FINAL REPORT

Long Term Model Development for 
Social Security Policy Analysis

By

Eric Toder, Melissa Favreault, John O'Hare, Diane Rogers, 
Frank Sammartino, and Karen Smith 
Urban Institute

and

Kent Smetters of the University of Pennsylvania, 
and 
John Rust of Yale University.

The Urban Institute 
2100 M Street N.W. 
Washington, D.C. 20037

January 31, 2000

This research was funded by the Social Security Administration, Office of Research, 
Evaluation and Statistics, Division of Policy Evaluation (Contract No.: 600-98-27332). We remain solely responsible for all errors and omissions. The nonpartisan Urban Institute publishes studies, reports, and books on timely topics worthy of public consideration. The views expressed are those of the authors and should not be attributed to the Urban Institute, its trustees, or its funders.
Abstract

Long-Term Model Development for Social Security Policy Analysis

Policymakers need to understand how Social Security reforms affect income distribution. Existing models range from simple representations of career earnings of typical workers to complex general equilibrium models. Population micro-simulation models, which project the earnings, wealth, and demographic histories of a representative sample of families, are useful for simulating many reform proposals. This report evaluates one such model – the projected cohorts model - and then discusses in detail three important issues in model development: 1) representing saving behavior, 2) capturing macro-economic effects, and 3) accounting for risk and uncertainty.
PREFACE

This report was prepared by the Urban Institute for the Social Security Administration, Office of Research, Evaluation, and Statistics, Division of Policy Evaluation (SSA/ORES/DPE) under Task Order 0440-99-36156, pursuant to Contract No. 600-96-27332.

The report was drafted by a number of authors, both Urban Institute staff and project consultants. Eric Toder of the Urban Institute directed the day to day research on the project and contributed to writing several of the chapters. Principal authors of each chapter in the report were:

Chapter 1 Melissa Favreault, Urban Institute
Chapter 2 Frank Sammartino and Karen Smith, Urban Institute
Chapter 3 John O’Hare, Urban Institute
Chapter 4 Diane Lim Rogers, Urban Institute and Kent Smetters,
   University of Pennsylvania
Chapter 5 John Rust, Yale University

Joseph Cordes, Daniel Dulitzky, and Frank Sammartino also contributed to the writing of chapter 3. Robbie Howell provided research assistance.

In preparing this report, we received numerous helpful comments on earlier drafts from David Pattison and Benjamin Bridges of the Social Security Administration and from a panel of outside experts. The outside experts were Jonathan Gruber of the Massachusetts Institute of Technology, Eric Hanushek of the University of Rochester, Olivia Mitchell of the University of Pennsylvania, John Shoven of Stanford University, and Douglas Wolf of Syracuse University. We also received helpful comments and suggestions from Rudy Penner, Lawrence Thompson, and Sheila Zedlewski of the Urban Institute and Gary Burtless of the Brookings Institution.

Theresa Plummer prepared the manuscript for publication.
# TABLE OF CONTENTS

PREFACE .................................................................................................................................................. I

INTRODUCTION AND SUMMARY ........................................................................................................... 1

I. OUTLINE OF REPORT ............................................................................................................................ 1

II. MAIN FINDINGS .................................................................................................................................... 3

1. Microsimulation Models. .......................................................................................................................... 3
2. The Projected Cohorts Model. ................................................................................................................ 5
3. Modeling Saving. ..................................................................................................................................... 6
5. Risk and Uncertainty. ............................................................................................................................. 8

REFERENCES TO INTRODUCTION AND SUMMARY ............................................................................ 11

CHAPTER 1: REVIEW OF SOCIAL SECURITY MICROSIMULATION ALTERNATIVES ......................... 12

I. INTRODUCTION ..................................................................................................................................... 12

II. WHY MODELS ARE IMPORTANT AND CLASSES OF MODELS FROM WHICH ANALYSTS CAN CHOOSE .......................................................................................................................... 13

III. DESIRED CHARACTERISTICS OF A MODEL OF SOCIAL SECURITY ........................................ 15

1. Capacity: Ability to Capture Proposed Reforms .................................................................................. 15
2. Outcomes ............................................................................................................................................. 22
3. Population Definition/Aging Modules ................................................................................................. 24
4. Time Horizon ....................................................................................................................................... 25
5. Integration of Social and Economic Research on Social Security Parameters ................................. 26

IV. ASSESSMENT OF DIFFERENT MODEL TYPES .................................................................................. 28

1. Criteria ................................................................................................................................................ 29
   - Representativeness and Completeness of Starting Data .................................................................. 30
   - Cohort Size (Aside From Stochastic Replications) ......................................................................... 30
   - Generalizability ................................................................................................................................. 30
   - Ability to Oversample Parts of The Population ............................................................................. 30
   - Theoretical and Empirical Validity of Updating Algorithms (If applicable). ............................... 30
   - Autocorrelation/Serial Correlation (If applicable) ....................................................................... 31
   - Feedbacks ....................................................................................................................................... 31
   - Outcomes ....................................................................................................................................... 31
   - Programmatic Detail ......................................................................................................................... 31
   - Projection Assumptions .................................................................................................................... 32
   - Sensitivity Analyses ......................................................................................................................... 32
   - Stochastic Replications Per Individual/Cell ..................................................................................... 32
<table>
<thead>
<tr>
<th>V. WHICH MODEL WHEN?</th>
<th>61</th>
</tr>
</thead>
<tbody>
<tr>
<td>VI. CONCLUSIONS</td>
<td>63</td>
</tr>
<tr>
<td>REFERENCES TO CHAPTER 1</td>
<td>6</td>
</tr>
</tbody>
</table>

**CHAPTER 2: REVIEW AND ASSESSMENT OF THE COHORTS SIMULATION APPROACH**

<table>
<thead>
<tr>
<th>I. INTRODUCTION</th>
<th>76</th>
</tr>
</thead>
<tbody>
<tr>
<td>II. AN OVERVIEW OF THE COHORTS MODEL</td>
<td>77</td>
</tr>
<tr>
<td>1. Applicability of The Cohorts Model</td>
<td>77</td>
</tr>
<tr>
<td>2. Types of Analyses Beyond The Current Scope Of The Model</td>
<td>78</td>
</tr>
<tr>
<td>Economic Well-Being of the Retired Population</td>
<td>78</td>
</tr>
<tr>
<td>Program Interaction</td>
<td>79</td>
</tr>
<tr>
<td>Post Retirement Earnings</td>
<td>79</td>
</tr>
<tr>
<td>Social Security Trust Funds Balances</td>
<td>80</td>
</tr>
<tr>
<td>Macroeconomic Effects of Social Security Reforms</td>
<td>80</td>
</tr>
<tr>
<td>III. KEY ISSUES FOR THE PROJECTED COHORTS MODEL: MISSING DATA AND EARNINGS PROJECTIONS</td>
<td>80</td>
</tr>
<tr>
<td>1. Starting Database</td>
<td>81</td>
</tr>
<tr>
<td>Incomplete and Underreported Survey Data From the 1994 CPS</td>
<td>81</td>
</tr>
<tr>
<td>Missing SSA Administrative Data</td>
<td>82</td>
</tr>
<tr>
<td>Missing Earnings Information For Deceased Workers</td>
<td>82</td>
</tr>
<tr>
<td>Missing Earnings Information For Former Spouses</td>
<td>83</td>
</tr>
<tr>
<td>Earnings Limited to Social Security Covered Earnings</td>
<td>83</td>
</tr>
<tr>
<td>Censored Earnings</td>
<td>83</td>
</tr>
</tbody>
</table>
2. Calculation of Expected Lifetime Social Security Benefits .......................................................... 85
   Differential Mortality. .................................................................................................................. 85
   Taxation of Benefits.................................................................................................................. 85
3. Labor-Force Participation and Earnings Projections................................................................. 85
4. Changes in Labor Force Participation and Earnings Patterns ................................................. 86
   Increased Labor Force Participation of Women. ........................................................................ 86
   Decreased Labor Force Participation of Older Men. ............................................................... 88
   Increased Relative Earnings of Women. ...................................................................................... 89
   Decreased Relative Earnings of Younger Men. ....................................................................... 89
   Increased Share of Wives Earnings........................................................................................... 91
   Increased Correlation of Spouses' Earnings. .............................................................................. 91
   Increased Earnings Inequality..................................................................................................... 94
   Changing Macroeconomic Environment. .................................................................................. 97
5. Alternative Methods For Projecting Earnings ......................................................................... 97
6. Demographic Projections .......................................................................................................... 98

IV. THE PROJECTED COHORTS MODEL AS A FRAMEWORK FOR A MORE COMPREHENSIVE
   MODEL .......................................................................................................................................... 101

   1. Behavioral Response at the Micro Level ................................................................................ 101
   2. Macroeconomic Feedback Effects ....................................................................................... 102

V. SUMMARY/CONCLUSIONS ........................................................................................................ 103

REFERENCES TO CHAPTER 2 ...................................................................................................... 107

CHAPTER 3: MODELING SAVING BEHAVIOR ............................................................................. 110

I. INTRODUCTION .......................................................................................................................... 110

II. ISSUES IN MODELING SAVING BEHAVIOR ........................................................................ 111

   1. What Questions Do We Want to Answer? ........................................................................... 111
   2. What Definition of Saving Should We Use? ......................................................................... 112
   3. Stylized Facts With Respect to Saving ................................................................................. 112
      The U.S. Private Saving Rate Has Declined ........................................................................... 113
      Saving Rates Vary By Age, But Have Declined for All Groups ............................................. 114
      Saving Rates are Higher for Higher Income Households ...................................................... 114
      Many Individuals Reach Retirement with Little Savings ...................................................... 115
      Social Security and Pension Assets Constitute Most of Retiree Wealth ................................ 115
      The Elderly Spend-Down Only Some of Their Wealth ......................................................... 115

III. OVERVIEW OF THE SAVING AND CONSUMPTION LITERATURE ................................... 117

   1. Life-Cycle Models 117
   2. Other Models of Saving ......................................................................................................... 120
      Hyperbolic Discounting .......................................................................................................... 120
      Behavioral Life-Cycle Model ................................................................................................ 121

IV. APPROACHES TO MODELING SAVING BEHAVIOR ............................................................ 123
Table of Contents (Continued) January 2000

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Dynamic Programming (DP) Methods</td>
<td>124</td>
</tr>
<tr>
<td>2. Estimation and Projection of Wealth at Retirement</td>
<td>125</td>
</tr>
<tr>
<td>3. Projection of Wealth Based on Econometric Estimates of Annual</td>
<td>127</td>
</tr>
<tr>
<td>Consumption</td>
<td></td>
</tr>
<tr>
<td>Consumption a Fixed Fraction of Current Income</td>
<td>128</td>
</tr>
<tr>
<td>Consumption as a Function of Permanent Income</td>
<td>129</td>
</tr>
<tr>
<td>Consumption as a Function of Cash-on-Hand</td>
<td>130</td>
</tr>
<tr>
<td>4. Issues That Models of Wealth Accumulation Must Address</td>
<td>130</td>
</tr>
<tr>
<td>Quality of Data</td>
<td>130</td>
</tr>
<tr>
<td>Treatment of Uncertainty</td>
<td>130</td>
</tr>
<tr>
<td>5. Saving and Wealth Accumulation in Existing Simulation Models</td>
<td>131</td>
</tr>
<tr>
<td>MINT</td>
<td>132</td>
</tr>
<tr>
<td>CORSIM</td>
<td>133</td>
</tr>
<tr>
<td>DYNASIM</td>
<td>134</td>
</tr>
<tr>
<td>PRISM</td>
<td>134</td>
</tr>
<tr>
<td>DYNACAN</td>
<td>135</td>
</tr>
<tr>
<td>EBRI/SSASIM2</td>
<td>135</td>
</tr>
<tr>
<td>Auerbach-Kotlikoff Overlapping Generations Model</td>
<td>135</td>
</tr>
<tr>
<td>V. SAVING AND OTHER BEHAVIORAL DECISIONS</td>
<td>136</td>
</tr>
<tr>
<td>1. Retirement Decision</td>
<td>136</td>
</tr>
<tr>
<td>Estimates of the Retirement Decision</td>
<td>137</td>
</tr>
<tr>
<td>Joint Retirement Decisions of Spouses</td>
<td>137</td>
</tr>
<tr>
<td>2. Bequests</td>
<td>139</td>
</tr>
<tr>
<td>3. Other Sources of Income</td>
<td>140</td>
</tr>
<tr>
<td>4. Portfolio Choice</td>
<td>140</td>
</tr>
<tr>
<td>VI. CONCLUSIONS</td>
<td>141</td>
</tr>
<tr>
<td>REFERENCES TO CHAPTER 3</td>
<td>143</td>
</tr>
</tbody>
</table>

