parity and forecast their future based on their assessment of how *the future* will be different from *the present*. The Bongaarts and Feeney approach has demonstrated predictive validity and is consistent with the dominant influence of period factors (NiBhrolchain 1992). Greater confidence in forecasting these **quantum** trends into the future would result if adjusted intended parity data were consistent with the Bongaarts-Feeney levels of quantum. In effect, intended parity rests on women's answer to a personal question that parallels the one faced by the forecaster: based on present circumstances how many children will women/you have? Expert judgment will be needed to reconcile any differences between Bongaarts and Feeney quantum-based estimates and adjusted intention data. Let us provide an illustration, for the total U.S. population intended parity for those in their early 20s hovers slightly above 2.0 (adjusted estimates will be approximately the same, see Appendix II). There is little trend over the past decade. The Bongaarts-Feeney adjusted TFR likewise shows underlying period quantum as relatively stable and approximately 2.0. Thus current behavioral and intention data show no downward secular trend and thus portend no additional fertility decline. Perhaps other factors suggest a decline, but those forecasting very low fertility in the U.S. must identify relevant factors, demonstrate their antinatalist impact, and justify the persistence of such effects decades into the future.

Summary

This paper revisits old issues: can fertility intentions be used to predict actual fertility behavior? The evidence is clear: intentions strongly predict subsequent behavior. However, the predictions are far from perfect and errors in prediction need not cancel. Thus aggregate predictions can contain substantial error. We propose models (embodying a research agenda) that would likely improve our ability to translate fertility intentions into more accurate estimates of completed cohort fertility. However, we caution that these estimates are of only moderate value in making fertility forecasts. Completed cohort fertility estimates do not fit nicely into the mechanics of cohort component projection techniques, i.e., the completed cohort quantum does not apply to any period. The Bongaarts and Feeney (Bongaarts and Feeney 1998) adjusted TFR does provide a period based estimate of quantum. Forecasters must assess whether the future will be like the present, thus justifying projection of current levels, or whether changed circumstances will be anti- or pronatalist justifying scenarios of decreased and increased levels of future fertility. Confidence in period quantum estimates would be heightened if they equaled recent estimates of intended parity for women in the heart of the childbearing years. Such a secondary role for intentions does not place predictive validity at the top of a fertility forecasting agenda. Nevertheless, predictive validity remains a useful line of inquiry for forecasters and is

[15] Changes in tempo, the second component of period fertility, have effects that are transitional. Thus, tempo is important for short and medium range projections but has more modest effects on long term projections (i.e., longer than 25 years).

of substantial substantive interest in understanding contemporary fertility. Where varying predictive validity is problematic for forecasting, it is an interesting phenomena relevant to various hypotheses about fertility decision-making.

References

Ajzen, I. and M. Fishbein (1977). "Attitude-Behavior Relations: A Theoretical Analysis and Review of Empirical Research." *Psychological Bulletin* **84**: 888-918.

Barber, J. S. (2000). "Intergenerational influences on the entry into parenthood: Mothers' preferences for family and nonfamily behavior." *Social Forces* **79**(1): 319-348.

Barber, J. S. (2001). "Ideational influences on the transition to parenthood: Attitudes toward childbearing and competing alternatives." *Social Psychology Quarterly* **64**(2): 101-127.

Barber, J. S. and W. G. Axin (1998). "The Impact of Parental Pressure for Grandchildren on Young People's Entry into Cohabitation and Marriage." *Population Studies* **52**: 129-144.

Becker, G. (1960). *An Economic Analysis of Fertility*. Princeton, Princeton University Press.

Beckman, L. J., R. Aizenberg, et al. (1983). "A Theoretical Analysis of Antecedents of Young Couples' Fertility Decision and Outcomes." *Demography* **20**(519-533).

Blake, J. (1974). "Can We Believe Recent Data on Birth Expectations in the United States?" *Demography* **11**: 25-44.

Bongaarts, J. and R. A. Bulatao, Eds. (2000). *Beyond Six Billion: Forecasting the World's Population*. Washington, DC., NAS Press.

Bongaarts, J. and G. Feeney (1998). "On the Quantum and Tempo of Fertility." *Population and Development Review* **24**: 271-291.

Bongaarts, J. and G. Feeney (1998). "On the quantum and tempo of fertility." *Population and Development Review* **24**: 271-291.

Bongaarts, J. and R. G. Potter (1983). *Fertility, Biology and Behavior*. New York: Academic Press.

Calhoun, C. A. and J. de Beer (1991). Birth expectations and population forecasts: the case of the Netherlands. *Future demographic trends in Europe and North America: what can we assume today?* W. Lutz. San Diego, California/London, England, Academic Press: 197-226.

Coombs, L. C. (1974). "The Measurement of Family Size Preferences and Subsequent Fertility." *Demography* **11**: 587-611.

Coombs, L. C. (1979a). "Reproductive Goals and Achieved Fertility: A Fifteen Year Perspective." *Demography* **16**: 523-534.

Coombs, L. C. (1979b). "Underlying Family-Size Preferences and Reproductive Behavior." *Studies in Family Planning* **10**: 25-36.

Coombs, L. C. and D. Fernandez (1978). "Husband-Wife Agreement About Reproductive Goals." *Demography* **15**: 57-73.

Davidson, A. R. and L. R. Beach (1981). "Error patterns in the prediction of fertility behavior." *Journal of Applied Social Psychology* **11**: 475-488.

Davidson, A. R. and J. J. Jaccard (1979). "Variables that Moderate The Attitude-Behavior Relation: Results from a Longitudinal Survey." *Journal of Personality and Social Psychology* **37**: 1364-1376.

de Beer, J. (1991). "From birth expectations to birth forecasts: a partial-adjustment approach." *Mathematical Population Studies* **3**: 127-144.

Freedman, R., D. S. Freedman, et al. (1980). "Changes in Fertility Expectations and Preferences between 1962 and 1977: Their Relation to Final Parity." *Demography* **17**: 365-378.

Heaton, T. B., C. K. Jacobson, et al. (1999). "Persistence and Change in Decisions to Remain Childless." *Journal of Marriage and the Family* **61**: 531-539.

Hendershot, G. E. and P. J. Placek (1980). *Predicting Fertility: Demographic Studies of Birth Expectations*. Lexington, Lexington Books.

Henshaw, S. K. (1998). "Unintended pregnancy in the United States." *Family Planning Perspectives* **30**(1): 24-29+46.

Lee, R. D. (1980). "Aiming at a Moving Target: Period Fertility and Changing Reproductive Goals." *Population Studies* **34**: 205-226.

Long, J. F. and S. I. Wetrogan (1981). The utility of birth expectations in population projections. *Predicting fertility: demographic studies of birth expectations*. P. P. Hendershot GE. Lexington, Massachusetts, D.C. Heath: 29-50.

Monnie, A. (1989). "Fertility Intentions and Actual Behaviour. A Longitudinal Study: 1974, 1976, 1979." *Population: An English Selection* **44**: 237-259.

Morgan, S. P. (1981). "Prayerfulness in America." *Chicago Studies* **20**: 237-252.

Morgan, S. P. (1982). "Intention and uncertainty at later stages of childbearing: The United States, 1965-70." *Demography* **18**: 267-286.

Morgan, S. P. (1985). "Individual and couple intentions for more children." *Demography* **22**: 125-132.

Morgan, S. P. (1991). "Late nineteenth and early twentieth century childlessness in the United States." *American Journal of Sociology* **97**: 779-807.

Mosher, W. D. (1982). "Infertility Trends Among U.S. Couples: 1965-1976." *Family Planning Perspectives* **14**(1): 22-27.

Namboodiri, N. K. (1972). "Some observations on the economic framework for fertility analysis." *Population Studies* **26**: 185-206.

Namboodiri, N. K. (1981). "On factors affecting fertility at different stages in the reproduction history: An exercise in cohort analysis." *Social Forces* **59**: 1114-1129.

NiBhrolchain, M. (1992). "Period paramount? A critique of the cohort approach to fertility." *Population and Development Review* **18**: 599-629.

Noack, T. and L. Ostby (2000). *Free to choose - but unable to stick to it? Norwegian fertility expectations and subsequent behavior for the following 20 years*. FFS Flagship Conference, Brussels.

O'Connell, M. and M. J. Moore (1977). "New Evidence on the Value of Birth Expectations." *Demography* **14**: 255-264.

O'Connell, M. and C. C. Rogers (1983). "Assessing Cohort Birth Expectations Data from the Current Population Survey, 1971-1981." *Demography* **20**: 369-384.

Pollard, M. S. and S. P. Morgan (2001). *Emerging Parental Gender Indifference? Sex Composition of Children and the Third Birth*. Annual Meeting of the Population Association of America, Washington, DC.

Remez, L. (2000). "Degree of certainty about plans to have children strongly predicts whether individuals will do so." *Family Planning Perspectives* **32**: 46-7.

Rindfuss, R. R. and L. L. Bumpass (1978). Age and the sociology of fertility: How old is too old. *Social Demography*. K. E. Taeuber, L. L. Bumpass and J. A. Sweet. New York, Academic Press: 43-56.

Rindfuss, R. R., S. P. Morgan, et al. (1988). *First Births in America*. Berkeley, U. of California Press.

Rindfuss, R. R., S. P. Morgan, et al. (1998). *First Births in America*. Berkeley, University of California Press.

Ryder, N. B. and C. F. Westoff (1965). *Reproduction in the United States*. Princeton, NJ, Princeton University Press.

Schoen, R., N. M. Astone, et al. (1999). "Do fertility intentions affect fertility behavior?" *Journal of Marriage and the Family* **61**: 790-799.

Thomson, E. (1997). "Couple Childbearing Desires, Intentions, and Births." *Demography* **34**: 343-354.

Thomson, E. and Y. Brandreth (1995). "Measuring Fertility Demand." *Demography* **32**: 81-96.