CHAPTER 4: MACROECONOMIC GROWTH MODELING FOR MICROSIMULATION OF SOCIAL SECURITY REFORM.................................................148

I. INTRODUCTION ..................................................................................................................148

II. UTILITY-BASED MODELS (EXPLICIT HOUSEHOLD OPTIMIZATION) ..................151

1. Ramsey Model ..............................................................................................................151
2. The Life Cycle Overlapping-Generations Model .................................................152
3. Deterministic OLG Models.....................................................................................152
4. Uncertainty in OLG Models....................................................................................153
5. Making Utility-Based Models More Realistic ......................................................155

III. AD HOC MODELS (NON-EXPLICIT SPECIFICATIONS OF HOUSEHOLD BEHAVIOR) ........156

1. Advantages and Disadvantages of Ad Hoc Specifications ................................156
2. Overview of Ad Hoc Modeling Approaches and Examples ..................................158
3. Macroeconomic Modeling with Microdata .............................................................160

IV. HOW MODELS ADDRESS PARTICULAR ISSUES ..........................................................160
CHAPTER 5: STRATEGIES FOR INCORPORATING RISK, UNCERTAINTY, AND PRIVATE INSURANCE MECHANISMS IN MODELS OF SOCIAL INSURANCE

I. INTRODUCTION

II. THE IMPORTANCE OF ACCOUNTING FOR RISK AND BEHAVIORAL/EQUILIBRIUM FEEDBACKS

1. Some Preliminary Concerns
   Risk vs. Uncertainty
   Modeling Priorities: Risk vs. General Equilibrium Effects

2. Examples Illustrating the Need to Account for Risk and General Equilibrium Feedbacks
   Example 1: Projecting the Financial Impacts of an Aging Population
   Example 2: The Value of Social Security in a World of Incomplete Annuity Markets
   Example 3: Social Security’s Effect on Intra and Intergenerational Risk-Sharing
   Example 4: Social Security’s Effect on Markets and Institutions

3. Examples that Illustrate the Benefits of a Comprehensive Treatment of Risk in Social Security Policy Models
   Example 1: Interaction Between Old Age and Medicare Insurance
   Example 2: Interactions Between Old Age and Disability Insurance
INTRODUCTION AND SUMMARY

The Social Security Administration (SSA) is undertaking a major effort to improve its capability to evaluate the distributional effects of Social Security reforms. As part of this effort, SSA has supported the development of two large-scale micro-simulation models – the Model of Income in the Near-Term (MINT) and the Projected Cohorts Model (PCM). In addition, SSA is considering for internal use the Cornell Microsimulation Model (CORSIM), a model developed by a team of economists and sociologists at Cornell University to analyze the effects of policy reforms. These new models supplement the models used by the Office of the Chief Actuary (OCT) by adding much more detail on how Social Security affects different sub-groups of the population.

This report provides a review of the major issues in developing models to analyze the effects of reforms of the Social Security system on the distribution of income within and between cohorts. We consider the ability of models both to project accurately how reforms affect economic variables and to evaluate the net social benefits of alternative outcomes. While the emphasis is on micro-simulation models that can provide detailed projections of the economic status of sub-groups of the population under current law and proposed alternatives, we also consider alternative approaches for modeling the long-term macro-economic effects of policy changes. We discuss the strengths and weaknesses of alternative modeling strategies, in relation to the information policy-makers need to evaluate alternative reforms. In doing so, the report considers in detail some of the most complex and difficult issues in using models to analyze the effect of policy reforms. These include the choice of how to model household saving and alternative ways of incorporating risk and uncertainty in Social Security policy analysis.

I. OUTLINE OF REPORT

Chapter 1 of the report provides a detailed review of alternative models used to analyze Social Security policies. We explore the relationship between the options that policymakers want to analyze and the capabilities of alternative types of models to answer different questions. We list detailed criteria for evaluating models and then assess the strengths and weaknesses according to these criteria of four general model types – representative worker models, cell-based models, population microsimulation models, and computable general equilibrium models. The chapter concludes with a general discussion of which types of model are appropriate under which circumstances.

Chapter 2 reviews and assesses one population microsimulation model currently under development at the Social Security Administration – the Projected Cohorts Model (PCM). The general approach of the cohorts model is to begin with a complete earnings history of a recent cohort of retirees (the 1930 birth cohort). It then projects characteristics of future cohorts by adjusting earnings to reflect earnings growth and changes in labor force participation rates and relative earnings of women. The chapter provides an overview of the PCM, summarizing the
types of analyses it can currently perform and those it cannot perform without further enhancements. It then considers problems and issues that the cohort model approach must address. These issues include how to correct for gaps in the current database, improving the calculation of lifetime Social Security benefits, and adjusting labor-force participation rates and earnings patterns to take account of observed changes for more recent cohorts. We conclude the chapter by discussing the strengths and weaknesses of using the cohorts model approach as a platform for a more comprehensive model that would incorporate behavioral responses and macro-economic effects.

Chapter 3 discusses issues in modeling household saving in microsimulation models. We begin by reviewing some stylized facts about saving that any model must explain and then discuss the main approaches to modeling saving in the academic literature. The central approach is the life-cycle model (LCM), which offers a simple and powerful framework to explain how households allocate consumption over their lifetime. We review empirical research that contradicts the implications of the simple form of the LCM and that either suggests ways of modifying the LCM to improve its explanatory power or offers alternative modeling approaches. We then outline broad approaches for modeling saving in microsimulation models and review the treatment of saving in current models. We conclude with a brief discussion of how to integrate modeling of saving with modeling of other behavior, especially the decision of when to retire.

Chapter 4 reviews approaches for modeling the effects of Social Security reforms on the macro-economy and for linking macro-economic models to microsimulation models that can be used to analyze income distribution. We discuss two broad approaches to macro-economic modeling – utility-based models that represent household labor supply and saving as a solution to a lifetime optimization problem and ad hoc models that impose a less formal structure. These latter models rely more on econometric analysis from historical data to derive behavioral equations for labor supply and saving. We consider the strengths and weakness of the two alternative approaches. We then discuss alternative ways of linking macro and micro models. The chapter then reviews specific features of existing and potential macro models for Social Security policy analysis, including the level of aggregation at the producer and consumer levels and how the models treat household saving and labor supply, the government sector, investment adjustment costs, and international capital flows. We conclude the chapter by exploring the alternative ways that utility-based and ad hoc models would approach key policy questions, such as the effects of cuts in Social Security benefits, mandatory private accounts, and a decision to invest the trust fund surplus in equities.

Chapter 5 discusses strategies for incorporating risk, uncertainty, and private insurance mechanisms in models of Social Security reform. The chapter begins with illustrations of why it is necessary to account for the effects of policy changes on risk-bearing and how failing to do so can lead to misleading implications on the net gains and losses from policy reforms. It then discusses the types of risks people face and the mechanisms people use (private and public) to
reduce risks and explains why models of Social Security should ideally be able to address a wide variety of risks in a comprehensive way.

The final sections of Chapter 5 review different methods of incorporating risk and insurance in policy models. Because formal modeling of risk is extremely complex, we provide illustrations of “short cut” approaches to modeling risk, noting their potential applicability and possible pitfalls of some simplified approaches. We then survey the recent literature on dynamic programming/stochastic general equilibrium (DP/SGE) models, showing how they have been used to generate analyses of the welfare gains and losses from the current Social Security system and from transitions to alternatives, such as a system of individual accounts.

The report is intended to provide information to the SSA on the wide range of policy models available and on the advantages and disadvantages of alternative approaches to model development. The Urban Institute has not recommended an overall strategy for model development, although we hope that the information in this report will help SSA in deciding how best to proceed with their own model development. We hope also that the report will inform others on the “state of the art” in this rapidly expanding area of policy analysis.

An Appendix to Chapter 5 presents the views of Professor John Rust of Yale University, a consultant to this project and the author of chapter 5, on a strategy for SSA to expand model development. While the Urban Institute takes no position on the material in this Appendix, we believe the views of Professor Rust are worth serious consideration.

II. MAIN FINDINGS

1. Microsimulation Models.

A number of different types of models of varying complexity are available for analyzing the effects of Social Security reforms. They vary along many dimensions, including the extent to which they accurately represent the composition of the population and subgroups of interest (by sex, race, education, and income level), incorporate details of the Social Security retirement program, and allow for behavioral responses to policy changes. The simplest are representative worker models that can be used to simulate how Social Security reforms would affect the lifetime taxes and benefits of a worker with an assumed “representative” lifetime pattern of earnings. The most theoretically sophisticated are stochastic general equilibrium models that permit estimation of the economic effects of policy changes by representing household labor supply and saving behavior as the solution to a complex optimization problem, given household earnings capacity and initial wealth and government policies.

---

1 See, for example Citro and Hanushek (1997) and the Appendix to Chapter 5 of this report by John Rust.
Ideally, models should allow policy makers to estimate the impact of policy changes on income and lifetime welfare of different age cohorts of the population and sub-groups within them (by income level, race, sex, and level of education). They should be flexible enough to incorporate alternative assumptions about demographic changes and behavioral responses to policies, based on recent research findings. They should be sufficiently transparent so that users understand what specific assumptions are producing the model’s results. At a technical level, their computing cost should be low enough and computing speed fast enough to allow users to perform numerous simulations incorporating alternative scenarios and policy changes in response to requests for information from policymakers.

Our survey finds that no one model meets all concerns; all model types have both strengths and weaknesses. Representative worker models typically incorporate substantial details about the Social Security program. This enables them to be used to simulate the effects of different changes in payroll taxes (such as a change in the rate or the maximum earnings level) and the benefit formula (such as changes in the bend points in the replacement rate formula, cost-of-living adjustments, and spousal benefits.) But representative worker models incorporate little population detail, so they cannot be used to generate estimates of how policy changes affect income distribution. They also cannot be used to estimate overall effects of the fiscal balance of the Trust Fund or the state of the economy.

Cell-based models represent the population by age, sex, and cohort subgroups and incorporate average earnings profiles, Social Security benefit replacement rates, and tax rates in each population cell or subgroup. They are useful for tracking aggregates such as Trust Fund balances. But they are inadequate for studying the effects of Social Security proposals on detailed population sub-groups because they do not generate sufficient detail on the distribution of individual earnings histories.

Population Microsimulation models simulate the lifetime earnings, wealth, and demographic profiles of a representative sample of families. They include models such as the Model of Income in the Near Term (MINT) and Projected Cohorts Model (PCM) that project retirement incomes of sample individuals based on historical relationships, but do not incorporate a full year by year evolution of economic and demographic processes. They also include models such as the Cornell Microsimulation Model (CORSIM) and the Urban Institute’s Dynamic Simulation of Income (DYNASIM) model that simulate annual changes in demographic characteristics and economic outcomes for the sample population. The advantage of population microsimulation models is that they can capture the interactions of multiple program changes and social and economic processes. This enables users to simulate the effects on lifetime incomes and benefits of different cohorts and demographic groups of changes in Social Security rules. But the complexity of these models can make their results difficult to interpret and dependent on assumed future behavioral processes that may be improperly specified. While microsimulation models typically incorporate some behavioral responses based on statistical research results, they do not incorporate a theoretically consistent model of economic behavior. They also cannot be used to
generate overall economic forecasts without being linked to some type of external macroeconomic model that represents the production side of the economy and solves for market clearing wages and interest rates.

**General equilibrium models** provide a consistent behavioral framework for analyzing the effects of policy changes on labor supply, saving, total economic output, and social welfare. But these models are extremely complex and are often based on strong *a priori* assumptions about the structure of household preferences. To make them tractable, they typically represent the population and the parameters of taxes and benefit programs in a highly stylized fashion. This makes them less useful than dynamic microsimulation models in generating detailed distributional estimates and in analyzing changes in program details. But their advantage is they are only the modeling tool that provides new insights on the “big picture” effects of reforms on saving and work effort.

**The Projected Cohorts Model.**

The Projected Cohorts Model (PCM) is a population microsimulation model currently under development at SSA. As part of this contract, we were asked to evaluate this model in some detail.