Thomson, E., E. McDonald, et al. (1990). "Fertility Desires and Fertility: Hers, His, and Theirs." *Demography* **27**: 579-588.

Trent, K. and K. Crowder (1997). "Adolescent birth intentions, social disadvantage, and behavioral outcomes." *Journal of Marriage and the Family* **59**: 523-535.

Trussell, J. and B. Vaughan (1989). "Aggregate and Lifetime Contraceptive Failure In the United States." *Family Planning Perspectives* **21**(5): 224-226.

Van de Giessen, H. (1992). Using birth expectation information in national population forecasts. *National Population Forecasting in Industrialized Countries*. N. a. H. C. Keilman. Amsterdam, Swets & Zeitlinger: Chapter 12.

Westoff, C. F. (1981). The validity of birth intentions: Evidence from U.S. longitudinal studies. *Predicting Fertility*. G. E. Hendershot and P. J. Placek. Lexington, Mass., Lexington Books: 51-59.

Westoff, C. F. and J. McCarthy (1990). "Reproductive Intentions and Fertility Rates." *International Family Planning Perspectives* **16**: 84-89+96.

Westoff, C. F., E. G. Mishler, et al. (1957). "Preferences in Size of Family and Eventual Fertility Twenty Years After." *American Journal of Sociology* **62**: 491-497.

Westoff, C. F. and N. B. Ryder (1977). "The Predictive Validity of Reproductive Intentions." *Demography* **14**: 431-453.

Williams, L., J. Abma, et al. (1999). "The Correspondence Between Intention to Avoid Childbearing and Subsequent Fertility: A Prospective Analysis." *Family Planning Perspectives* **31**: 220-227.

Wilson, F. D. and L. Bumpass (1973). "The Prediction of Fertility Among Catholics: A Longitudinal Analysis." *Demography* **10**: 591-597.

Wu, Z. and H. Wang (1998). "Third birth intentions and uncertainty in Canada." *Social Biology* **45**: 96-112.

Appendix I. Adjusting Fertility Intentions Data

Developing a firm scientific basis for projecting cohort completed fertility constitutes a major research agenda. In this appendix we: i) discuss issues that need to be resolved, ii) present illustrative examples of how fertility intention data might be adjusted using estimates of past predictive validity and iii) describe what is needed to put a data collection system in place to routinely estimate intended fertility and evaluate its predictive validity.

i) Issues to be Resolved When Adjusting Fertility Intentions

As noted in the text, a number of important details need to be resolved/considered in conjunction with the analytic strategy outlined. We discuss them briefly below:

- *Nonresponse to intention questions*. Substantial numbers of women report that they "don't know" how many children they will have or refuse to provide any answer to this question. For instance, in the 1980 CPS over 16% and 11% of women aged 18–21 and 22–25, respectively, did not provide numerical estimates to the fertility intention question. In a report available from the author, we show that there are significant race/ethnic differences in the likelihood of reporting fertility intentions. There is a substantial published literature on nonresponse to this question, its possible substantive meanings, and the bias that can result from ignoring it. At a minimum, investigators should be very concerned if nonresponse exceeds 10% and if nonresponse varies substantially across time/groups. The preferred strategy would involve estimating the bias produced by nonresponse and attempts at correcting this bias.
- *Top-coding of fertility intentions questions*. Most women intend fewer than four additional children. However, a nontrivial group report intentions for 6–10 children. For instance, 2.5% of women in the NLSY-79 intended 6 or more children. In contrast the CPS top codes intentions at 6 or more. The NLSY-79 data produce a mean additional intended number of births of 2.49, 2.47 and 2.2 if top-coded at 10, 6 and 3, respectively. We recommend top-coding if very high values are given (i.e., numbers above 4). The primary justification is substantive—there is likely little substantive difference between a woman who intends 4 and ten additional children. That is, both intend many and intervening events will largely determine the ultimate number that they bear. As we have noted in several places, the predictive validity of intentions for large numbers of children need to be assessed. We expect that predictive validity declines as number of additional children exceeds three. Preliminary calculations by the author from the NLSY-79 supports this claim.
- *Extent of disaggregation*. Disaggregation will likely improve estimates of intended parity. Specifically, there is substantial evidence that predictive validity varies by subgroups. If separate estimates of predictive validity by subgroups can be estimated and if the size of subgroup can be determined, then the standard decomposition (of aggregate differences into that due to rates and to distributions) can be applied. The advantages of this approach are offset by the demands that the distributional variables be measured similarly across time. A key variable for Census Bureau forecasts is race/ethnicity. Forecasts by group depend on consistent measurement of race/ethnicity.

ii) Estimating Predictive Validity

We will provide two extended examples of how one can estimate predictive validity with existing data. One example uses repeated cross-sections from the CPS; the second uses data from a multi-wave longitudinal survey, the 1979-NLSY.

Current Population Survey

Table AI-1 shows data on intended and current parity from repeated cross-sectional CPS surveys. The first row shows results for all women aged 22–25 in 1980. These women had a mean current parity of .74 and mean intentions for additional births of 1.05. Thus, their intended parity at this age was 1.79 (.74+1.05). For the second cross-section, we use the 1998 CPS the most recent data that includes parity. In 1998 these women would be 18 years older (i.e., ages 40–43) than in 1980. We exclude women who immigrated to the U.S. after 1980 from the 1998 sample, and calculate a mean parity of 1.87. Current parity exceeds intended parity 18 years earlier by a factor of 1.05. Births observed in the 18 year period exceed those intended at the beginning of the period by a factor of 1.07. This is impressive correspondence of intentions and behavior but indicates a surprising underestimation of completed fertility. Based on these data alone, one would inflate mean intentions for subsequent cohorts at 22–25 years of age by 1.07 when predicting completed cohort fertility. Replication of such estimates are certainly advised prior to actual application of this estimate.

Subsequent rows (of Table AI-1) assess predictive validity for race/ethnic groups. The aggregate predictive validity for non-Hispanic whites is nearly perfect, only slightly worse for blacks and significantly worse for white Hispanics. Again given only this information, one could apply these estimates of predictive validity to intentions data from subsequent cohorts. Overall cohort estimates would be obtained by calculating a weighted mean reflecting the specific cohort's race/ethnic composition.

National Longitudinal Survey of Youth

Figure AI-1 shows (for whites) intended parity and current parity by age for the set of cohorts included in the 1979 NLSY. The ages observed vary by cohort with older age ranges observed for the earlier cohorts, e.g., '57 cohort observed 22–41 and '65 cohort observed 14–33. The dominant pattern is one of convergence of intended and current parity, as would have to be true as cohorts reached the end of the childbearing years. But this convergence could result because of increasing parity or because of reductions in intended parity. Note that there are some reductions of intended parity at young ages; but intended parity is quite stable from the mid 20s until the oldest ages observed. The two most recent cohorts do show evidence of declining intentions into the 30s.

Figure AI-2 shows parallel data for blacks. Results are broadly similar. However, there is less tendency, even at young ages, for intended parity to decline with increasing age. Instead intended parity increases slightly with age and convergence results from increasing current parity.

To more explicitly compare the estimates of predictive validity for whites and blacks, we present Table AI-2. The final panel shows the ratio of estimated predictive validity for the two groups. For c1957 and c1959, blacks show greater predictive validity; we estimate the opposite result for c1958. Thus, there is not clear and consistent evidence of a racial difference in predictive validity using the NLSY79 data.

CPS and NLSY79 Compared

Reconciling these estimates provides a challenge. Most serious is the substantial shift in predictive validity across surveys. The CPS estimates all exceed 1.0 indicating more births than births intended. The NLSY79 results show the opposite result. This difference illustrates the importance of data comparability/data quality. We do not now have an explanation for these differences.

iii) Needed data for routine estimation of intentions and predictive validity

While preparing this report the problems and promise of CPS data for ongoing estimation of intended parity and its predictive validity became clear. From 1976 to 1988 and in 1990, 1992 and 1998 CPS June Supplements collected data on future intended births (for women 18–34) and on current parity for women aged 15–44. These data allow for tabulations like those shown in Table 4. One can follow cohorts as they age in such repeated cross-sections. For a population closed to migration and with no deaths, net of sampling variability the repeated cross-sections would allow perfect monitoring of changing aggregate intentions and of predictive validity of earlier intentions. Recent CPS surveys contain information on date of immigration, an important innovation for identifying cohorts of residents x years prior to the survey. It is probably safe to assume that the cross-sectional samples (of females aged 18–44) are not heavily biased by emigration or by mortality.

Features of current data collection efforts hinder a research agenda on predictive validity:

- The biggest problem is lax interest in collecting data on intentions. The baseline estimates are in place but subsequent data collection is needed. The need for yearly data can be questioned, but regular surveys at 2 or 3 year intervals is needed. (Intentions questions combined with retrospective fertility/marriage histories on a five year cycle is also attractive. See Below)
- The value of earlier data collection can be increased if the age range for collecting intentions data were increased (from the current 18–34) to ages 18–39. This change would acknowledge the substantial increase in birth at ages 35–39.
- Reinstatement of the CPS supplement on marriage and fertility histories coupled with the intention questions would have a number of benefits. Most importantly, it would allow one to define populations disaggregated by parity and marital status at time x–y. Thus, in year x one could identify retrospectively a cross-sectional sample of women who were married and parity 0 in year x–y. Their current parity and intentions (in year x) comes from the current survey. Their parity and intent in year x–y comes from current characteristics in the x–y cross-sectional survey. In this way, predictive validity for a host of characteristics could be estimated from the repeated cross-sections. With the CPS intentions supplement alone, predictive validity can only be estimated for fixed characteristics, e.g., birth cohort and race.

In short, the CPS has invested substantially in collecting data on fertility intentions. To continue monitoring at 2, 3 or 5 year intervals permits a serious research agenda on the long term predictive validity of intended parity. Modest changes in procedure and reinstatement of birth and marriage histories to the CPS would add substantially to a research agenda on predictive validity of reproductive intentions.