The PCM uses as its base data a sample drawn from an exact match of the 1994 U.S. Census Current Population Survey (CPS) and Social Security Earnings Records (SSER) of individuals who became eligible for Social Security retirement benefits in 1992 and thus were born in 1930. Based on the experience of the 1930 birth cohort, it projects earnings histories of future cohorts by adjusting the earnings histories of the base year cohort. In its present form, it can be used to simulate the effects of changes in the Social Security benefit formula and payroll taxes on different groups of earners. Adding imputations of other sources of income (pensions and income from non-pension saving) and linking the PCM with a macro-economic model could potentially make the model usable for performing broader analyses of the impacts of policy reforms on the level and distribution of income of future cohorts of retirees.

The cohorts model is superior to representative worker models and cell-based models for analyzing the distributional effects of Social Security reform proposals, such as changes in the structure of the benefit formula. Its advantage is that it incorporates the full actual earnings and demographic history of a cohort. There are some gaps in the data because of missing observations and because earnings subject to Social Security are censored at levels that were much lower in relation to the average wage in earlier years than today. But SSA has developed a method for correcting censored earnings and imputations for missing earnings histories of deceased workers and former spouses could be added.

The main strength of the cohorts model is that it is able to simulate an entire lifetime of earnings for members of a cohort, using a simple computational process. Its main weakness it
that the simple adjustment it uses will not adequately project earnings histories of future cohorts if their lifetime earnings do not resemble earnings histories from the 1930 cohort. Although the model adjusts labor force participation rates and relative earnings of women to 1990 levels, it may be missing important changes that are already evident in earnings patterns for later cohorts. These include further increases in female labor force participation rates, especially for middle-aged and older women, a more positive correlation of earnings levels between spouses, and an increase in the inequality of the earnings distribution. The best way to fix these problems is to develop methods that incorporate information available in the partial histories of later cohorts.

The PCM in its present form cannot be used to analyze complete future total incomes because it does not include all income sources of beneficiaries and does not project marriage histories. In particular, it does not project income from pensions, other assets, and “partial retirement” earnings of Social Security beneficiaries. The PCM could be enhanced by modeling saving and retirement behavior and by including macro-economic feedback effects. The general approach used by the PCM has advantages for projecting saving because it simulates the entire earnings history of a worker. But simulating other behavioral responses (such as the timing of retirement) requires more complex algorithms that depart from the simple scaling adjustments of the cohorts model.

**Modeling Saving.**

The economics profession is far from a consensus on a theory of saving or on how to interpret results of econometric studies on how saving incentives affect net saving. Different approaches can yield very diverse predictions about how proposed policy reforms would affect household saving. Moreover, evaluation of policy reforms, such as a shift to individual accounts, is highly sensitive to assumptions about saving responses of households.

The most powerful and widely used theory of saving is the life cycle model (LCM). The basic premise of the LCM is that individuals are forward looking in planning to save to finance the decline in earnings that accompanies old age. The simple version of the LCM ignores uncertainty and assumes that utility is time separable and depends only on consumption and leisure in different periods. It also assumes that individuals have unlimited access to credit markets at the market interest rate and neither leave nor receive bequests. The simple LCM yields strong predictions about the paths of consumption and saving, given any pattern of lifetime earnings. Key implications are that consumption in any year depends only on total wealth (including the present value of future earnings and Social Security wealth) and not on either current income or on the types of assets in which wealth is held. But a significant amount of empirical research contradicts the implications of the simple form of the LCM.

Enhancing the LCM to account for uncertainty, liquidity constraints, the effects of government transfer programs, and bequests makes the model more consistent with observed data. Modified approaches that account for these factors, and less traditional approaches such as
the “behavioral” life cycle model, suggest that both current and lifetime income could affect current consumption and that the propensity to consume out of different forms of wealth may differ.

Existing large-scale microsimulation models typically contain very rudimentary treatments of household consumption. A standard approach is to project wealth of future retirees by a simple reduced form relationship that projects wealth at retirement based on wealth in some earlier period and other household characteristics.

An alternative would be to introduce a consumption function directly into the model, compute saving as the difference between income and consumption, and build up household wealth year by year based on projected saving and an assumed distribution of rates of return. Such an approach would allow for direct simulation of changes in policies (such as individual accounts) on wealth accumulation but the results from such simulations would be highly sensitive to the exact specification of the consumption function. Because the current state of research still provides only limited guidance on how best to model consumption, models should be structured to facilitate testing of the consequences of alternative representations of consumption behavior.

_Macroeconomic Growth Modeling for Microsimulation of Social Security Reform._

Social Security reforms may have significant effects on labor supply and saving. As a consequence, their effects on wages, interest rates, and the growth rate of output may be too large to be disregarded. Micro-simulation approaches, on their own, cannot account for these overall economic effects. Some type of macro-economic model is needed to supplement microsimulation models.

There are a wide variety of macroeconomic models that are used for different purposes. For the purpose of modeling the long-run effects of Social Security reforms, demand side models that are useful for forecasting short-run fluctuations in output and employment are not helpful. What is required is some type of neo-classical growth model that projects long-run output as a function of labor supply, the stock of capital, and state of technical knowledge. But within this grouping, there are a number of available options.

One useful way of classifying empirical macroeconomic models is by whether or not they explicitly model optimizing behavior. We classify models in these two groups as _utility-based_ models and _ad hoc_ models.

The advantage of using utility-based models is that they are theoretically rigorous and consistent and, because they are based on explicit modeling of preferences, can generate measures of the effects of policy changes on economic welfare. In addition, structural models are less subject to the “Lucas critique”, which states that relationships estimated under one policy regime
will not be stable if the policy regime changes, than *ad hoc* models. Utility-based models are less subject to the Lucas critique because their parameters are arguably “deep” parameters that reflect underlying preferences and technical production relationships instead of behavioral responses observed in an environment that is likely to change.

The advantage of using *ad hoc* models is that they typically include more sectoral details and can be tailored to model policy reforms that have been observed in the past. If their behavioral specifications are more detailed, they may be more suitable than structural models that represent household preferences in an over-simplified fashion. The problems of some *ad hoc* models result from excessive aggregation (for example using a single consumption relationship for all households, regardless of age) and can be reduced by estimating separate equations for subgroups that behave differently.

Microsimulation and macro-economic models can be linked with either a “bottom-up” or “top-down” approach. In both approaches, some type of macro model must be used to generate production relationships and demands for labor and capital and to solve for equilibrium prices. The difference between the two approaches is in the treatment of household decisions. In the “bottom-up” approach, household decisions in the microsimulation model determine aggregate factor supplies. These are then fed into the macroeconomic model, which generates factor demands and solves for prices. The two models run simultaneously until demand and supplies are equilibrated. In the “top down” approach, the calculation of equilibrium prices and outputs is done wholly within the macro model, which generates both demand and supply for factors of production. The microsimulation model is then used to allocate aggregate factor supplies from the macro model among households.

**Risk and Uncertainty.**

Individuals face a wide variety of risks that can adversely affect their income security in retirement. In the context of proposals to supplement or partially replace the current Social Security system with individual accounts, discussion has focused on risks associated with fluctuations in returns on equity. But there are many other potentially more important risks that people face. These include financial risks associated with wage declines, unemployment, uninsured health care costs, divorce or death of a spouse, living longer than expected (thereby using up assets), loss of pension coverage, and potential changes in Social Security benefits and taxes. It is important to model these various risks in a comprehensive way. It is also important for models to consider the interactions of different public programs in altering risks, including Social Security retirement and disability benefits, Unemployment Insurance, Medicare, and other Federal and state transfer programs. Finally, government measures to reduce risks may interact with actions the private sector takes, such as private health and life insurance, employer pensions, privately purchased annuities, and within-family transfers.
There is now a rapidly growing economics literature that examines how individuals use Social Security and other public and private mechanisms to cope with risk. In reviewing this literature, we contrast the more formal dynamic programming/stochastic overlapping generations (DP/SGE) models with “short cut” approaches.

Short cut approaches provide important advantages in simplification and ease of communication of results, but have several limitations. First, to the extent they are not based on a consistent utility maximization framework, they are unable to assess the effects of alternative policies on individual welfare. Second, they typically fail to account for potentially large behavioral and general equilibrium feedback effects. This could lead to errors in the direction as well as the size of effects of policy changes. Finally, they are usually subject to the “Lucas critique,” in that the estimated relationships they use are not invariant to changes in the policy regime.

DP/SGE models address these problems and are becoming sufficiently realistic to be taken seriously in forecasting and policy analyses. They have produced valuable insights on how household responses to reforms of Social Security and the changes in economic equilibrium that the responses produce may affect the welfare of different cohorts in the population. For example, as discussed in the chapter, some DP/SGE models show that Social Security’s adverse effect on household saving results in a net reduction in individual welfare, even though individuals benefit from the opportunity to purchase an “actuarially fair” annuity.

But DP/SGE models still suffer from important shortcomings. They typically lack a “unified” model of social insurance that provides a single, consistent, integrated framework of policy evaluation. The computational complexity of solving DP/SGE models makes them difficult and expensive to use. They require large amounts of data to identify and estimate their unknown parameters, which for this reason often lack sufficient empirical validation. They lack a good representation of employer behavior, such as the demand for labor in different age groups and the decision to offer pensions, health insurance and other fringe benefits. Finally, the models fail to capture the potential effects of macro-economic shocks and must assume an equity premium instead of solving for equity returns endogenously. Moreover, there is as of now no good explanation for the observed high equity premium and large volatility in equity prices.

DP/SGE models are still a long way from being empirically justified. They are unable to solve and estimate a full life cycle model with realistic treatment of labor/leisure and consumption/saving decisions. They are also unable to incorporate various additional choices, such as educational and career choices and marital and childbearing decisions.

Recent research with advanced DP/SGE models has found that the current PAYGO system reduces the welfare of most cohorts because of its adverse effects on work effort and capital accumulation. This raises the question of whether the large popularity of the current system reflects intergenerational transfers that benefited early cohorts in the system or some
component of risk reduction that is not adequately captured in the models. Further advances in this type of work could produce new insights on the political economy of the current system and how it might best be reformed.
REFERENCES TO INTRODUCTION AND SUMMARY

CHAPTER 1

REVIEW OF SOCIAL SECURITY
MICROSIMULATION ALTERNATIVES

I. INTRODUCTION

Tremendous economic and demographic shifts in our society have made Social Security reform all but inevitable. Many analysts (see, for example, Weaver, 1997: 161, Salisbury, 1996) have expressed grave concerns that, at a time when we need to be making choices that will affect the economic well-being of millions of Americans both today and in generations to come, our capacities for modeling the consequences of Old Age, Survivors, and Disability Insurance (OASDI) reforms are at best limited. The problem is particularly acute if one wishes to examine consequences of policy changes for detailed subgroups of the population. A number of microsimulation models recently developed both inside and outside of the Social Security Administration, including the Historical and Projected Cohorts Models (HCM and PCM), the Model of Income in the Near Term (MINT), and the Cornell Dynamic Microsimulation Model (CORSIM), hold promise for soon filling this analytical void. Yet even these new models may not be able to answer some important questions. This chapter addresses whether these and other existing micro-level simulation models are sufficiently equipped to permit analyses of prominent Social Security reform proposals in coming years. Where the models are judged to be deficient, we assess whether they could be enhanced to meet the pressing needs for reasonable predictions.

This project specifically focuses on the concern, highlighted in many recent model reviews (e.g., Anderson, 1999, Citro and Hanushek, eds., 1997, Hollenbeck, 1995, Technical Panel on Methods and Assumptions, 1999), that existing microsimulation models do not accurately capture the interactions between policy and behavioral changes and between those behavioral changes and the performance of the macroeconomy. That is, where models realistically detail differences among members of the population, they typically lack connections between these individuals’ collective behaviors at a point in time, their subsequent changes in behavior in response to policy changes, and their resulting individual and collective opportunities. Such worries are especially salient when Social Security reformers call for

---

2 In preparing this chapter, we received valuable comments and suggestions from Benjamin Bridges, Olivia Mitchell, David Pattison, Rudolph Penner, Douglas Wolf, and Sheila Zedlewski.

3 It may be analytically useful to distinguish between behavioral response at the micro-level to a policy change and the macroeconomic effects of policy change. C. Eugene Steuerle nicely describes the latter as the sum of the former, though the modeling challenges each imposes are, in some strategies, distinct.
massive fiscal changes (like substantial tax hikes or benefit cuts implemented over several decades), or fundamental structural changes (like new individual accounts or sizable investment of trust fund reserves in equities), as opposed to minor short-run technical adjustments to benefit calculation algorithms.