Table 1: Do Birth Intentions Predict Future Fertility?

STUDY	COUNTRY	INTENT DATE / OBS DATE	SURVEY	FINDINGS
(Beckman, Aizenberg et al. 1983)	USA	1977–78/ 1979–8	L.A. County sample	Empirical findings consistent with theoretical model specifying intentions as a primary mediating factor affecting contraceptive use and fertility.
(Barber 2001)	USA	1980/1993	Inter-generational Panel Study of Parents & Children	Used Coombs scales to measure intent and predicted transition to first birth. Study's relevance lies in the strong arguments and evidence that predictive validity varies by normative context (whether one examines marital or nonmarital fertility) and by presence of goals that compete with childbearing for women's time and energy in young adulthood.
(Calhoun and de Beer 1991)	Netherlands	1982/1988	Netherlands Fertility Survey	Use a bivariate probit model to model changing and uncertain birth expectations. The structural relationship between CEB and AEB was specified in the models. Their results indicate the 1988 Dutch population forecasts are consistent with the plausible assumptions on the explanatory variables in their multivariate model.
(Coombs 1974)	USA	1962/1967	Gathered by researcher	New measures for the preference of number of children were developed. Scales obtained from data at the initial interview demonstrated a strong relationship to fertility at follow up. The I-scales (developed in this study) were found to be predictive of future fertility net of a number of other variables usually associated with differential fertility.
(Davidson and Beach 1981)	USA		Gathered by researchers	This study tests the subjective expected utility model with multiple data sets. Predictive validity is high when predicting that a couple will not have a child (90%). However, predictive validity is much lower when predicting that the couple will have a child.
(Freedman, Freedman et al. 1980)	USA	1962/1977	Gathered by researcher; Detroit sample drawn from vital records	Expectations are predictive of fertility 15 years later. There was a general tendency to have fewer children than expected. The correspondence between initial expectations and final parity was strongly affected by the parity at the initial interview. At the aggregate level, final parity was below expectations for those at parity 0,1,2 at initial interview. Final parity was above expectations for those at parity 4 at initial interview.

Table 1, continued

STUDY	COUNTRY	INTENT DATE / OBS DATE	SURVEY	FINDINGS
(Heaton, Jacobson et al. 1999)	USA	1988/1994	NSFH	Researchers found a large number of shifts in childbearing decisions. Of those who desire to remain childless at the baseline interview, only 45% remained childless at the follow-up and still desired no children. An additional 25% had a child. The remaining individuals changed their earlier decision to remain childless.
(Monnie 1989)	France	1974/1979	INED	At the aggregate level, predictive validity was high. 484 children were 'intended' at the beginning of the five year period and 496 children were born in the five year period. At the individual level, substantial differences between intentions and behavior are found. The author judges the results as disappointing "from the point of view of the forecaster."
(Noack and Ostby 2000)	Norway	1977–99 1988–99	Survey with follow-up using pop. register	Women in most groups did not realize their stated intentions. This was especially true for those who were young, or unmarried or intended more than 2 children.
(O'Connell and Moore 1977)	USA	1967/1971	CPS	Little evidence suggests that the 1967 and 1971 CPS aggregate data on birth expectations will underestimate actual fertility. In fact, the authors believe that these expectations will actually overestimate actual fertility slightly.
(O'Connell and Rogers 1983)	USA	1971/1981	CPS	At the aggregate level, the average number of births expected by married women 18-39 in June 1971 will come close to their eventual completed family size. Expected lifetime births seem to be overstated for single women and slightly understated for married women.
(Remez 2000)	USA	1987–1988 / 1992–1994	NSFH	An individual's intention to have a child predicts whether or not they will do so. The relationship between certainty of having a child and fertility is strong and significant even when other life-course variables are controlled.

Table 1, continued

STUDY	COUNTRY	INTENT DATE / OBS DATE	SURVEY	FINDINGS
(Rindfuss, Morgan et al. 1988)	USA	1973/1979	NLS of High School Class of 1972	Respondents were asked a parity- and time-specific question: do you expect to have a child in the next year, 2–3 years, etc. Research focuses only on the first birth. Answers strongly predicted whether respondents had births in the subsequent 3 and 6 year period. However, a strong pattern of fertility delay was observed, i.e., women were much more likely to delay births vis-à-vis intentions than to advance them.
(Schoen, Astone et al. 1999)	USA	1987–1988 / 1992–1994	NSFH	There is a strong relationship between 1987–88 intentions (and their certainty) and the percentage having a birth in the five year period. Overall, only 10% of those who are "very sure, no" have a birth, and 64% of those who are "very sure, yes" have a birth. Effects are similar for men and for women and are visible at all parities.
(Thomson 1997)	USA	1987–1988/ 1992–1994	NSFH	The husband's intentions contribute to predictive validity. The effects of partners' desires or intentions were not additive. Disagreement of intentions shifted the couples towards not having a child. Author claims results demonstrates the importance of collecting intentions from both partners.
(Thomson, McDonald et al. 1990)	USA	1957/1967	Princeton Fertility Survey	For a sample of couples at parity two, both husbands' and wives' intentions are predictive. Moreover, the effects of partner's intentions are additive. Disagreement reduces the likelihood of either spouse achieving desired fertility.
(Trent and Crowder 1997)	USA	1979/1984	NLSY	Fertility intentions were predictive for young women in this five year period. However, intentions did not mediate the effects of important variables like race and poverty status on the likelihood of a birth.
(Van de Giessen 1992)	Netherlands	1982/1988	Netherlands Fertility Survey	Author concludes that women in surveys seem to be able to provide accurate fertility forecasts at the aggregate level. However, adjustments need to be made. Adjustment methods are discussed.

Table 1, continued

STUDY	COUNTRY	INTENT DATE / OBS DATE	SURVEY	FINDINGS
(Westoff and Ryder 1977)	USA	1970/1975	National Fertility Survey	Aggregate intentions in 1970 overstated 1971-75 fertility intentions by same degree as did a projection based on period fertility. At individual level, intentions have substantial predictive power.
(Westoff and McCarthy 1990)	World	Mixed	WFS, DHS	Strong correlation between percentage of women who desire no more children and TFR. Significant inverse relationship between the two. This relationship is for 84 countries, 68 of which are considered developing countries.
(Westoff, Mishler et al. 1957)	USA	1937/ 1953–54	Gathered by Kelly (1955)	At the aggregate level, predicted fertility and achieved fertility were very similar. The mean number of children born per couple was 2.62 and the mean number of children desired was 2.64 for men and 2.79 for women. However, on the individual level there was substantial inconsistency.
(Williams, Abma et al. 1999)	USA	1988/1990	NSFG	Intentions were predictive of short-term fertility behavior. But group differences were found, e.g., women below the poverty level (compared to those above) were 2–3 times more likely to have an unpredicted birth.
(Wilson and Bumpass 1973)	USA	1965/1969	Sample of Catholics from ’65 NFS	Intentions were predictive of subsequent behavior.

Table 2. Does Including Information From Partner Increase Predictive Validity?

STUDY	FINDINGS
(Beckman, Aizenberg et al. 1983)	Compared to the husband's intent, wife's intention had a much larger net effect on contraceptive use and fertility in the 18–20 month period after first interview.
(Schoen, Astone et al. 1999)	For those married at the initial interview, considering the intentions of the spouse significantly improves model fit. One cannot reject the hypothesis that husbands' and wives' intentions have equal effect.
(Thomson 1997)	Husbands' desires and intentions influence the couples' fertility. The effects of partners' intentions were not additive. When intentions diverged, behavior shifted toward not having a child.
(Thomson, McDonald et al. 1990)	The effects of partner's intentions are additive. Disagreement reduces the likelihood of either spouse achieving desired fertility.

Table 3. Evidence that the predictive validity of intentions varies by subgroup

STUDY	VARIED BY	NOT VARY BY	DETAILED RESULTS
(Coombs 1974)		Parity, Wife's Education, Income, Religion	Predictive validity observed for Coombs I-scales for a range of subpopulations defined by education, income, religion, parity
(Freedman, Freedman et al. 1980)	Parity, Wife's Education	Husband's Education, Wife's Religion, Income	The relationship between expectations and final parity was affected by the initial parity. Higher education for the wife led to higher expectations than actual fertility. No consistent differences in predictive validity were found by husband's education, wife's religion, or income.
(Heaton, Jacobson et al. 1999)	Age, Race, Education	Gender, Employment	As people get older, their intentions change. Blacks are more likely to have children than whites, and blacks are less likely to decide to be childless or change their minds from wanting children to not wanting children. People with higher education are less likely to change from not wanting children to wanting children than are people with lower levels of education.
(Monnie 1989)	Parity		Predictive validity was highest for women who had 1 child at time 1 and wanted a second and for women who had at least 2 children and did not intend to have any more. Predictive validity was much lower for women who had 1 child and did not want a second and for women who had 2 children and desired a third.
(O'Connell and Rogers 1983)	Age, Marital Status		The youngest age group of married women overestimated their future fertility more than the older age groups of married women over a 5 year period. These differences were not as distinct for the 10 year period. The birth expectations of married women at parity 1 at time 1 were slightly exceeded. Married women at parity 0 at time 1 will probably not meet their birth expectations. Expectations data for single women are especially susceptible to subsequent marital events.
(Remez 2000)	Marital Status	Age, Parity, School Enrollment, Employment, Education, …	Only marital status predicts actual fertility as well as fertility intentions. Women whose husbands were more sure about having a child were more likely to have a child. Age, parity, school enrollment status, employment status, education level, respondent's mother's level of education and income were not significant predictors when in the model with intentions.
(Schoen, Astone et al. 1999)	Marital Status, Spousal Intentions		Those married at time 1 are more likely to achieve intended fertility than are those single at time 1.