Our objectives for this topic are threefold. First, we identify the Social Security policy approaches that are highest on the public agenda (and thus most likely to be factors in OASDI reform debates). Second, we use the best available research to illuminate some of the likely effects of these proposals on the larger economy and society. (Later chapters will probe these issues in more detail, especially with respect to saving and macroeconomic effects.) In general terms, we try to determine the extent to which social science research has uncovered fundamental structural relationships between government-imposed changes to incentives and funding mechanisms and individual and family decisions about labor supply, saving, and program participation. In examining these relationships, we necessarily consider changes to constraints that individuals face as a consequence of their own and others' choices (i.e., macroeconomic effects). Third, with findings from this literature survey in hand, we then assess existing micro-level models' capacities for simulating the distributional consequences of changes to the Social Security law in ways that incorporate these proposals' presumed higher-order effects on the larger economy and thus the population. We suggest which existing policy questions current models have a capacity to inform, and which policy questions they could soon inform with reasonable extensions.

With these three objectives in mind, we have organized this chapter into three parts. First, we define what we mean by a model and identify four types of models with features that permit distributional analysis of Social Security reform proposals. Next, we outline criteria for evaluating these models. This requires identifying which Social Security reforms are of interest, what outcomes we want to explore before and after Social Security reforms, for which groups we want to make comparisons, and what macroeconomic and behavioral responses to reform we wish to integrate into a model. Third, we describe several existing classes of models, focusing on their capabilities and limitations in terms of meeting these needs.

II. WHY MODELS ARE IMPORTANT AND CLASSES OF MODELS FROM WHICH ANALYSTS CAN CHOOSE

As Anderson notes, models are representations that can be used for identifying and illustrating key features of systems, for simulating systems under alternative conditions, and for projecting these systems into the future (1999: 1-2). Because social scientists rarely have the opportunity to conduct large-scale experiments, they cannot typically observe directly what would happen if a single variable or series of variables in a program like Social Security were to change. Generally speaking, we need models to help us to understand historical events about which only incomplete historical data are available, and thus to determine “what happened.” We also need
models to explore historical and projection-period hypotheticals to address “what would have happened if” or “what would happen if” questions. The alternative to modeling these events or hypotheticals is to rely on comparisons and intuition, and social research consistently finds dangers in either of these approaches. Intuition is prone to errors in perception, and comparisons suffer because of both observed and unobservable differences between the cases one compares. Further, even the most capable of analysts are simply unable to hold all the necessary information on such complex, detailed systems as Social Security in their heads.

A further benefit of constructing models is that it forces precision in policy analysis. If a policymaker wishes for researchers to conduct distributional and/or cost estimates of his/her favorite proposal, than he/she must specify exactly which tax and benefit parameters are to be changed, when, and by how much. Also, the researchers who build and use models need to be explicit about their assumptions regarding behavior. They must, for example, make detailed assumptions about future patterns in mortality and retirement.

Models of the retirement income system in general and Social Security in particular come in many varieties. If one's goal is to evaluate effects of retirement income policies, especially with respect to Social Security, on individuals and families, then one needs a model that either explicitly represents individuals and families or that disaggregates the population into a large number of cells. At a minimum, any model that we seriously examine here needs to have a few dozen cells, differentiating members of the population by sex, birth cohort, and lifetime income. Ideally, though, we would like much more detail, including, for example, information on timing of individuals’ first receipt of Social Security benefits and/or timing and duration of marriages, as these are factors that are central to certain kinds of OASDI reform.

Four kinds of models have the current potential to meet these initial, minimal criteria: representative worker models; cell-based models; population microsimulation models; and applied (computable) general equilibrium models, including overlapping generations (OLG) models and multi-agent stochastic overlapping generations general equilibrium (SGE) models. We provide a brief description of each of the four types below, in our third section (more complete descriptions are available in Anderson, 1999, and Citro and Hanushek, eds., 1997). Chapter 5 of this report reviews more academic Social Security models, including “textbook models” and single-agent dynamic programming (DP) models, and provides more extensive detail on SGE models.

4 In some cases, representative worker models could be considered to be a subset of cell-based models. We distinguish between the two types, defining cell-based models as those in which, whether using stylized, median, or average workers, the model population aggregates up to national totals, and representative worker models as those in which the model population (and its Social Security experiences) need not approximate any system-wide total.
III. DESIRED CHARACTERISTICS OF A MODEL OF SOCIAL SECURITY

Determining from among these varied options the type of model of retirement income that one should use for a given application requires that one consider a wide array of model characteristics. These characteristics span two major areas: first, a model's substantive content and then its technical features. At least six major areas compose substantive content: a model's programmatic detail (whether and how it incorporates Social Security and other federal, state, and local tax/transfer regulations), its detail on social and economic processes (like fertility, mortality, labor force participation, savings behavior), its population (i.e., the distinct groups one can analyze using the model), its time horizon, the extent to which social and economic research inform the relationships embedded in the model, and the extent to which model outcomes have been validated against external data, whether historical or projected. Technical features, in contrast, include factors like a model's run time, the language in which it is written, the order in which it processes records, the types of computers on which it generally runs (and hence its processor, memory, and storage requirements), the methods analysts use for accessing its output, and its modularity, the extent to which developers can add new pieces to the model as research interests expand. We address each of these distinct substantive areas in turn, setting priorities and suggesting which are the model features on which users concerned with Social Security reform are more and less able to compromise. While our discussion touches upon a few of the more salient technical issues, most notably run times, it does not focus on them. (Sabelhaus, 1999, provides one discussion of technical issues associated with micro-level modeling.)

1. Capacity: Ability to Capture Proposed Reforms

Several researchers have produced detailed reviews of prominent Social Security reform proposals (see, for example, Concord Coalition, 1998, ERISA Industry Committee, 1998). Noteworthy proposals include programs that President Clinton advanced in his State of the Union address, the recommendations of the 1994-1996 Social Security Advisory Council, legislation introduced in Congress, plans of expert groups and scholars, and position papers from research institutions and advocacy organizations. We classify provisions contained in these proposals into two major groups: structural reforms that would shift Social Security away from its current system of pay-as-you-go financing and programmatic reforms that would generally maintain the current structure, but change certain aspects of the program. Structural reforms include integrating individual accounts (using either an “add on” or “carve out” strategy), using general revenue financing, and diversifying the Trust Fund (e.g., strategies to invest fund balances in the stock market). Programmatic reforms include changes to: benefit calculation (e.g., changes to number of computation years, bend points, bend percentages, or cost-of-living adjustments); benefit levels (e.g., shifting spouse benefits to time of survivorhood); options for benefit timing (e.g., increasing and/or indexing the normal and/or early retirement ages); work incentives during benefit receipt, also known as "partial retirement" (e.g., changing or eliminating the earnings test);
the **taxable wage base** for Social Security (e.g., increasing or eliminating the taxable maximum); the **taxation of Social Security benefits**; and the **population covered** by Social Security (e.g., extending the Social Security system to integrate new hires in state and local governments not covered by Social Security); and also provisions to provide **minimum benefits**. Each of these types of reforms has distinct and important implications for model design. Few, if indeed any, existing models have the current capability to examine all of these proposals in a comprehensive manner.

Many have characterized current debates in Social Security reform as hinging on whether to "privatize" the system, but Geanakoplos, Mitchell, and Zeldes (1998) point out that these debates are more complex. These authors organize structural reforms that move Social Security away from the current system of modified pay-as-you-go financing along three dimensions—**privatization** (defined as replacing the current defined-benefit system with a defined-contribution system of individual accounts held by contributors), **diversification** (defined as investing funds in a broad range of assets), and **prefunding** (defined as the reduction of implicit and explicit debt)—and note that a reform can incorporate any one of these three elements without necessarily including the others. Reform discussions have addressed all three of these dimensions, both individually and in various combinations.

Personal account plans incorporate elements of all three dimensions. They have enjoyed particularly widespread attention and support, and at the same time have drawn heavy criticism. Two of the Advisory Council plans (those advanced by Gramlich and the Schieber/Weaver group) and a large proportion, if not the majority, of the plans recently introduced in Congress (see, for example, Moynihan and Kerrey's proposal) incorporate some type of individual account, often in tandem with some change to payroll tax rates. These accounts take a variety of forms in different plans. Sometimes they are mandatory, other times voluntary; sometimes they require that workers annuitize account balances on retirement, and in other cases workers are permitted to withdraw the balances in their accounts in a lump sum.\(^5\)

---

\(^5\)In a comparative analysis, Thompson (1999) identifies twenty important dimensions of variation in personal account programs and proposals, each of which has important distributional implications. These dimensions fall into four general areas: **general characteristics, investment management, withdrawal of funds, and guarantees.** **General characteristics** of personal account proposals include the following: whether a program is compulsory or not (yes/no); its contribution rate, and, in the U.S. case, whether contributions are added to the current rate; its budget financing (i.e., whether there is use of general revenues); who collects contributions; who remits benefits; who keeps the records; and frequency of employer reporting. **Investment management** dimensions include: selector of money manager; selector of investment strategy; number of options available to workers; and maximum lag time between contributing to and crediting of the fund. Issues surrounding the **withdrawal of funds** include: whether lump sum withdrawals are allowed; requirements of annuitization; requirements of price indexing; and selection of annuity provider. **Potential guarantees** encompass four elements: the absolute rate of return; the relative rate of return; any minimum benefit provision; prior law benefit; and solvency of investment companies.
Proponents of these plans justify their creation using two primary arguments: first, that the potential growth effects of increased saving would lead to higher general welfare, and, second, that taxpayers should have the right to control their own money and seek out the highest possible returns. Opponents of private accounts point to, among other concerns, the need to figure in both transition costs (due to the massive unfunded Social Security liability to both current workers and current retirees) and transaction costs, which will reduce returns to Social Security contributions. Many also express concern that implementing private accounts would lessen Social Security’s important redistributive role in U.S. society, though one can envision systems that rely on private accounts that are more redistributive than the current system. Some implications for models, then, of the need to model reforms which create private accounts include the following:

- Full, detailed individual earnings histories to capture contributions and accruals to accounts (and to the prior and remaining defined benefit system) accurately across the life course;
- Flexibility to capture variation in transaction costs, and posting, annuitization, and indexation options and regulations;
- Ability to model expected returns and risks to different investment choices;
- Ability to capture effects taking into account the variety of risks against which Social Security offers protection, including disability and death;⁶
- Ability to cost out any promised benefit or rate of return guarantees;
- Where account funds are self-directed, ability to capture the portfolio choices of new investors (taking into account the select nature of current investors);
- Ability to capture shifts in existing portfolios by current investors;
- Where accounts are fully voluntary or contain a voluntary component (e.g., additional contributions with favorable tax treatment), ability to consider differential take-up across social groups;
- Ability to model transition costs and implementations;
- Ability to model individual responses, including changes in retirement age and in saving, both within and outside of accounts;
- Ability to consider program interactions (see pp. 8 and 9 below);
- Ability to integrate macroeconomic feedback effects on economic growth and factor prices (interest rates and wage rates).

Plans to diversify the holdings of the Social Security Trust Fund are integrated into one version of the Ball (i.e., Maintenance of Benefits) Advisory Council plan and a proposal advanced by Aaron and Reischauer (1998). While several controversial issues surrounding diversification concern corporate governance issues (e.g., whether/how the government would vote any shares

⁶Chapter 5 will present an extensive discussion of issues surrounding risk and uncertainty.
of corporate stock that it held), and are thus not necessarily in the domain of an individual or family-level model, such plans do raise important distributional issues. Models will be more effective at considering these types of reforms if they:

- Explicitly model individual portfolio allocation to capture shifts in response to the perceived risk of Social Security;
- Consider and account for changes in prevailing rates of return that will result from changing demands for equities, either from government or household investment.

Other reform proposals are less fundamental in nature than creating private accounts or investing Social Security reserves in the stock market. Some examples are changes to specifics of the Social Security benefit formula, including the number of computation years used in determining benefits or the indexing formula for the benefit formula's bend points. Such proposals appear in many plans (see, for example, the Maintenance of Benefits and Individual Account plans of the 1994-1996 Advisory Council). Changes to the replacement percentages for spouse and survivor benefits (see, for example, Burkhauser and Smeeding, 1994) are similar in nature, though a bit more demanding in terms of their requirements for a model's life histories (i.e., histories must incorporate features of both spouses' trajectories in great detail). If one assumes that people do not alter their behavior in response to the change in benefit policy, these types of reforms are fairly straightforward to represent in several types of models that reproduce the Social Security benefit calculator in detail. In many cases, for example a minor change in the indexing formula, such an assumption may be reasonable. Implications of these separate proposals for models include the following:

- Models need to include a detailed benefit OASDI calculator;
- They must contain individual earnings histories that are realistic, generating not just reliable Average Indexed Monthly Earnings (AIME) as calculated under current law, but also annual streams of earnings that duplicate observed patterns;
- They need to consider marriage histories, including durations of marriages that end in divorce or widowhood;
- They need to include earnings histories that incorporate the appropriate correlations between spouses' earnings.