Table 3, continued

STUDY	VARIED BY	NOT VARY BY	DETAILED RESULTS
(Thomson 1997)	Wife's Age, Parity, Age of Youngest Child, Religion, Gender Role Attitudes	Gender Role Attitudes Held by Husband	Several control variables included in the full model had direct relationships with achieved fertility. Wife's age, couples parity, age of youngest child and religious preference had direct effects on birth risks. Amount of hours worked by the husband and gender role attitudes held by the wife also affects birth risks.
(Westoff and Ryder 1977)	Education, Wife's Working Status	Parity, Religion	Predictive validity shows little variation by parity. Individual inconsistency for both women who intend to have more births and those who intend not to have more births, declines with education but this relationship disappears when a duration control is introduced. There is no difference between Catholics and non-Catholics. Predictive validity is lower for working wives (compared to those not working).
(Westoff, Mishler et al. 1957)	Planning		Planners (those who planned pregnancies) and non-planners (those who had unwanted births) desired virtually the same number of children. Planners, however, averaged about 1 less child per couple than did non-planners.

Table 4. Current and Intended Births per 1,000 women: CPS Selected years

Year	Total: 18-34		18-24		25 to 29		30 to 34	
	Births to date	Lifetime births expected	Births to date	Lifetime births expected	Births to date	Lifetime births expected	Births to date	Lifetime births expected
1998	1,104	2,045	532	1,936	1,150	2,082	1,662	2,127
1992	1,135	2,098	521	2,053	1,181	2,137	1,679	2,106
1990	1,130	2,116	537	2,062	1,152	2,152	1,695	2,135
1988	1,095	2,073	459	2,045	1,163	2,116	1,674	2,057
1987	1,125	2,074	503	2,057	1,219	2,111	1,689	2,055
1986	1,114	2,099	497	2,087	1,171	2,117	1,734	2,094
1985	1,098	2,062	508	2,046	1,193	2,113	1,674	2,029
1983	1,096	2,079	481	2,071	1,220	2,082	1,786	2,088
1982	1,086	2,023	453	1,994	1,241	2,026	1,792	2,059
1981	1,136	2,048	517	2,033	1,273	2,012	1,857	2,106
1980	1,127	2,059	507	2,023	1,238	2,022	1,905	2,150
1979	1,144	2,072	514	2,033	1,269	2,033	1,942	2,170
1978	1,171	2,113	496	2,033	1,324	2,060	2,066	2,297
1977	1,197	2,133	485	2,052	1,345	2,049	2,150	2,351
1976	1,263	2,160	528	2,030	1,442	2,098	2,266	2,445

(Source: http://www.census.gov/population/socdemo/fertility/tabH8.txt)

Table AI-1. Predictive validity of reproductive intentions from two CPS cross-sections: 1980 and 1998

Sample:	1980 CPS Data (Age 22-25)				1998 CPS Data (Age 40-43)		Predictive Validity	
	Parity (1)	Intentions (2)	Intended Parity (3)	(N) (4)	Current Parity (5)	(N) (6)	Final Parity (9)/(8) (7)	Ad. Births [(5)-(1)]/(2) (8)
All Women	0.741	1.053	1.793	5621	1.87	3856	1.04	1.07
Race								
White	0.686	1.101	1.787	4800	1.84	3285	1.03	1.05
Black	1.129	0.736	1.866	656	1.96	438	1.05	1.13
Other	0.755	0.903	1.658	165	2.29	133	1.38	1.70
Hispanic origin								
Hispanic	1.044	0.841	1.884	370	2.35	265	1.25	1.56
Non-Hispanic	0.718	1.072	1.789	5146	1.83	3591	1.02	1.04
Race/Ethnicity								
White Hispanic	1.041	0.870	1.911	355	2.36	250	1.23	1.52
White Non-Hispanic	0.658	1.126	1.783	4346	1.80	3035	1.01	1.01
Black Hispanic	1.500	0.818	2.318	11	1.70	10		
Black Non-Hispanic	1.122	0.741	1.862	640	1.96	428	1.06	1.14

Note: Seven percent of the 1998 sample immigrated after 1980 and, thus, were excluded from 1998 estimates.

Table AI-2. Predictive validity for Whites and Blacks: NLSY79 Selected Cohorts

Race	Cohort	Age Range	Final Parity/ Intended Parity	Additional Births/ Intended Births
Whites	c57	20-39	0.836	0.816
		24-39	0.846	0.772
	c58	21-40	0.876	0.850
		24-40	0.906	0.858
	c59	22-41	0.841	0.810
		25-41	0.938	0.905
Blacks	c57	20-39	0.905	0.875
		24-39	0.914	0.826
	c58	21-40	0.837	0.745
		24-40	0.822	0.656
	c59	22-41	1.062	1.084
		25-41	0.992	0.986
Blacks/ Whites				
	c57	20-39	1.083	1.072
		24-39	1.080	1.070
	c58	21-40	0.955	0.877
		24-40	0.907	0.764
	c59	22-41	1.263	1.337
		25-41	1.059	1.089

Figure 1. Strong Model Linking Intentions and Behavior

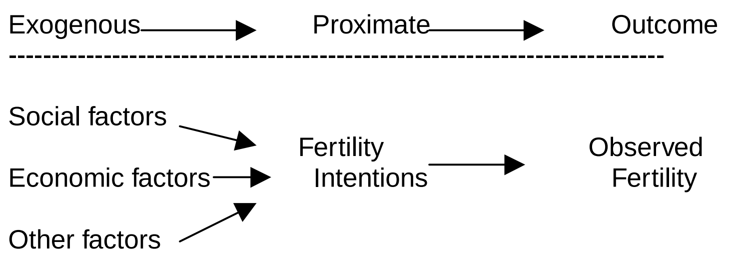
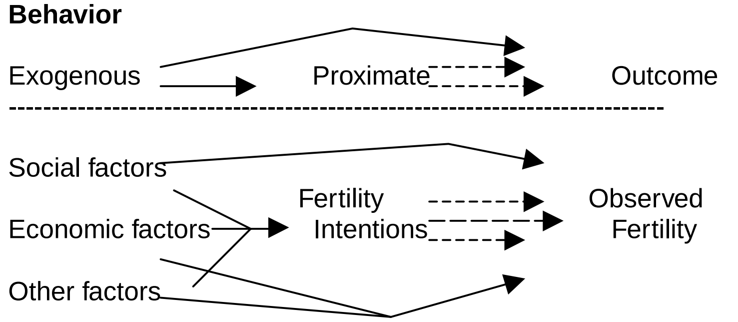


Figure 2. More Complicated Model Linking Intentions and Behavior



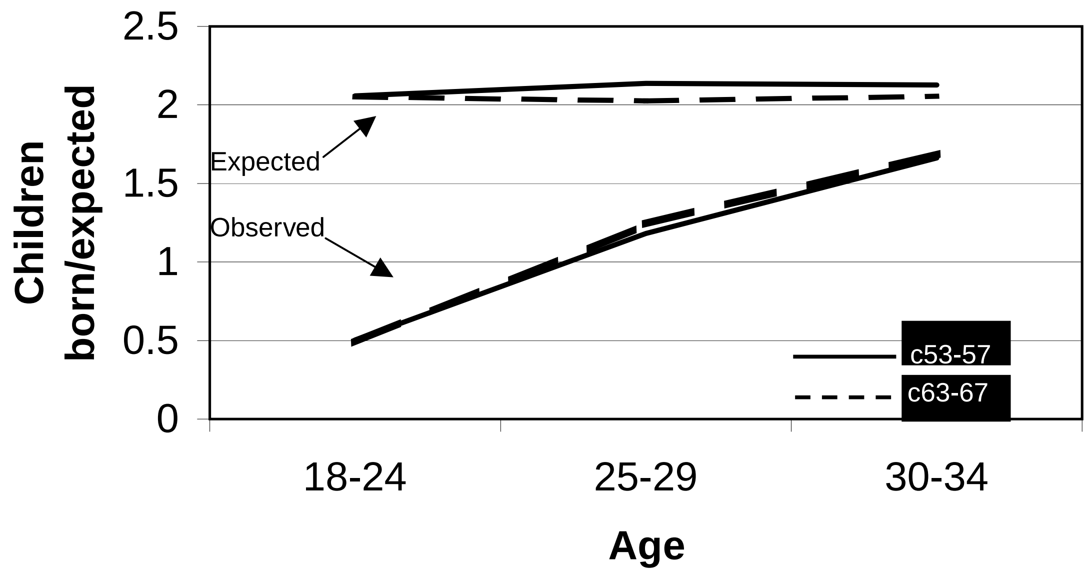
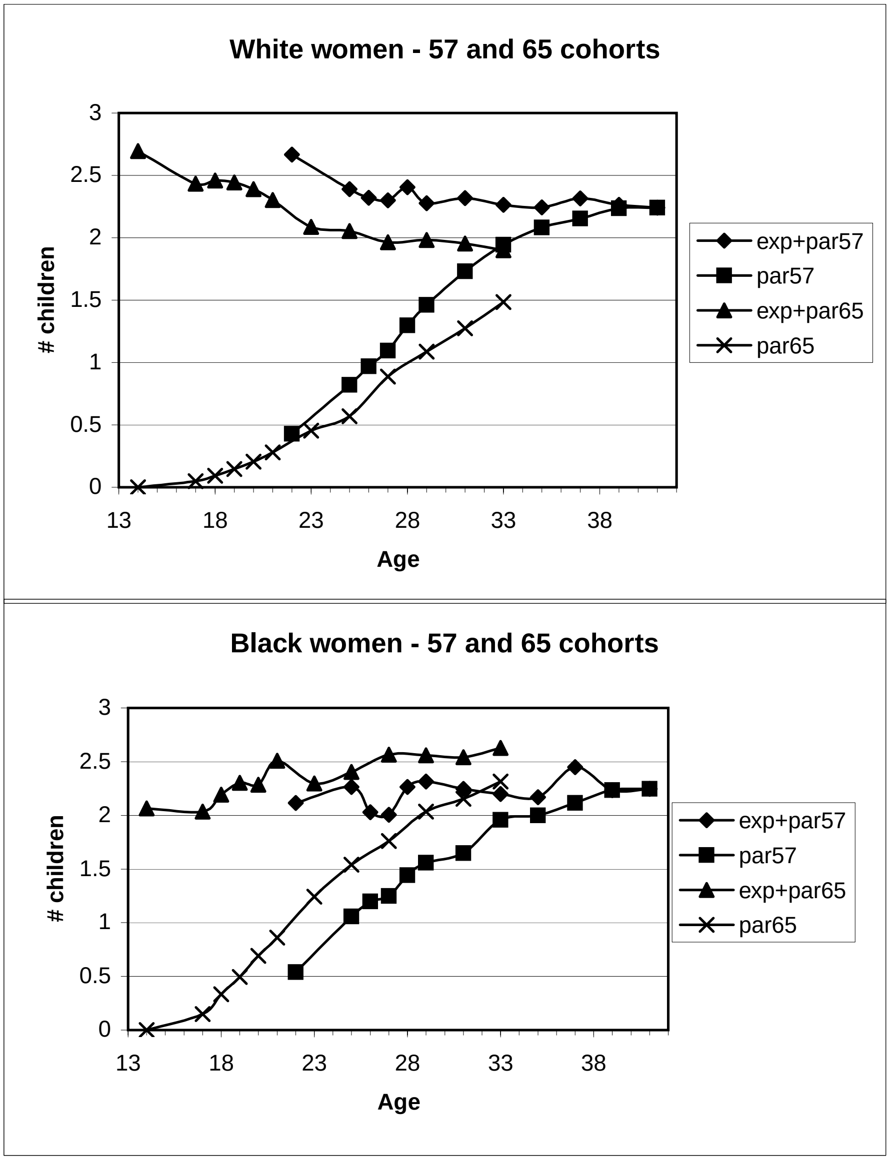


Figure 3. Current and Intended Parity: Selected Cohorts

Source: Table 4.

Figure 4. Current and Intended Parity: NLSY-79 Selected Cohorts



Discussion

Nico Keilman, *Discussant*

The paper gives a systematic discussion of various theoretical issues connected to the measurement of fertility intentions, and it presents an outline of a model that transforms aggregate intentions into Completed Cohort Fertility values for currently young women.

Phil mentions several problems that arise when one intends to measure fertility intentions.

- The fixed-target nature of intentions in many of the studies, e.g. those that simply add current parity to the number of additionally intended births, either at the level of an individual woman or at the aggregate level of birth cohorts.

- The absence of a time reference
  Phil notes that a woman, who intends to have her first child within a year, behaves very differently compared to a woman of the same age who wants to delay her first child for at least five years. This problem is easy to solve: In a review of 21 fertility intention surveys in Western Europe and the U.S., Wim van Hoorn and I found that roughly in half of those surveys, a question on the *timing* of the next child was included. The USA (CPS surveys of 1971–1994) was *not* among them.

- The absence of contextual reference, such as the sex of the future child, or attitudes to alternatives to parenthood (work, education).

- Ignoring the level of certainty Phil stresses the importance of including a range (one or two children) for those women who cannot be precise. In earlier work from the 1980s he has argued that "don't know" answers should be included in the analysis, because they give valuable information, in particular when subsequent surveys are studied.

- Period factors may distort intentions, compared to actual behaviour. In addition, age/period interactions may be important: in certain periods, postponement has been stronger than in others, but in both cases postponement led to too high (short-term) fertility expectations.

- The role of marriage is important: married women have more reliable intentions than unmarried women, Phil notes. To this I add that a possible increase in unmarried cohabitation in the U.S. in the future will diminish the usefulness of birth expectations for fertility forecasting.

One problem not mentioned by Phil is the representativeness of the data set. In many cases (11 out of 21) we were informed that specific groups of women were underrepresented in the survey: young women, childless women, women in the big cities are frequently mentioned. In many cases this non-response bias is strongly linked to childbearing intentions. Thus weighting the data set may be considered to improve intentions estimates, but this does seldom help. Nor is it always possible.

I was less enthusiastic about Phil's conclusion that contemporary U.S. cohorts predicted their fertility quite well. He reaches that conclusion by plotting actual parity and intended parity for various ages, see for instance his Figure 4. Here I am not convinced.

First of all, we see that intended parity is not stable over the life course. For white women it falls by at least half a child, for black women it increases by the same amount for the two cohorts concerned (presumably as a consequence of unintended births). Other white cohorts in Figure AII-1 show the same pattern; for black cohorts it is less clear. The apparent stability in intentions in Figure 3 could be the result of opposite trends for white and black women.

Thus intentions expressed at young ages are easily off by half a child, and therefore not very reliable indicators for mean parity at age 40.

This conclusion agrees with what we found for French, Finnish, and Norwegian women born in the 1960s, when their parity was checked in the early 1990s: intentions that women expressed in their 20s were too high by between 0.5 and 0.9 children per woman. At the same time, we found a near linear relationship between the Period Total Fertility Rate in the survey year, and expected total number of children for women aged 20-24 in nine European countries. Thus current norms and attitudes still seem to have a strong impact on stated birth intentions, as Westoff already concluded in 1981.

Because of the high levels of teenage childbearing in the U.S., we have to focus on young women, those aged 18–20 years, say. For forecasting purposes then, the intentions of these young women ideally should reflect actual parity at age 45, but these intentions are not very reliable, not even in the aggregate (completed cohort fertility). Aggregate intentions vary as much as the period TFR does.

In the final part of the paper, Phil proposes to develop adjustment factors that adjust stated mean fertility intentions for various subgroups, defined by race/ethnicity/marital status. The adjustment factors are observed for cohorts of which the whole childbearing history is known, and next they are projected for younger cohorts. This way one obtains extrapolated cohort TFRs for current young cohorts, say those born around 1980, who gave their intentions. These have to be translated into period TFRs for the year 2000–2020, say. Phil proposes to use the Bongaarts-Feeney method (which corrects period TFRs by parity for tempo distortions) as a starting point for this translation.

I have two remarks concerning Phil's proposal.

1. The method based on adjustments factors is attractive because it builds on earlier work by O'Connell and Rogers, Van de Giessen, and Calhoun and De Beer. But at the same time, a drawback is that it focuses on the mean number of children for each subgroup. While it is sufficient, for forecasting purposes, to predict mean numbers of children for subgroups, individual differences between women in a particular subgroup are masked—yet these will be important. They are important in order to understand patterns in the adjustment factors—after all, these have to be extrapolated. Also, as Phil mentions, a situation with three women having one child each is very different from one woman having three children. Therefore, in my view, an individual approach is to be preferred, in which the number of children a woman has at age 45, say, is a function of her childbearing intention at age 20, say, and a number of covariates: not only race/ethnicity and marital status, as Phil proposed, but also whether she is in first or second union, her religion, her attitudes to having children, her educational career, and her professional career, etc. Many of these covariates will be time-dependent.

   The dependent variable would be clearly ordinal: 0, 1, 2, 3, 4, 5, and perhaps 6 children. Special attention should be given to childless women: either intended at age 20, or actual at age 45 (or both). Probably these women are different from the rest in terms of partner status, or attitudes, or professional career. Uncertain women, who can only indicate an intended parity range, or who perhaps answered "don't know," should also receive special attention. In fact, the model I propose here is somewhat similar to model 2 in Phil's paper.

   While many important aspects clearly need to be resolved still, such an individual-based model, when successful, provides much more insight in the intentions/actual gap, than an aggregate approach.

   At the same time, by proper aggregation over covariate categories, one can compute implicit adjustment factors of the kind Phil proposed. Or, to turn the argument around, Phil's adjustment factors may be considered as a first descriptive step towards an individual-based model.

2. A second comment concerns Phil's proposal to use the Bongaarts-Feeney method when translating a series of CCFs in to period TFRs.

   I must say that I am not convinced of the validity of the B–F approach, and I am not alone in that. One important weakness of the B–F method is that it is based on age- and parity-specific fertility rates, that do not properly reflect fertility risks in the sense of occurrence/exposure rates. As a consequence, the parity-specific adjusted TFRs are influenced by current parity distribution, and B–F-adjusted period TFRs exaggerate tempo effects. (Similarly, the Total First Marriage Rate reported by Tomas Frejka in his section on Marriage and Cohabitation (4.1.3) underestimates period marriage levels by some 10–20 percentage points.)

   Fortunately, the translation from CCF to period TFR can ignore parity, and a very simple approach, based on extrapolated mean or median ages at childbearing, together with a reasonable parameterized age pattern (for instance on the basis of a Gamma or a Normal curve) are sufficient.

To sum up:

- I share Phil's *conceptual* scepticism.
- I am a bit more sceptical about the *empirical* correspondence between intended and realized parity.
- I would propose an individual-based model, rather than one which centers around aggregates.
- And I would give special attention to childless women, either intended or realized, to representativeness, and to "don't knows."

George Masnick, *Discussant*

Background

Before turning to Phil Morgan's excellent discussion of birth expectations, it will be useful to underscore some of the findings from Frejka's and Kingkade's discussion of recent U.S. fertility trends. At first glance, not much appears to be happening with U.S. fertility, especially when compared to recent trends in many industrialized countries. The total period fertility rate in the U.S. has hovered between 2.0 and 2.1 since 1990. During the same time period, the average total period fertility rate in 18 "western" European countries has declined steadily from about 1.7 to under 1.6. In the three "overseas" countries of New Zealand, Australia, and Canada, the total fertility rate has declined from an average of 1.9 to under 1.8 (all data discussed in this section are from Frejka and Kingkade).

If one looks at the completed fertility of U.S. cohorts, stable replacement levels of 2.1 have characterized women born between 1953 (47 years of age in 2000) and 1963 (age 37 in 2000). This stability however is the result of steadily declining non-white cohort fertility and slightly rising white fertility (most likely due to increasing representation of recent Hispanic foreign migrants in the white population).