When individuals continuously adapt to changes to Social Security provisions, modeling becomes more complex. Several important proposed changes to existing Social Security parameters, for example increases in the normal or early retirement age or the exempt amount for the earnings test, are straightforward to analyze in many models in the sense of changing a model's representation of Social Security law, but require more complex incorporation of behavioral responses. In general, changes to OASDI provisions affect two key person-level economic variables: the after-tax wage and wealth. Both affect a person's work effort (a lower after-tax wage reduces work effort, holding wealth constant, while lower wealth increases work),
and saving (lower wealth tends to increase saving). The complexity lies in how changed provisions affect the marginal return from working and saving.

For most people, the Social Security tax and benefit structure has complex effects on incentives to work and save. Individuals with significant attachment to the covered labor force and earnings that, on average, exceed about half of their spouse's, face an effective tax rate in their prime working years of less than the statutory tax rate for OASDI contributions because additional earnings raise future benefits, though not in proportion to taxes paid. A forward-looking person with full information should be able to evaluate the ways that reforms affect future income streams, and adjust current income (via labor supply and saving choices) to compensate for the future gains or losses that reform imposes. Married persons with less attachment to the labor force who have spouses with more significant covered labor force experience can face very different the incentives from this program (see, for example, Feldstein and Samwick, 1992, on marginal tax rates). For them, the effective tax rate will be equal to the statutory rate if their benefits are determined by their spouses' earnings instead of their own. The tax is likely to reduce labor supply because the "substitution effect" due to the reduced return from an additional hour of work is likely to outweigh the income effect due to the (smaller) proportional reduction in total family income.

When considering likely changes to Social Security incentives, we need nonetheless to keep in mind that individuals operate in a constrained environment. They face minimum income needs in every period and are likely to face limits in their ability to borrow against their human capital. Further, they confront a labor market in which opportunities often cluster around certain normative arrangements (e.g., a forty-hour work week), rather than a continuous set of labor/leisure choices. Finally, they may have cognitive limitations. Some individuals may be completely myopic. A growing literature, which Chapter 3 will discuss, suggests that many individuals discount future income guarantees at a very high rate. As a consequence of these separate factors, individuals' responses to changing incentives may be more modest than theory might suggest for individuals operating in less constrained environments.

Ideally, a model can be structured so that the members of its population would respond automatically to changes in after-tax wages and wealth that result from changes in OASDI parameters. For example, a model's retirement timing function could depend on a person's expected Social Security and pension benefits at the normal and early retirement ages, so that an increase in either age could result in people delaying the choice to receive Social Security retirement benefits and could increase labor supply and earnings for those who have chosen to receive benefits. People could also respond to these types of reforms by changing their saving behavior (i.e., if they expect to receive less from Social Security, or to wait longer before they can receive benefits, then they may save more when younger). This implies that models need:
Chapter 1: Review of Social Security Microsimulation Alternatives January 2000

- A capacity, whether at the micro level or through calibration, to account for different behavioral responses between groups (e.g., different responses to the earnings test for men and women);
- Information about opportunities to earn wages in retirement;
- Information about pensions and other wealth.

Some of these same proposed changes to Social Security, both structural and programmatic, are likely to shift the composition of benefits within the program and between Social Security and other government and private programs. A model that anticipates and accounts for these changes will be more effective than a model that is more limited. Raising the normal and early retirement ages, for example, is likely to raise participation rates in the Social Security Disabled Insurance program and perhaps in the Supplemental Security Income (SSI) program as well, thus offsetting some cost savings (see, for example, Bovbjerg, 1998 or Wittenburg, Stapleton, Scrivner, and Hobbie, 1998). Whether Medicare continues to be available to workers at age sixty-five is likely to influence decisions about first benefit take-up for Social Security (see, for example, Rust and Phelan, 1997). Minimum Social Security benefit provisions, as proposed in the National Commission on Retirement Policy (NCRP) plan (1998), could serve to reduce SSI costs.

On the private-sector side, Social Security reforms could also have significant effects on the costs for employer pension and insurance programs. ERISA Industry Committee (1998) documents that approximately half (fifty-one percent) of defined benefit pension plan participants are in plans that are explicitly integrated with Social Security, as are many defined contribution plan participants. Slusher (1998) and Bender (1999) explore similar issues, considering the characteristics of individuals with integrated pensions. They find that individuals with integrated plans differ significantly from the population at large.

If a model is to capture these interactions accurately, it needs to contain complex modules of the DI program (see, for example, Benitez-Silva, Buchinsky, Chan, Rust, and Sheidvasser, 1999) and SSI, rather than simple arrays of participation rates in the two programs. Such an SSI model, for example, might take into account the relationship between sizes of SSI benefits, which are state-specific, sizes of Social Security benefits, and levels of take-up. These variables matter because social science research consistently shows take-up to vary directly with benefit levels (see, for example, literature on the SSI program, including McGarry, 1996; Warlick, 1982; and Zedlewski and Meyer, 1989). Models might also need to incorporate extremely detailed pension rules, including those about integration with Social Security, in order to capture these interactions. The modeling implications are that:

- Models that describe Social Security in isolation are likely to have less applicability and validity for exploring certain reforms than those models in which Social Security outcomes are tied in detail to private pension and health insurance plans and government social assistance programs such as SSI or Food Stamps;
Models should have a detailed module of DI receipt;
Models should consider the Medicare eligibility age in timing functions;
Social insurance and social assistance models should incorporate differential take-up functions, and not assume full participation based on eligibility.

Proposals that would change income taxation of Social Security benefits (see, for example, the Maintenance of Benefits, Individual Accounts, and Moynihan-Kerrey plans) impose similar demands on models. In order to model such proposals, one again needs information on non-wage income (both one's own and one's spouse's), because this determines the tax bracket in which an individual or couple falls. Information on asset and pension accruals is also useful for identifying those populations that are most vulnerable. Proposals to divert funds (i.e., budget surpluses) from general revenues to the Social Security Trust Funds pose parallel challenges because they change the overall distribution of the tax burden. Shifting from payroll tax financing of retirement and disability benefits toward federal income taxation of these benefits makes the system more progressive, because payroll taxes are imposed on earnings up to the taxable maximum at a flat rate, while income taxes are levied on both wages and non-wage income at graduated rates. Other Social Security reform proposals, like that of Kotlikoff and Sachs (see Kotlikoff, 1997), would institute a broad-based consumption tax to finance the transition to a regulated, privatized social insurance system. Distributional analyses of such a plan need to incorporate the differential impacts of income and consumption taxes. Income and consumption tax financing changes require, then:

- Ability to model existing federal taxes and proposed new taxes, which itself implies a need for measures of individuals’ non-wage income (i.e., at least crude information about pension and asset income) in addition to their earnings;
- Information about consumption and how it differs among households and over the life cycle.

Among the OASDI reforms most popular with the general public are proposals to increase the taxable wage base (see, for example, polls of National Public Radio/Kaiser/Kennedy School, 1999, and Americans Discuss Social Security, 1998). A few plans (for example, the Moynihan-Kerrey proposal) have such provisions. Models relying on administrative records may not include all earnings, and, if they do not, developers need to pay special attention to imputing and projecting earnings above the taxable maximum. Implications of these proposals for models include needs for:
Information about wage income above the taxable maximum;
Mechanisms with which to integrate likely employer-side feedbacks (e.g., employers may change the composition of compensation packages) if Social Security begins to cover more earnings.

A number of proposals (all three of the Advisory Council plans and the NCRP plan) would extend Social Security coverage to previously excluded populations. This is a particularly challenging reform scenario for modelers, because so little high quality micro-level data exists on these groups. Implication of these particular reforms for models is that:

- A model population needs to cover all members of the Social Security population, not just workers with insured status and/or some covered earnings.

The wide array of proposals to reform the Social Security program that we have just discussed is not, of course, exhaustive. Analysts will likely wish to use models to consider both policy changes that have been advocated in the past and future ideas for reform. If models are flexible and modular, they will be capable of responding to the dynamic, evolving demands of the policy community. As noted earlier, no model can now meet all of these needs. Certain provisions may need to be estimated outside of a particular model and then fed back into it, for example as a constraint on other model predictions.

2. Outcomes

In addition to knowing which proposals they want to evaluate, model developers need to identify the proposals’ effects that they want to capture. A model must provide reasonable estimates of a variety of measures of Social Security outcomes as well as measures of other economic outcomes. These outcomes are cross-sectional and longitudinal, and aggregate and distributional. Table 1-1 suggests a number of important outcomes related to OASDI and economic well-being more generally that models should produce.

Table 1-1 enumerates person-level, cohort-level, and aggregate (or system-wide) social and economic outcomes connected with Social Security and well-being over the life course. The table differentiates outcomes measured at a point in time—cross-sectional measures (the lower portion of the table)—from outcomes based on full life experiences—longitudinal or life-course measures (the top portion). Cross-sectional measures, such as the level of one's Social Security benefits, the number of participants in the program in a given year, or the number of older

---

7Earnings sharing proposals, for example, received considerable attention in the eighties (see, for example, Congressional Budget Office, 1986), but are now primarily discussed in connection with privatization (see, for example, the Kotlikoff and Sachs proposal in Kotlikoff, 1997).
### Table 1-1

**Important Outcome Measures for Social Security Analyses**

<table>
<thead>
<tr>
<th>Longitudinal OASDI Outcomes</th>
<th>System-wide</th>
</tr>
</thead>
<tbody>
<tr>
<td>Person-level and Cohort-level</td>
<td>Fiscal balance over long term</td>
</tr>
<tr>
<td>Present value (by program) of lifetime contributions to OASDI</td>
<td>Present value (by program) of lifetime benefits from OASDI</td>
</tr>
</tbody>
</table>

**Cross-Sectional Outcomes**

<table>
<thead>
<tr>
<th>OASDI Program</th>
<th>Other Social and Economic</th>
<th>Related Social Welfare</th>
</tr>
</thead>
<tbody>
<tr>
<td>Person-level</td>
<td>Aggregate</td>
<td>Person-level</td>
</tr>
<tr>
<td>Covered by OASDI?</td>
<td>Proportion in covered work (including underreporting, evasion estimates)</td>
<td>Alive?</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pay FICA or SECA tax? How much?</td>
<td>Total number of contributors; total value of receipts</td>
<td>Labor force participation, wages</td>
</tr>
<tr>
<td>Receive disability, retirement, or survivors' benefits? How much? What type (worker/dependent)?</td>
<td>Total number of beneficiaries and amount of benefits paid, by program and beneficiary type</td>
<td>Consumption, (non-pension) saving and wealth, income from wealth</td>
</tr>
<tr>
<td>Pay income tax, including on Social Security benefits?</td>
<td>Total income tax; tax on benefits; tax rate on earnings</td>
<td>Poverty status (pre- and post-OASDI)</td>
</tr>
<tr>
<td><strong>Total number of beneficiaries and amount of benefits paid</strong></td>
<td><strong>Total number of beneficiaries and amount of benefits paid</strong></td>
<td><strong>Total number of beneficiaries and amount of benefits paid</strong></td>
</tr>
</tbody>
</table>
Americans that Social Security benefits remove from poverty, are essential for considering both Social Security costs and the adequacy of Social Security benefits. Life-course measures, including the internal rate of return and the present value of net benefits, which Leimer (1995) and Geanakoplos, Mitchell, and Zeldes (1999) detail, contrast, and evaluate, help to summarize lifetime experience with Social Security and thus facilitate comparisons across groups, including full generations. These latter measures are helpful for communicating information about the equity of the OASDI system, and have the advantage of accessibility even for non-technical audiences. One must use care in interpreting them in isolation, though, because higher rates of return can occur when one receives lower lifetime income (see, for example, the discussion in Bosworth and Burtless, 2000). Such measures also have limits if one considers Social Security in a tax-transfer framework rather than as a forced saving vehicle (see, for example, the analysis in Choi, 1991).

Nearly all of the other cross-sectional measures that Table 1-1 lists have longitudinal analogues, both at the person and aggregate level. For example, one can aggregate the cross-sectional measures of an individual's consumption over the person's lifetime to derive a measure of lifetime well-being, and can likewise add up these lifetime consumption paths over an entire cohort or population to gain insight into long-term, aggregate social welfare. Similarly, one can consider a person's lifetime poverty experience, rather than just whether one is in poverty at a point in time.