Although European fertility levels, on average, have been trending downward, this is the result of considerable variation in trends across countries. Completed cohort fertility levels in Scandinavian nations have remained fairly stable for the past decade, similar to the U.S. Some Scandinavian levels are approximately equal to or slightly above the U.S. (Sweden and Norway), and some slightly below (Finland and Denmark).

More importantly for possible future U.S. fertility trends, completed cohort fertility in some European countries has recently begun to decline after years of relative stability (Switzerland, France, England and Wales, and the Netherlands). Some have experienced longer-term declines and now stand well below replacement fertility (Greece, Italy, and Spain). Underlying these fertility declines in Europe has been a fundamental shift in the age pattern of fertility toward delay of the first birth.

While U.S. completed cohort fertility has been stable, there is definite evidence that a shift toward delay in childbearing is also occurring here. One measure of the extent to which fertility is being delayed is the proportion of completed fertility achieved by age 27. Even though the proportion of completed fertility achieved by women by age 27 is highest in the United States compared to other western countries, and the decline in this measure of timing between the 1950–51 and 1960–61 cohorts in European countries has been dramatic, the U.S. decline has nonetheless also been significant.

The measure of the proportion of completed fertility can only be made for cohorts having already reached age 40 or so. An alternative measure of the delay in childbearing will confirm that the delay in U.S. childbearing has continued through the early 1970s cohorts as well. This alternative measure is the average number of person-years of motherhood experienced by each cohort by age 27 (equivalent to the area under the cumulative cohort fertility curves through age 27 in Frejka and Kingkade Figure 9). The average person-years of motherhood can be divided by the cumulative fertility level through age 27 to give an average number of person-years per child. Focusing on person-years of motherhood for just first births can further refine this alternative measure of timing.[16]

Other measures of timing such as the mean age of cohort fertility has also shown a steady increase among U.S. women during the past 25 years. This trend is occurring for both whites and non-whites. There had been a slowdown in the increase in mean childbearing age for whites (but not for non-whites), but again this could be explained by the increasing representation of recent Hispanic immigrants (who on average have earlier fertility) in the white population base.

Will the increasing delays in childbearing continue in the United States, and perhaps even be accelerated to the degree it has in other Western countries? Will the increasing delay in childbearing eventually lead to lower completed fertility? Or will the delays be fully compensated in the later half of the cohort's childbearing years, as current birth expectation data would perhaps let us believe? Are birth expectation data even useful in developing forecasts of completed fertility?

Birth Expectations

Because both period and cohort TFRs have remained close to 2.0 in the U.S., and because U.S. women have stated their intentions early in their reproductive lives to have on average two children, some analysts have concluded that birth expectations are good predictors of completed fertility. Morgan presents new data to show that there here are several pitfalls in reaching such a conclusion.

First, there is some evidence that birth expectations are not fixed over the life course of individual cohorts, but change to reflect the realities of their childbearing experience as it unfolds. For example, the NLSY data examined by Morgan indicate that the delayed childbearing of white women in the 1965 birth cohort has been accompanied by a fairly steady decline in their average birth expectations. For black women in the same cohort, fairly rapid early childbearing has caused them to re-evaluate their initial average birth expectations of only 2.0 children stated when they were in their teens. They have subsequently raised their average expectations to above 2.5 completed total births when they were last interviewed in 1998 (when they were 33 years old).

Second, Morgan shows with NLSY data that individual American women have not been very good at predicting their own fertility. Many of those initially stating an intention to have zero children actually went on to have one or more, and many of those expecting to have three or more actually had two or less. It appears that *average* cohort fertility has been fairly close to *average* cohort birth expectations over the past decade or so only because of the happenstance that these errors for individual women cancelled each other. With both inability to accurately predict one's own fertility very well, and with average cohort birth expectations apparently changing to reflect a cohort's actual childbearing experiences, it would appear that expectation data, the way they have been collected, are structurally flawed as helpful indicators of future fertility.

Morgan proposes an equation that explicitly recognizes that birth expectation data could perhaps be made more useful if adjusted, with more research needed on what would cause some women to either underestimate or overestimate their completed fertility. Perhaps a better direction for future research is how changing the focus to include critical tempo as well as quantum elements might refine birth expectation data. The current delay in the timing of first birth might have been better predicted if women were asked about their childbearing expectations before age 27, for example. For women at different parities at each age, a question might be posed about childbearing intentions over the next year or next five years. Finally, birth expectation data focusing on women having reached the last half of their childbearing years with zero or one child ever born might be particularly useful when attempting to better understand the implications of postponed fertility for completed family size.

[16] This measure of cohort fertility tempo was first suggested in George S. Masnick, Observing the Modern Fertility Pattern: An Essay in Methodology, PH.D. Dissertation, Brown University, 1970.

Floor Discussion

Goldstein

With regard to Masnick's comment that it takes very little to change a birth forecast by several hundred thousand. The message should be that the Census Bureau can't be any more precise than that. That would be a useful thing for Census to accomplish.

It is also possible that expectations data seem to do OK in the U.S. because the average tends to be near 2.0—about where the TFRS is—so the apparent success of birth expectations is a coincidence.

Bean

Regarding the debate about cohort versus period data. We may be confusing methodology with usefulness. There is no doubt the period data are needed from a policy point of view.

Regarding the Morgan chart which showed how badly people predicted individual parities. What he noticed was offsetting errors—the tendency to overshoot at low parities and undershoot at parities 3+—with the average doing reasonably well. He thinks we could model these error tendencies.

Morgan

He couldn't find the data, but recalls that those who intended to be childless actually averaged "1", and those who said "3" actually averaged "2.25"—That's not so good to me.

O'Connell

He agrees with Morgan–remembers that in the early 80s cohorts always said lifetime childlessness would be 10%—actually for all post-WWII cohorts it is going to be closer to 18 to 20%—basically a doubling of the initial intentions.

Ahlburg

Looking at the schematic of models proposed by Morgan—Ahlburg has tried to estimate similar models. He found that adding fertility expectations made the forecast worse. The reason is that actual fertility is a leading indicator of birth expectations. He thinks we have a problem with reversed causality in these models and that is why he thinks expectations will not help.

McDonald

He is working on a study where women were asked how important is it to you to have a birth in next 3 years—then reinterviewed them 3 years later to find out who had a birth and who did not and why. It turns out that what happened to their relationship was the most important factor. Those who remained in a relationship guessed pretty accurately. Those who stayed out of a relationship guessed well also. It was those in changed relationships who had fewer or more children than they originally intended.

Murphy

There must be a lot of information uncertainty in the future at the individual level—this suggests that the idea of individual predictability is almost meaningless. He has heard nothing that would indicate that intentions can identify turning points—He thinks that is because they are a really an extrapolation of past trends.

O'Connell

We did use such information in the 80's to predict a turning point. What happened was that a lot of childless women(20 to 30% of them) continued to say they expected to become mothers. And that is what eventually happened. Now, to be fair, more said they would have children than actually ended up with children. But my point is that they did imply a change in trend which eventually did occur.

Schoen

He is also skeptical about intentions and does not think much of their validity as a "target" type model, especially at the individual level. But nothing else has been found to be as good an aggregate level predictor—After all, its not as if we have the ability to forecast the future with extraordinary precision. It does seem that it is more accurate if you wait until 22 or 23 years of age.

Van de Kaa

Ronald Inglehart has developed a measure of "materialism" (law and order and economic well-being) as compared to "post-materialism" (individual well-being, environmentalists). Dirk divided up Europe according to these measures. He found that the post-materialists had higher family size desires, but ended up with much lower fertility than materialists. This is important if our societies are trending towards either type.

Morgan

Birth expectations simply mean how many more additional children do you expect to have? No other referents at all. I think people interpret this as how many kids they will have if things change as they expected. I agree with Schoen that nothing seems to work any better. I would first study the quantum component of period fertility—but I wouldn't ignore intentions.

O'Connell

In early 70s we asked both how many kids to you expect and how many do you expect in next 5 years? It was dropped during the 70s because fertility fell so rapidly that the 5 year question was disastrously over-estimated. Then by early 90s we had had many years of people saying all expectations data was useless and we now collect it only sporadically. We lost the retrospective fertility history question because NICHD funded the question and it was very, very expensive. They eventually decided to fund other surveys.

Weed

It seems that the high levels of divorce and remarriage may play a role. Especially if we look at how different the White and Black patterns are in terms of postponement and separation.

Wednesday Morning

Open Discussion

What Did We Learn?

Wetrogan

What have we learned? It seems that nothing will help us forecast the turning points in American fertility. It is no surprise that births expectations data seem to parallel the current period fertility patterns rather than predicting turnarounds. The discussion appears to suggest that we need to disaggregate the fertility even further than we have to date. Perhaps we need to pay closer attention to the data by parity, generation, nativity, and even geography. We are still faced with the question—Will America Remain an Outlier in Fertility?

Will America Remain an Outlier in Fertility?

Bean

He thinks we need to disaggregate by both nativity and origin as all foreign born people are not the same. He suggests we may need nativity information more than race. He also asked: are Hispanics an immigrant population or are they the Latino culture group.——Bean thinks immigrant differences will be more temporary than are cultural differences.

Frejka's table 10 shows most of the fertility differential is due to nativity.

Bean suggests that tempo is more important in low fertility environments. Is there any reason fertility might increase because of tempo compression alone?

Frejka did a wonderful job of looking at rich data and he showed courage by postulating conclusions from that data.

Day

Politics does make a difference—limits on abortion could raise fertility.

McDonald

Low fertility in Europe was due to the delay in first births which was never compensated for because of high unemployment, need for education, etc.

The United States has very early ages at first birth. How long can that continue? This is important as teen births predict higher births.

Women can combine work and children by delaying childbearing but catching up eventually.

What is the purpose of fertility projections? The total is not that important to him—-rather it is the geography of kids that matters.

Morgan

Tempo is an important factor only for the short-term. Teenage fertility trend in the U.S. is down and the total TFR is still 2. But U.S. teen fertility is still four times that of any other country.