Analysts can provide more comprehensive accounts of winners and losers from Social Security reforms when they have access to full distributions of the person-level outcomes that Table 1-1 identifies, rather than just mean or median values. If a model sample is complete and representative, one can derive the measures in the aggregate columns in the table by adding up outcomes from the person columns and grossing up total estimates by the ratio of the model population to the historical one at the model's start. Otherwise, analysts can apply differential (rather than uniform) weights when grossing up to ensure that the model's sample reflects a national population or subgroup (for example a cohort or series of cohorts).

3. Population Definition/Aging Modules

When defining the population of interest for effective retirement income studies, analysts often want a lot of detail. Ideally, analysts would have a capacity to assess these cross-sectional and longitudinal Social Security outcomes along as many social dimensions (including, for example, race, sex, lifetime earnings, cohort, marital status, and benefit type) as is practical with as much precision as possible. Such detail enables one to identify clear winners and losers, both within and between cohorts, from changes to given Social Security regulations or structures.

Unfortunately, as a model's population becomes more detailed, a model also becomes larger and more complex. When making projections, one may need to vary, for example, fertility and mortality projections along all of the population dimensions that one wishes to use the model.
to explore.\footnote{One can frame this discussion of groups differentiated in a model a bit differently, by instead describing the desired interconnections and dependencies, or endogeneity, of model processes.} If one wishes to compare Social Security outcomes by cohort, marital status, lifetime income, and sex, one likely needs to be sure that the model's implied mortality rates vary, in accordance with the most recent estimates and best social and economic theory, along each of these dimensions. Using a probability table approach, this requirement quickly leads to huge arrays with hundreds of elements and hence an enormous demand for data. Using regression-based approaches, data demands are great as well, and other sorts of complexities also arise.

Given the wide array of population groups in the U.S. and the diversity of U.S. social programs, it is virtually impossible to differentiate members of the population along all major lines of within- and across-cohort interest. Assessments about priorities for model population definition are both necessary and, of course, subjective. They can, however, be informed by the regulations of the Social Security program, which differ depending on one's birth cohort and explicitly consider marital and work histories, Social Security literature, for example the presence of comprehensive and useful historical pieces about OASDI that are based on administrative data (see, for example, Moffitt, 1984), understanding of the current political climate (e.g., the presence and strength of advocacy groups that represent certain constituencies, and the corresponding likelihood that Congress might implement certain types of reforms), examination of trends in components of OASDI (e.g., increasing duration of first receipt spell among disabled workers, described in Rupp and Scott, 1998), and recognition of high levels of economic distress among certain population subgroups.

4. Time Horizon

Closely connected to the question of the definition of a model's population is the time horizon over which a model simulates Social Security outcomes (and, implicitly or explicitly, demographic and economic events). Just as model size and complexity increase with a more expansive population definition, so too do they increase with a longer time horizon. If one specifically wishes to model long-run fiscal balance of Social Security and long-run revenue effects of competing proposals, one needs to simulate contributions from taxpayers and payments to beneficiaries for a minimum of seventy-five years (the long-range time horizon employed by the Social Security actuaries). Full cohort analyses of OASDI experiences can, in certain implementations, require even longer time horizons, and usually require extensive treatment of censoring (that is, unobserved and/or unsimulated portions of cohort members' lifetimes). Specifically, they demand that either one survives the model population to the point of death of the last remaining member of a cohort (to at least age one hundred for cohorts now reaching retirement age) or one adjusts future income and taxation streams to account for probability of survival to each age. Similarly, full cohort analyses require that one has a measure of people's experiences with Social Security before they are observed in a survey. Modelers can impute prior
experiences of members of a cohort into a model and in doing so probably wish to account for mortality and divorce that occurs among members of the cohort of interest prior to the initial observation.

Simulating Social Security outcomes over such lengthy time horizons requires, at a minimum, development of reasonable projections of long-range future patterns, including both levels and differentials, in marriage, divorce, labor force participation, wages, and mortality. For certain processes, like fertility or divorce, one can reasonably assume that in the future age-specific patterns will be roughly consistent with those recently observed. For other processes, for example mortality or wage levels, such an assumption is less plausible, so developers must carefully integrate trends. The length of the historical interval over which one measures trends can have profound consequences for forecasts.

5. Integration of Social and Economic Research on Social Security Parameters

Especially in the presence of extensive time horizons, useful models of the Social Security system need to integrate the best economic and social research on processes, like labor supply and saving, related to Social Security taxation and benefit receipt. This is true whether a model relies on relatively simple updating mechanisms, like applying aggregate rates to population cells, or more complex updating procedures, including either reduced form or structural equations calibrated to external targets. For modelers undertaking complex projects that replicate fertility, marriage, work patterns, saving, portfolio allocation, and consumption, the task of maintaining consistency among multiple assumptions is even more daunting.

Where no reform that resembles a proposal has ever been implemented in the U.S., however, it is difficult to assess what reasonable assumption parameters for model feedbacks should be. A substantial fraction of the reforms that we have discussed thus far fall into this category. Economic and sociological theory, and analytic models based on these theories, can provide a compass in such uncharted waters. Frequently, though, social scientists have failed to reach consensus on the nature of relationships between key variables, and tests of these relationships in complex empirical settings, in which multiple changes take place concurrently, are inconclusive.

Besides needing to impose consistency between micro and macro outcomes, models should also incorporate well-documented relationships, ideally theoretically grounded, into micro-
level processes themselves. Anderson (1999), for example, describes and expresses concern about the lack of economic content in many microsimulation models. Rather than relying on individual and family-level utility functions, as would be consistent with economic theory, these models tend to employ descriptive functions that characterize, for example, earnings and wealth patterns as resulting from age, past earnings experiences and wealth holdings, and a few other demographic variables. While such functions may be useful for replicating current joint distributions, when laws and incentives change, the modeler must generally make ad hoc assumptions about responses at the micro level. Adding prices and proxy measures of expectations to micro-level decision functions is one possible, if partial, solution to this existing limitation. Imposing behavior from “above” through integration with/calibration to a macroeconomic or CGE model is another.

Unfortunately, the literature has not come to a consensus on some of the key behavioral responses to Social Security. The effect of changes to Social Security benefit levels on saving behavior is still a subject of great contention, as one recent, comprehensive literature review suggests (Congressional Budget Office, 1998). Likewise, research on the effect of tax incentives on private saving has yielded a wide range of estimates, with some analysts suggesting that incentives create large amounts of new saving and others suggesting that they spur virtually no new saving (see, for example, Poterba, Venti, and Wise, 1995, Attanasio and DeLeire, 1994, and Engen, Gale, and Scholz, 1994). The spread in estimates like these gives the micro-modeler a shaky foundation from which to start when attempting to model integration of private accounts into the Social Security program. (As noted earlier, Chapter 3 discusses the saving literature, and ways to integrate its findings into micro-level models, in greater detail.)

Where empirical results are more robust, research often suggests that magnitudes of responses to various policies are likely to vary over time (e.g., to differ depending on whether one is early or late in the life cycle or where the economy is in the business cycle) and across population subgroups (e.g., married women are likely to respond differently than married men). Many researchers believe, for example, that workers respond differentially to hikes in employer/employee payroll taxes based on their place in the life course (see, for example, Congressional Budget Office, 1997: 29).\footnote{We do not discuss employer response to payroll tax hikes, since most economic research suggests that workers rather than employers bear essentially the full burden of the payroll tax (under the theory that employer are indifferent to whether they pay the government or the worker).} Most research suggests that for younger and middle aged men and unmarried women (those not yet near retirement), labor supply is relatively
inelastic. To the extent that men and unmarried women do respond to tax increases, in some cases they may actually work more when their taxes go up. For married women, in contrast, labor supply is found to be relatively elastic. As their taxes go up, married women tend to work less, and the size of the effect is greater than the size of the effect for men. At older age ranges, though, labor supply response to changes in Social Security tax rates may differ, with the earnings test playing a potentially important role for responses in certain groups (see, for example, Friedberg, 1997).

Another relatively consistent finding is that Social Security benefit cuts should lead to modest increases in labor supply and delays in first receipt of retirement benefits. Most studies suggest that an increase of a year in the normal retirement age should lead to delays of a few months in first Social Security benefit receipt, with greater effects for men than for women (see, for example, Fields and Mitchell, 1984; Burtless and Moffitt, 1984; Gustman and Steinmeier, 1985). There is less research and consensus on the effects of likely changes to Social Security's early eligibility age, though many speculate that the behavioral response is likely to be more substantial (Congressional Budget Office, 1999). Other research presents strong evidence of significant effects of the timing of Medicare eligibility on the labor supply of older workers (Rust and Phelan, 1997).

The level of complexity and uncertainty surrounding the effects of many of these reforms suggests two priorities for modeling efforts. Because so many of these issues are unresolved, models will be more effective if they can quickly and easily simulate alternative responses to many Social Security reforms. If we are not sure, for example, how much people will delay their retirement in response to changes in the normal retirement age, then modelers can (and probably should) simulate several plausible alternatives that cover the middle range of estimates. Second, because most reform packages include multiple provisions (e.g., changes to benefit computation in addition to increase in the retirement age) that will surely interact with one another, modelers should consider integrating related programs (e.g., OASI and DI) and building in likely interactions between them, employing a range of possibilities for these interactions wherever possible. Again, using the example of changes to the normal retirement age, modelers should consider the increased period of eligibility for DI benefits that such a change implies, and perhaps utilize a range of alternative worker responses.

IV. ASSESSMENT OF DIFFERENT MODEL TYPES

In this section, we describe and assess whether and how conclusions about relationships between key variables (e.g., tax rates and labor supply, benefit levels and saving) from the literature review are currently incorporated into each of several classes of models that we describe below. These classes include representative worker models, cell-based models, population microsimulation models, and general equilibrium models. This framework allows us to evaluate the suitability of alternative models for performing different policy analyses, as a models' suitability for an analysis is clearly a function of the questions one wishes to address. A model or
strategy that is appropriate for evaluating the effects of implementing individual accounts on private saving may not rank highly if one wishes to examine the effects on aged poverty of increasing the number of computation years in the Social Security benefit formula.

While we focus on comparative advantages of these generic types of models for addressing different aspects of Social Security reform, we frequently use prominent or innovative models from each group as examples. Examples make it easier to illustrate precisely how important economic relationships could be integrated into the models as now structured and as potentially extended. We raise both issues that are common to all or nearly all models and issues and challenges that are unique to particular strategies. Within each category or issue type, we try to focus on the importance of various issues for accurately simulating the specific sorts of OASDI reform issues and hypothesized consequences of such reforms described earlier.

While we present our discussions of these modeling strategies in a given order, generally from the least detailed in theoretical content to the most detailed, we do not mean to imply any inherent ranking of the model types for any given analysis. For some questions, simpler models may be the most cost-effective.

One particularly acute question common to all four types of models that we examine concerns appropriate baselines. If one is using a model from any of these classes to compare a baseline and a policy scenario, then one needs to consider what an appropriate baseline is for long-term analyses of the Social Security system. Current OASDI payroll and income tax levels with the current benefit structure is one candidate baseline, though it is clearly both unrealistic and unsustainable. Current law constrained to solvency is another option, but how one can best impose this constraint, through payroll or income tax increases, benefit cuts, or some combination, is an open question. The 1999 Technical Panel on Assumptions and Methods (1999) points out, for example, that Social Security law is ambiguous on how benefits are paid at the point when the Trust Funds run out of money. It recommends that the SSA actuaries thus produce estimates under two distinct scenarios: one in which current benefits are maintained and taxes are raised, and a second in which current taxes are maintained and benefits are reduced proportionately. Analyses that neglect the unfunded liability of the current Social Security system to future retirees, like those based on the assumption that current taxes and current benefits continue indefinitely, can be useful for illustrating the disparate impacts of alternative policies, but analysts need to qualify such results carefully.

1. Criteria

The Panel on Retirement Income Modeling (1995: 19-23) identifies a number of criteria for assessing the usefulness of microsimulation models. In the comments that follow, we draw extensively from these ideas, as well as from suggestions by Pattison (1999b). In describing the four classes of models suitable for micro-level analysis, we shall address the following common questions:
Representativeness and Completeness of Starting Data.

How good is the database that the model relies upon (starts from)? In making assessments about starting sample quality, we attach special importance to the validity of earnings and marital trajectories and correlations between spouses earnings over the life course given their importance for Social Security benefit determination. The file’s performance in accounting for pension and asset wealth may also be an important factor in some modeling strategies.

Cohort Size (Aside From Stochastic Replications).

Within a birth year group, how many people and/or family types does a model represent?

Generalizability.

What are the groups about which one can use the model to make inferences? This dimension of a model is, of course, clearly related to the first two dimensions that we just described. Using any model, one is limited to making conditional statements (e.g., “if a given set of assumptions about fertility, mortality and wage growth hold, then one can expect a particular outcome”). Whether one can legitimately make conditional statements about a group may depend upon whether the baseline population and updating and/or tabulation processes adequately account for differentials between that group and others in the population.

Ability to Oversample Parts of The Population.

Can one examine distributional differences within small population groups using the model? Many groups of interest for Social Security reform are relatively small. If a model permits oversampling, it can enable the user to consider the effects of reform on a small but targeted group.\(^{11}\)

Theoretical and Empirical Validity of Updating Algorithms (If applicable).

How good are the updating algorithms that the model employs? How consistent are they with best available social and economic research? Do they allow for forward-looking behavior by agents? Are these algorithms derived from real data? If so, what are the strengths and limitations of the data from which they are derived? Aggregate rates and/or adjustment factors, regression-based approaches, statistical match procedures, reweighting, or some combination of these four,

\(^{11}\) As computer processing and storage capacities increase, permitting models to employ larger starting samples, this capacity/criterion should become less important.
are alternative approaches to updating a model population. Each has associated advantages and pitfalls, although in theory the distinct strategies could yield similar results.

**Autocorrelation/Serial Correlation (If applicable).**

By what means does the model account for autocorrelation and serial correlation? Findings from social and economic literature suggest that effective models of many individual- and family-level processes and outcomes central to policy-oriented microsimulation models, like earnings paths over the life course, often need to incorporate complex error structures. Models that fail to account for correlation patterns in outcomes within persons (over time), within families, and even within birth cohorts or geographic spaces will require far greater calibration than models that do. Unfortunately, implementing these sorts of structures into model processes can be challenging both from estimation and computation perspectives.

**Feedbacks.**

Does the model contain linkages between individuals' circumstances and choices in early periods and their choices and circumstances in subsequent ones? If not (or if existing links are inadequate), how easily could these types of linkages be integrated into the model?

**Outcomes.**

About which Social Security outcomes (cross-sectional, longitudinal, aggregate, distributional) can one use the model to make inferences? As already noted, some of the Social Security outcomes of great interest to policy makers include measures of the long-term fiscal balance of the system, of how the system and proposed alternatives affect poverty rates, and of lifetime money's worth and redistribution for different groups in the population.

**Programmatic Detail.**

With what level of detail does the model replicate Social Security and other government programs (e.g., are SSI and DI payments incorporated, does Medicare eligibility interact with labor supply decisions)? What are the implications of these choices? Most of the micro model approaches that we review replicate the Social Security program in great detail, including calculation of AIME and PIA, comparison of competing eligibility criteria (e.g. retired worker versus spouse/survivor), and application of actuarial reductions and credits and the earnings test. More detailed aspects of Social Security, such as income taxation of Social Security benefits, children's benefits, or family maximum benefits are sometimes not explicitly modeled. In contrast to most micro models, those models that are more successful at integrating macroeconomic considerations tend to have only crude representations of Social Security benefits, often based on mean replacement rates. Could a more detailed representation of the Social Security program be added to these more aggregated models?
**Projection Assumptions.**

What assumptions are embedded in the model's projection period analyses? Many micro models are calibrated to the intermediate assumptions of the OASDI Trustees' Report. It is also quite common to assume that patterns in the most recent period are carried forward into the projection period. What, for example, do a model's assumptions imply about wage differentials between men and women into the future? Do gender and interracial mortality differentials decline, increase, or stay the same? If distributional analyses are the goal, then these choices about the projection period are critical. Theory and empirical research should inform them.

**Sensitivity Analyses.**

Does the model give analysts the capacity to explore the sensitivity of estimates to key assumptions, for example across high, intermediate, and low alternatives? The consequences of any changes to the financial picture of Social Security will surely vary based on demographic outcomes (including fertility levels and life expectancy) and economic trends (including real wage growth). Future values of these variables are uncertain, but will likely fall within a range of plausible alternatives. How easy is it to explore alternative outcomes?

**Stochastic Replications Per Individual/Cell.**

Does the model automatically allow a user to examine the sensitivity of estimates to a different series of random shocks (as represented by random draws) for each individual? If so, how many alternative draws? If not, with what ease and speed could multiple replications be conducted? Or, alternatively, are any methods employed that reduce random variability? One can conceive of virtually all life-course processes as resulting from a combination of systematic and stochastic forces. While what we have labeled as sensitivity analyses examine changes to the systematic elements in processes (e.g., the central tendencies in levels and predicted differentials across groups), stochastic replications and/or variance reduction concern changes to the individual random environment (e.g., whether the person is lucky or unlucky in the labor market that year, whether he or she experiences the onset of a disability). As model samples sizes increase, the importance of explicitly performing stochastic replications decreases, since stochastic replications are implicitly performed in simulating outcomes for large numbers of highly similar individuals.

---

12 Researchers have different views on whether or not variance reduction is desirable. Supporters of variance reduction techniques point to closer tracking of outcomes that a model's parameters imply. Opponents note that employing variance reduction techniques makes computing standard errors for an estimate of a population mean impossible.
Stochastic Replications of Aggregate Demographics/Economics.

Does a model allow the user to consider the consequences of random or cyclical fluctuations in demography, such as another flu epidemic or baby boom, or economic risks, such as a stock market crash, massive natural disaster, or oil embargo, on Social Security outcomes? Can the analyst weigh the likelihood of such shocks based upon their historical likelihood? Can one examine distributions of alternative assumptions? If so, how many replications does the model make? If not, with what ease and speed could one conduct multiple replications? Do correlation patterns in these assumptions follow patterns observed historically? Just as individual-level processes are characterized by uncertainty, so too are aggregate processes. Analysts (e.g., Lee and Tuljapurkar, 1994) have pointed out the dangers of relying on high, medium, and low alternatives for adequately expressing the uncertainty of estimates. One can avoid these pitfalls by considering joint distributions of assumptions rather than single estimates.

Ease of Use.

How large and detailed is the model? Is its source code easy to follow? Can an analyst trace the sources of programming bugs and apparent anomalies easily? How well documented is the model? Models which replicate many processes and rely upon thousands of lines of code may have advantages in terms of the level of detail with which one can examine the population, but they can also be quite difficult to debug and maintain.

Accessibility.

Is the model in the public domain? Can researchers outside of the model's parent organization easily access its source code or documentation of its key parameters? Have developers disseminated materials about the model's structure and assumptions widely? Model developers often have conflicting goals. Some developers are charged with maintaining the privacy of individuals represented in the data upon which they rely. Others intend to generate revenue/profits by marketing their models to outside users. Others still may hope to advance a political agenda through model estimates. While accessibility and openness of models facilitates their development, evaluation, debugging, and enhancement, they can compromise some of these other objectives.
Development and Maintenance Time and Cost.

How much does it cost to develop a model of this type? Are there significant maintenance costs? In discussing these issues, we present relative rather than absolute cost estimates, since each of the model types that we will examine can be developed in greater or lesser detail. Not surprisingly, the more detailed the model, the greater the time and expense required for development and maintenance.

2. Type 1: Representative Worker Models

Representative worker models are probably the simplest type of model that we examine. Researchers have used them quite widely to produce several influential analyses of Social Security's effects on a range of typical workers (see, for example, Steuerle and Bakija, 1994). The Office of the Actuary of the Social Security Administration frequently utilizes composite workers in its estimates of distributional consequences of major reform proposals (see, for example, Goss, 1997). In these models, workers earn either steady wages (minimum wage, average wage, maximum taxable wage, or some fraction of one of these three) throughout their career or else age-sex specific average or median wages. The models expose these stylized workers to the rules of the Social Security program annually so that in every year that they work, the individuals pay taxes to the Social Security system, and then at some arbitrary age (typically the normal retirement age) begin to collect benefits. Some of these models account for differences in life expectancy within and across cohort-sex groups, and the implications of these differences for survivor benefits, while others are more simplistic. To model changes across cohorts, the modeler adjusts the input parameters, including values of the wages in earnings trajectories and probabilities of death. These workers' lifetime contributions and benefits can be tabulated in a number of ways (e.g., lifetime transfers, internal rates of return, replacement rates) to reveal patterns in Social Security adequacy and equity across birth cohorts, sex, or stylized income/marital status groups.

A concern with representative worker models is that they could yield misleading estimates if observed worker profiles differ systematically from the hypothetical profiles they use. For example, these models rarely incorporate interruptions in work histories or divorces, events that are both prevalent in our society and consequential for Social Security outcomes. The usefulness of representative worker models could be enhanced through the use of multiple wage profiles that are more realistic and conform more closely to observed histories, as explored in recent work by Toder et al. (1999) and Burtless, Bosworth, and Steuerle (1999).

Representative worker models also lack the capability to replicate important aggregate totals such as Trust Fund balances. These models could conceivably, with some disaggregation and the application of weights to each profile, be useful for making broader, population-wide inferences. This option requires that the model population be comprised of composite individuals who reflect representative experiences.
Another shortcoming of existing representative worker models is that they rarely incorporate information about other government taxes and transfers (e.g., income tax and SSI), and even the full spectrum of Social Security components (especially DI). There is, however, no inherent reason why these programs could not be integrated in these models in future work.

Using this type of model, one represents a given Social Security reform by changing the value of existing program parameters, such as the payroll tax rate or replacement rates in the benefit formula, and then determines the consequences for the different worker/couple types. Reforms that integrate individual accounts are simulated by assigning the representative workers accounts with a given rate of return, annuitization structure, and rules about transaction costs. Several recent studies, for example by Beach and Davis (1998) and Shirley and Spiegler (1998), have used these types of models for exploring privatization proposals. These studies make highly simplistic assumptions about characteristics of work profiles. Burtless, Bosworth, and Steuerle (1999) provide a more sophisticated treatment. They use historical data on lifetime earnings to create representative workers with different average levels and patterns of lifetime earnings.

It is also possible to make certain ad hoc adjustments to worker profiles in representative worker models in order to account for anticipated feedbacks, whether of individual accounts or of policy changes for which a precedent exists (for example, increases in the retirement test exempt amount), but this is rare in practice. In some representative worker models, the user can examine the sensitivity of model outcomes to underlying economic and demographic assumptions.

While we have thus far discussed representative worker models as a separate type of model, in the tradition of the Steuerle-Bakija analyses, it is worth noting that representative worker analyses can serve as components of other modeling strategies. Numerous micro-level studies of Social Security systems trace out individual or family “case studies” (see, for example, Falkingham and Hills, 1995, Caldwell et. al, 1999, or Holmer, 1996) to illustrate individual marital, earnings, and tax/transfer paths over the life course. Further, many models in a variety of classes employ representative workers both for model development and debugging and to provide clearer illustrations of the effects of policy (see, for example, discussion of the Cohorts model in Pattison, 1999b). Software embedded in the DYNACAN model allows analysts to trace paths easily for any person in the model, and the Modgen development engine for the LIFEPATHS model has a “biography browser” that similarly enables a user to examine visually the life histories it generates (Gribble, 1997). It is possible to make these same types of
enhancements with any model that contains individuals, illustrating paths of all workers or selecting particular workers to profile because they have earnings close to a group average or because they are members of (and can thus represent) a target group of special concern (e.g., never married disabled mothers).

**Summary Of Representative Worker Models.**

*Representativeness and completeness of starting data:* Atypical, incomplete, families stylized, though some work is beginning to employ more realistic trajectories based on empirical estimates.

*Cohort size (aside from stochastic replications):* Usually less than twenty.

*Generalizability:* Minimal, though results serve extremely useful illustrative function.

*Ability to oversample parts of the population:* N/A.

*Theoretical and empirical validity of updating algorithms:* Minimal, although empirically derived worker profiles are a major improvement.

*Autocorrelation/serial correlation:* N/A.

*Feedbacks:* Typically none; could be integrated on an *ad hoc*, limited basis.

*Outcomes:* Present value of lifetime taxes/transfers (sometimes disaggregated by program), net and expressed as a percentage of present value of lifetime income, replacement rates, and internal rates of return.