Rao

Turning points may be driven by changes in the foreign born—-but the foreign-born is too small in the CPS to ever know it.

He suggests that projections should only have a horizon of 25 years or so.

Henshaw

The effect of restricting abortion on fertility is unclear. It might go up or down. He sees little chance of abortion being made illegal.

O'Connell

The sample variability in CPS means you can't disaggregate too many ways. The American Community Survey, by contrast is expected to have 3 million cases a year and will be a better tool for detailed analysis of groups such as the foreign born.

Chamie

He agrees that mortality should be the most accurate forecast because all people want to live. He suggests that one problem many countries have with migration forecasts is the constraint that the forecasters must use official predictions. All sorts of differences in fertility behavior have disappeared in the past few decades. Fertility is still where the most serious errors are found. In general, he likes our national projections.

Murphy

There is still no complete explanation for the baby boom or the baby bust–we do not have a model that works.

Zlotnik

In the short term it is still important to look at immediate factors—-We should now use nativity in projections because it is currently an important variable in population growth.

Frejka

He discussed his method of calculating the percentage of births which are mistimed. He says Henshaw has promised to do some additional calculations to try to eliminate the effect of mistimed births. Henshaw also plans to take into account the educational factors which were discussed yesterday.

We need to focus on the most important leading indicators of fertility. This might be similar to economic indicators which predict reasonably well out 6 months or so. Frejka gives as an indicator example the cohort analysis that he and Morgan both did. The incomplete cohorts can serve as a leading indicator.

Bean

He asked the European fertility experts for their views on future prospects that European fertility might rise in the short to medium term.

George

The only evidence he sees for a rise in Canadian fertility is the birth expectations plus the end of timing delays. He uses current fertility levels as the middle assumption for 50 to 75 years which is long enough to see age structure effects.

Official migration levels are set by Canadian government policy for next 5 years or so. So that is what we have to use. Mortality is as important as fertility and one should make alternative assumptions. And if the projections are done often enough you can capture turning points.

Coleman

Our purpose here is not to discuss how to make projections. Aren't we really here to determine if U.S. fertility is peculiarly high and, if so, is it a precarious condition or a stable one?

His view is that European fertility will rise because of the end of postponement but also because expectations are for two kids still. This is stated repeatedly as a goal and he thinks society will adjust so women can achieve this two-child goal. Perhaps it is no surprise that the very rich and flexible United States is doing better right now at permitting women to achieve their goals. Maybe it is Europe which is out of step. He thinks no one else agrees with him. In Brazil at the International Union for the Scientific Study of Population meetings, the debate was confined to how far below two children would European countries end up—a lot or a little And it does matter. By the way, the participants are all hoping that the IUSSP will publish those Brazilian debates.

We have learned two things at this conference. First, we should analyze subgroups within the population in order to find explanations. Second, what matters is the fertility behavior of the core populations of each country—that requires the use of the kinds of correction techniques that Bongaarts advocates.

Even a large paper like Frejka's did not include any analysis of female labor force participation or gender equity.

He is curious about why Canada has such a low TFR(1.6) and New Zealand has such a high TFR(2.1)—These societies are not that dissimilar from the United States, so why are their fertility responses so different.

Van de Kaa

How valid is it to even compare Europe with United States? The U.S. has had significant migration throughout its history. It is just beginning to occur in Europe. European countries don't know whether they are now immigrant countries, or are they countries that admit people when they cannot refuse them entry?

It is also true in Europe that discussions of this type begin with the assumption of perfect contraceptive use so that all kids are seen as wanted—but the United States does not have perfect contraception. Achieving it might eliminate TFR differences with Europe.

In Germany we see society splitting into family oriented and childless couples. In Southern Europe it is hard to have kids outside of marriage so they delay marriage and remain childless. When they do marry, they have to have a kid. But they often stop at one. People want to have two kids, but there are competing priorities and the longer they postpone the less likely they get to two.

There are already declining national populations in 12 to 15 countries. He expects this to motivate nationalistic governments to promote childbearing and the value of children.

Look at the values of trend setters. Surveying their opinions will help forecasters. Look for changes in value orientation among the young, well-educated, socially mobile, secular group.

He expects U.S. fertility to remain an outlier because the United States is an outlier in many other ways. Look at the belief in God in the United States for instance.

Haub

Nobody cares about the historical failure to predict the baby boom. He agrees with Dirk that the United States is different and he things we have gone through a lot of change. He expects the United States to change relatively little.

He is surprised that no one has mentioned the rise in Sweden in the early 1990s which was strictly the result of a change in one law which made it advantageous to have a second birth within 24 months of the first.

He suggests that fertility in much of the developed world has found its equilibrium level in many places. Only economic fluctuations will affect it in the future.

Murphy

Maybe these numbers are not that important. Britain has had below replacement fertility for 30 years. And it is projected to be below replacement for the next 30. Yet, in only one of those 60 years is natural growth expected to be negative—even with relatively low levels of migration.

There is a huge range of variation in the desired number of children. Why do people make such a wide variety of individual choices?

One major difference with the United States is that European migrants are from Eastern Europe, whereas the U.S. migrants are largely from Mexico and Asia.

He agrees with Dirk about the existence of structural factors which inhibit fertility in parts of Europe. He sees more stability now—divorce and the age of childbearing are not increasing. He expects fertility to go up with the end of postponement, but he doesn't expect TFR levels much above 1.8. He notes that this level was predicted 30 years ago at a meeting of the Royal Society in London. So occasionally demographers do get it right.

Bean

The pattern of wage and income inequalities that emerged in the 1970s in the U.S. may have been only partially alleviated during the 1990s. The bottom part of the income distribution may have had actual declines in real purchasing power. Can we conjecture about the fertility effects of these changes?

Also, there have been substantial increases in housing prices which have outstripped increases in earnings. Do we see or expect to see fertility effects?

Morgan

He argues the whole idea of leading indicators comes from an over-determined view of social change and that it would be a big mistake to look for them.

He also commented that those who work with confidence intervals produce ranges which are much greater than any of us thinking on a substantive basis would have anticipated. As an example, he mentions a recent presentation where the confidence intervals in 2050 were between zero and six kids.

McDonald

He disagrees with Morgan. Australian fertility has been experiencing a linear decline for the last 10 years, so you might expect that trend to continue. But he found the age at marriage stopped rising 5 years ago—and this year fertility seems to have stopped declining. He thinks it is possible that age at marriage can serve as a leading indicator.

His main point is the need to remember the heterogeneity of the U.S. population by many different characteristics. He says this is very true in the U.S. parity data.

Frejka

Coleman is wrong about the possibility of fertility rising in the United States as the societies become more open, etc. The fact is that U.S. fertility is at 2.0 because of contraceptive failure, not because of the U.S. social structure.

He thinks leading indicators can be useful, and not necessarily deterministic—just part of human behavior—-like marriage.

Day

He quite agrees with Morgan that leading indicators don't do much. He described research on many countries where fertility and mortality were substantially controlled. The theory was that fertility would be low in those countries where you couldn't afford failure or where life was so good that kids served no economic purpose. He got no results at all.

He agrees with Dirk about the importance of values and changes in values. It is just hard to measure these attitudes.

What do governments want out of censuses? They need data to plan society and measure change, especially if they are welfare states. We need to be sure we keep getting the basic data. He doesn't like how the 2000 Census was treated in Congress.

Schoen

It is possible European fertility will rise if they start to get immigrants from high-fertility areas–like Africa and Asia.

To illustrate the increasing heterogeneity of the U.S. population—he noticed from Frejka's work that about one in 4 of women born during the 1960s are childless. The differences between those who have no children and those who do could become an important social issue.

Henshaw

He agrees with Frejka that unintended pregnancies have had a big impact on U.S. fertility rates. He expects that technological improvements in fertility control over the next 25 years may lead to lower fertility.

His comment on the idea that Americans are so rich that they can afford to have two children—-Even with such resources, you need to find the right partner. He thinks such issues plus divorce help explain why women want two, but average 1.8.

Where Do We Go From Here?

Hollmann, *U.S. Census Bureau*

The next set of U.S. projections must address two big issues. One major issue is the OMB-mandated race expansion. There are all kinds of unsettled issues which revolve around what categories to project and how to display the information. He thinks we will need to create 62 separate groups, but not all will be done with the same degree of elegance.

The second major issue involves our historical reluctance to do stochastic population projections, usually for some of the reasons alluded to by Morgan. In spite of these concerns, we have decided to make an effort to more formally express our uncertainty about our results.

He agreed that it was important to consider heterogeneity—especially in factors such as nativity, unintended pregnancy, and education.

In the past, we have tended to develop population categories which were useful to our customers. Perhaps we need to emphasize what categories are actually useful in explaining population growth?

In the past, we have usually treated the racial categories as closed populations where the race of the child was always the race of the mother. This is no longer tenable with these mixed-race categories.

He hopes to be able to use Mulder's results for historical errors to help develop certainty ranges in our next set of projections.

Because of the Kingkade/Frejka work to assemble a historical data base, Hollmann expects that we will be taking cohort fertility levels, tempo effects, and so on into consideration.

He thinks we have a wonderful opportunity to evaluate again the birth expectations data which we did use in our most recent projections. This is due to both Morgan and Frejka's papers.

Floor Discussion

Fullerton

He suggests we focus on four or five important sub-groups of the population—think about those very hard, and then derive the overall fertility from these group's shares.

We are becoming less and less "European" every day, so maybe we are looking at the wrong reference group.

Murphy

Making the race predictions will take a lot of effort because you need to tab the new data by race of child, father, and mother. This must be done soon.

Hollmann

He agreed with Murphy and estimated that the total possible combinations of two parents with a child was something like 238,000 cells. If we eliminate the ridiculous, combinations we might get down to perhaps 90,000.

Chamie

Whatever you do should be extremely transparent. This includes description of your goals and your problems.

Explanations should be simple as you are conveying this information to masses of people with different interests. What is the central message you wish to get out?

The image of the Census Bureau must be improved—it must be seen as very objective with no biases.

Hollmann

In response to Chamie, the Census Bureau is trying to improve the quality of the delivery system and the clarity of the product via use of the Internet.

On the other hand, we have largely abandoned the carefully crafted paper reports we used to do. Those were aimed at a pretty generalist audience. We struggle with timeliness versus explication.

McDonald

You might contact Statistics New Zealand about the issues surrounding multiple race information. They only have three groups, but even then their experience has been rather negative. They have had a hard time joining the census with the administrative data sets. This is probably very similar to the problems we would face.

Hollmann

I agree. We already have some such problems as when we try to explain to some people why American Indians apparently live longer than almost any other race group. It is difficult to explain that more people apparently mark themselves as "American Indian" in the census than are so marked on death certificates.

To make matters worse, we already know that other Federal agencies won't even start collecting the race data in the new format until 2003. That means it won't be available until 2005 or so.

Hodges

He argues for doing the so-called naive projections because most of what we have heard here suggests that these are actually similar to the most well-informed assumptions.

He reminds us that the categories of data we produce are as important as the methodology or accuracy.

Haub

African-American fertility convergence seems to be happening. Perhaps this will happen in other groups as they become more assimilated?

George

Is the Census Bureau planning to do projections using the cohort method? Or are you going to analyze the cohort data to develop period assumptions?

In Canada the race projections are done after the age and sex projections have been completed. He feels this procedure yields fewer complications. Do you think this could be done in the U.S.?

Hollmann

In response to George's first question, I expect we will continue to directly use the cohort-component method.

As far as race projections, he assumes they will be integrated with the projection of the other characteristics. Customers then know that they all came from a unified set of assumptions and consistent methodology.

Van de Kaa

His recommendation would be not to forget long-term trends that we have seen and the surprises that we have experienced in that respect. For example, when the first demographic transition began, the assumption was that fertility would stabilize nicely around the replacement level. We now know in Europe that this is not true and that we are in a whole new ballgame.

We have learned that the native-born White population has a TFR as low as that in Europe. I suggest the other race/ethnic groups might eventually follow exactly the same pattern. It may take a century.

Keilman

The real question is: Will American fertility come down to the Western Europe average of 1.5 or will it stay at 1.9? I think the answer has been given a number of times by the audience and that answer is 1.9.

Summary

Gregory Spencer

In retrospect, one of the most memorable aspects of this conference was that it took place only about 3 weeks after the terrorist events of September 11, 2001. Most participants' travel plans had to be redone because of airport closures, but there was never any movement toward cancelling the meeting. Overall, perhaps a dozen invitees declined to attend; most often people who did not feel comfortable flying to Washington, D.C.. But the vast majority still came and participated. It is interesting to note that there was scarcely any acknowledgment given the terrorist activities in these proceedings. I can't find a single mention in the transcripts. I wish I could say this was a statement about our devotion to the conference topic. I think it is more likely that we were all grateful for something familiar to argue about or debate, hoping that the larger world would regain its normalcy in the meantime.

What Will Happen to American Fertility?

This really was the critical point of the conference. We were interested in finding out if there was a consensus that the United States would eventually decline to "European-like" levels (let's suppose that means a total fertility rate below 1.5 or so). Many participants gave several different opinions during the meetings. Some declined to give an opinion. Those that I can find in the proceedings with an opinion generally belong to that group who prefer to use the word "stable" as opposed to those who emphasized the possibility of significant decline. First, however, I want to acknowledge Ahlburg and Alcantara as the only two participants who were willing to say that they expect American fertility to rise. Van de Kaa did state that he expected U.S. fertility to always remain an "outlier." But it was not clear what that meant for future fertility levels. It was far more common to find the opinion that American fertility would remain stable or decline slightly (Frejka, Schoen, Chamie, Goldstein, Keilman, Haub). Schoen, Coleman, and Bongaarts make the stronger argument that "Americans say they want 2 children and that is what they have. This is not a coincidence." Coleman added that it was no surprise that the rich United States with its flexible culture has enabled women to achieve their family size preferences. [He did not believe the United States will decline to European levels because the Europeans have had huge timing postponements, high youthful unemployment, continued rigid sex roles, and emphasis on the importance of families. None of that has happened or is true in America. He suggests that in the United States people have more space and more religion. Fertility might even rise if there is a renewal of family values, the end of delayed childbearing, or if new reproductive technology is introduced.] O'Connell was not sure that the achievement of the two-child family wasn't just a numerical coincidence. He noted that the birth expectations in America have been stuck on two children ever since they started surveying birth expectations back in the 1970s. Van de Kaa argued for a substantial downturn in future U.S. fertility as the young see the advantages of remaining childless for a while.

As I said, there was no general agreement on the future course of American fertility. Many experts do not believe the United States will follow the European pattern. As Fullerton observed, "America is becoming less like Europe every day. Why do we think it should serve as a model?"

Seeing as there was little agreement on whether American fertility would remain stable or decline, I thought I would at least create a brief list of the factors that influence American fertility. Many contributing factors were put forth during the course of the conference. Here I have attempted to summarize the major ideas, but clearly my listing is not exhaustive. It is also possible that I have misunderstood someone's point or inadvertently misrepresented their argument. Of course I am only able to list those people who directly mentioned each concept during the floor discussion. There are undoubtedly other participants who share the same ideas.

What Factors Contribute to American Fertility Being Relatively High?

I tried to group these ideas/statements/theses into extremely broad general categories, which I call social factors, economic factors, and biological factors. Many of these ideas could legitimately be placed in several different categories.

Social factors:

Religiosity—Several demographers singled the United States out as an unusually religious society (Frejka, Henshaw, Coleman, and Alcantara)

The United States is a more child-friendly society where childbearing is encouraged (McDonald)

Perhaps Americans have a more immature attitude about sex, which leads to more contraceptive failure (Henshaw)

There was a bigger "baby boom" in post-WWII USA than in Europe so bigger families are more acceptable in the United States (Murphy)

Inadequate education in the United States increases unwanted fertility in particular (Frejka)

Keilman says the degree of socialization is important (immigrants at younger ages have lower family sizes than do immigrants who are older on arrival)

Marriage as an institution may be more durable in the United States (Schoen)

Equality of the sexes in the United States facilitated childbearing (implying that men did more of the work) (Frejka, Schoen, Coleman, McDonald)

Economic factors:

More space in America makes it easier to have children (Goldstein)

Lower taxes than in Europe lessen the costs of childbearing here (Coleman and McDonald)

Childcare is more affordable in the United States because there is a larger wealth differential among population groups (Coleman and McDonald)

There is more financial support for children from government than commonly supposed (McDonald)

America has lower housing costs and this facilitates childbearing (McDonald)

Biological factors:

Immigrants, Ethnicity, Nativity, and Race. These terms were used in overlapping fashion by Haub, Rao, Frejka, Bean, Schoen, Rao, Goldstein, George, and McDonald. All saw these as important contributors to the comparatively high U.S. fertility.

Higher contraceptive failure rates lead to higher U.S. fertility (Henshaw, Frejka and Van de Kaa)

There is a significantly earlier start to childbearing in the United States (Murphy)

Are There Leading Indicators of Future Trends in Fertility?

Another topic of discussion was whether or not there were "leading indicators" that could help one predict the future course of American fertility. I should note that some (Morgan, Ahlburg in particular) thought the whole idea of leading indicators was a dead end. Macdonald and Frejka disagreed.

Geographic variation was mentioned as an predictive factor, but there was no agreement on the direction of any effect (Frejka, McDonald, Wetrogan)

Frejka gave as an example the type of analysis done by him and Morgan, arguing that analysis of incomplete cohorts can guide us.

Any renewal of family values would tend to boost fertility (Goldstein)

An actual measurable end to delayed childbearing would imply a future fertility decline (Bean, Goldstein)

The introduction of new reproductive technology might reduce fertility (Goldstein)

Any change in the values of "trendsetters" will change fertility, but in unpredictable directions (Van de Kaa, Frejka)

Welfare reform might increase fertility (Henshaw)

Changes in unemployment/economic growth should immediately affect fertility (Bongaarts, Haub, McDonald)

What will happen to teenager's fertility and first births are critical to future childbearing (McDonald)

Restrictions on access to abortion would increase fertility (Day). Henshaw says this is unlikely.

A rise in real housing prices should depress fertility (Bean)

There was also a fairly spirited discussion of the utility of birth expectations in this regard. Masnick asked if there was any expectations data from European countries before fertility fell. Have expectations changed since then? Bean argued that the errors in expectations by parity tend to offset. Schoen agreed but argued "what do we have that is any better?" Morgan and O'Connell said the errors are so big that the data can't be any good. Murphy said they were just extrapolations of past trends and not predictive of future shifts. There was no consensus.

Outline of Current U.S. Fertility Projections Research

The above notes are illustrative of the precept that "reasonable people can disagree." We all reviewed the same papers and listened to the same arguments. There is no evidence that anyone's mind was changed during the meetings. We did learn that there are many, many possible forces operating in American society that may influence American fertility. What was missing was the ability to rank these influences in importance. At this stage, we were only able to describe the possible influences of these factors. I think this is a valuable beginning.

The Census Bureau has begun work on a new set of national population projections. In closing, here are the initiatives we are currently undertaking (as of June 2002):

1) Projections will be based on Census 2000 enumeration
2) Projections will incorporate a probabilistic treatment of the components of change
3) Projections will be stochastic rather than scenarios. We may run 10,000 versions.
4) Projections will incorporate the new OMB race classification system.

At this time, it is not clear what changes may be made in our fertility assumptions in these new projections. It is not obvious that we need to make any significant changes from the current assumptions.

COPAFS
1249 Duke Street, Suite 402
Alexandria, Virginia 22314
703-836-0404

ISBN 0-9720596-0-1