*Programmatic detail:* Typically high for OASI, including benefit computation and actuarial reductions, but may omit details such as family maxima; few integrate SSI, Food Stamps, or Medicare, but there is no inherent reason why these cannot be included.

*Projection assumptions:* Typically limited to a narrow range of variables, including life expectancy and wage growth; model may constrain the OASDI system to solvency, thus explicitly assuming the nature and composition of actions government will take in future.

*Sensitivity analyses:* Often conducted employing Trustees' high, intermediate, and low assumptions, but rarely explore sensitivity of outcomes to features of the worker's path itself.

*Stochastic replications per individual/cell:* Usually one, but could in theory add random noise term (for example, based on the historical variance of earnings for a given subgroup) to earnings trajectory and resimulate; models with shorter run times have greater capacity for these types of analyses.
Stochastic replications of aggregate demographics/economics: Could, instead of using just high, medium, and low assumptions, resimulate using scenarios (combinations of assumptions) drawn from a distribution; models with shorter run times have greater capacity for these types of analyses.

Ease of use: A key strength of the strategy; models tend to run quickly, existing program parameters are relatively easy to modify, and results are summarized in convenient, understandable way.

Accessibility: Because they tend to have the smallest number of assumptions, developers can more easily express all major assumptions in a clear and concise format.

Development and maintenance time and cost: Typically the least costly of the strategies outlined, though more sophisticated versions with multiple, more representative earnings profiles and proper accounting for differential mortality and survivors' benefits require significant investment.

3. Type 2: Cell-Based Models

Cell-based models of the Social Security system vary widely in size and complexity. To construct a cell-based model, one divides a population into subgroups and then projects changes over time to the resulting aggregates. Subgroups may be fairly small, like a five-year age-sex-marital status category, but ultimately do not represent individual decision-making units. One assumes that all of the units within any given subgroup are identical. To move these groups through time and across states (for example, from unmarried to married), the modeler applies appropriate parameters (like a marriage or mortality rate or a tax schedule), estimated from the best available data, to the aggregates. Most models of this type compute Social Security benefits by relying on mean levels for each cell or by using simplified replacement rates. How benefits are computed has implications for the types of distributional analyses that one can conduct using a cell-based model. These models do not generate full distributions of Social Security outcomes, and this precludes certain analyses of winners and losers from policy changes.

The Social Security Administration Actuarial model, surely the most prominent of all models of this program, is a cell-based model. It has considerable detail on the population, sometimes projecting based on age-race-sex-marital status groups. It projects program benefits using indirect methods (rather than computing them from a person's own and his/her spouse's earnings histories). To cope with aggregate uncertainty, the Social Security Actuarial model employs three sets of assumptions about demographic and economic outcomes, detailed in Board of Trustees (1999). Critics complain that the SSA Actuarial Model lacks consistency among its many assumptions, and many researchers identify individual assumptions as either too optimistic or too pessimistic (see, for example, Lee and Tuljapurkar, 1997, on mortality assumptions).
Martin Holmer’s EBRI/SSASIM2 is another example of a cell-based model. The Holmer model incorporates macro-economic and macro-demographic uncertainty in complex ways, simulating hundreds of alternative scenarios using a wide variety of economic and demographic assumptions drawn from a distribution, thus producing interval estimates (see, for example, Holmer, 1996; 1997). It also integrates macroeconomic feedback through a growth model. Many view its distributional detail as inadequate, because micro-units are differentiated only on the basis of cohort and sex (i.e., there are only a few typical persons of each sex in each cohort).

The Macroeconomic-Demographic Model (MDM), developed by Capital Research Associates for the National Institute on Aging (Anderson, 1990), is another example of an important cell-based approach. The population of MDM is quite diverse, with hundreds of family types represented, though the model does not operate on a micro-level data base. It also incorporates a neoclassical macro-economic growth model that links households with the business sector.

A major limitation of cell-based models is that they do not typically incorporate histories at either the individual or family level. One thus cannot observe this specific aspect of heterogeneity of experiences within cells, heterogeneity that could be pivotal with changes like earnings or benefit sharing or private accounts with different fee and contribution structures. Even the effects of relatively simple changes to provisions like the number of computation years that determine a Social Security benefit can only be simulated indirectly in a cell-based model, leaving no real opportunity to draw inferences about distributional impacts of the proposal, except perhaps across age-sex groups.

Cell-based models nonetheless have several comparative advantages relative to other model types. Developers often express their assumptions in a relatively parsimonious way; less aggregated strategies tend to rely on a greater number of more detailed assumptions. Computational demands, and thus expense, are typically less significant than for the general equilibrium and population microsimulation approaches that we will discuss next. As Bongaarts notes these types of models are “likely to be preferred in cases where accurate projections are required of relatively rare events or of family characteristics that occur infrequently” (1983: 34). There are important Social Security applications for which this is the case. Cell-based models also may have a comparative advantage for comparing aggregate amounts, including Trust Fund balances. Only very well-calibrated dynamic microsimulation models (i.e., models that represent the full Social Security area population and that accurately account for nearly all of the wages subject to Social Security tax both in the past and into the future), of which few if any are currently in operation, are possible competitors in this regard.

13 The thirteen key assumption parameters in the SSASIM2 model include: disability incidence and disability recovery rates; female and male labor force participation rates; the hours worked growth rate; the inflation rate; the mortality decline rate; net immigration; nominal interest rate; productivity growth rate; the total fertility rate; the unemployment rate; and the wage share growth rate.
Summary of Cell-Based Models.

Representativeness and completeness of starting data: Varies; may represent entire population through thousands of cells, or through as few as two per cohort; usually do not contain families, though simplifying rules (e.g., husbands and wives ages reflect average differences in ages observed among couples in the population) may be used in some tabulations.

Cohort size (aside from stochastic replications): N/A; although full cohorts may be represented implicitly, individual agents/families are never explicitly represented.

Generalizability: Can sometimes generalize about groups as small as an age-race-sex-marital status or beneficiary type group.

Ability to oversample parts of the population: N/A.

Theoretical and empirical validity of updating algorithms: Rates reflect observed patterns, usually holding constant structure; extrapolations may not always pick up compositional changes within an aggregate (e.g., changes in educational attainment).

Autocorrelation/serial correlation: N/A.

Feedbacks: Sometimes integrated through macro functions (e.g., neoclassical growth model) in a uniform way for all cell members.

Outcomes: Often both cross-sectional and longitudinal group means (groups may be highly disaggregated), no distributions within cells; no individual-based longitudinal information (earnings histories, marital histories); often quite successful with cost aggregates.

Programmatic detail: Variable; programs typically represented by participation rates and either mean benefit levels or replacement rates; integration of OASDI and DI may be captured, though integration of Medicare and/or SSI is rare.

Projection assumptions: Variable; specific to cells.

Sensitivity analyses: Variable; frequently consider alternative age-sex patterns in rates.

Stochastic replications per individual/cell: Usually one.

Stochastic replications of aggregate demographics/economics: Variable; can be as detailed as stochastic simulation with hundreds of alternative assumptions.

Ease of use: Variable; some require extensive training, others simpler and user-friendly.
Accessibility: Variable; some are proprietary; aggregate assumptions frequently very accessible, but detailed assumptions often less so.

Development and maintenance time and cost: Variable; as great as many person-years.

4. Type 3: Population Microsimulation Models

Population microsimulation models start with a broader population than models in these prior two classes. 14 Typically, the starting sample is a representative sample or subsample of individuals in families and/or households from a census or survey that is matched to administrative records which provide the individual's earnings histories, though some public-use models rely on imputed rather than administrative earnings records. (Using a census or survey is not a strict characteristic of this strategy. One could develop a population microsimulation model using an entirely synthetic population, though presumably one would want this population to share, at least roughly, the characteristics of some current or historical population.) The model then ages this population, usually a cohort or series of neighboring cohorts. Aging generally occurs in discrete steps, such as a month or year, though some models of this type rely on time until next event as the time unit in some or all of their modules. Klevmarken (1997) identifies six different ways that a population may be aged, depending upon the sequencing of model processing loops. 15

Through the chosen aging process, the model simulates life histories for each decision unit, whether an individual, family, household, or even firm, 16 in the population separately. The life histories generally incorporate the outcomes of economic, demographic, and social processes, and a dynamic microsimulation model typically employs many separate procedures to model these outcomes for different subgroups of the population. 17 Aging procedures span a variety of

14 One can trace the origins of most models in this larger school to the work of Guy Orcutt. Orcutt first laid out the mechanics and rationale for this strategy (1957), and then was involved in the development of an early prototype model (Orcutt, Greenberger, Korbel, and Rivlin, 1961) and also the first large-scale forecasting model of this type (see Orcutt, Caldwell, and Wertheimer, 1976).

15 Klevmarken distinguishes between the “time-period,” “process,” and “individual” loops, any one of which can be inner, middle, or outer in the simulation.

16 While microsimulation models of firms do exist, see, for example, Eason (1997), and were indeed part of Orcutt’s vision for microsimulation, we do not discuss them here.

17 Distinct procedures for subgroups of the population allow the effects of all important factors to vary (as opposed to including interaction terms for just a select fraction of variables). This is necessary when one believes that given processes differ fundamentally across social groups. In such a case, a variable that is predictive of outcomes for one group may be irrelevant for determining the outcomes of a second group and may act in an opposing direction for a third group. For example, one might speculate that the presence of young children in a household influences the labor market behavior of men and women in distinctive ways, increasing the probability that a man works, but decreasing the probability that a woman works. If this were the case, a model might be more effective with separate equations.
techniques, ranging from the purely deterministic (i.e., based strictly on rules), to partially structured and partially stochastic (e.g., based on equations with systematic components and Monte Carlo components), to strictly stochastic (i.e., based solely on random draws). Aging procedures may be as simple as reweighting (the typical "static" microsimulation approach), or as complex as an imputation of some attribute/characteristic through a cross-sectional regression or the application of a hazard function. Most models of this type apply aging algorithms sequentially by process (in a modular fashion) within their time-unit loop, though there is some variation.

Unlike the cell-based or overlapping generations models, these population microsimulation models simulate much more detailed characteristics of the population. These models typically incorporate detailed rules of the Social Security program including features such as the program's early retirement deductions and delayed retirement credits, the earnings test, family caps, and disabled worker benefits, many of which are not always included in representative worker models. These detailed models are capable of integrating a wide array of program interactions, for example between the DI and OASI components of Social Security and between Social Security and SSI or the Food Stamp program. They have the further advantage of permitting analysts to explore full distributions of cross-sectional OASDI outcomes rather than just expected values and/or aggregates. In some cases, it may be difficult to compute lifetime measures using such a model because of the portions of lifetimes (and fraction of the population) that are not observed, for example, individuals who die before the survey that serves as the starting population. Analysts can compute aggregate longitudinal outcomes by adding missing populations (for example, persons who are residing in institutions, as they are also excluded from many prominent survey files).

Most any sort of Social Security trajectory that can arise in real life, for example a spell of disability followed by re-entry to the work force and then retirement, can be replicated using a population microsimulation model. One can model programmatic Social Security reforms, then, by changing program parameters, and one can model structural reforms like individual accounts by introducing the new accounts with assumed contributions, returns, and requirements. One can simulate complex reforms affecting couples' benefits (like earnings sharing) using many models of this type because they usually generate full trajectories of life events, including work, marriage, divorce, and death. In principle, in many of these models, the model user could allow people in the model population to change their behavior in response to a reform. As most of these models are currently designed, this can be implemented in at least two ways: ad hoc, top-down adjustments to alignment targets, and bottom-up integration of more sophisticated micro-equations with economic parameters. Anderson (1990) and Baekgaard (1995) discuss these issues, as does Chapter 2 of this report.
In evaluating reforms to Social Security using this type of model, one has a rich opportunity to see the consequences of benefit calculator types of reforms (see, for example, Sandell and Iams, 1997, Historical Cohort Model estimates in Fontenot, 1999). Proposals to set up individual accounts can be integrated with precise detail on parameterizations (e.g., alternative annuitization and indexation assumptions) and variance in returns based on a person’s observed characteristics. While in theory certain microsimulation population models can provide information about nearly all of the types of outcomes listed in Table 1, calibration of the model sometimes favors one type of outcome over another. For example, in some models emphasis on reaching cross-sectional targets (e.g., national earnings aggregates and subgroup mean wage rates) may lead to less emphasis on longitudinal validation (e.g., the validity of earnings sequences). There are thus tradeoffs among these outcomes, regardless of the type of micro model that one chooses, and an attendant need to prioritize among them.

Despite their richness, population microsimulation models have encountered criticism for lack of behavioral content and inadequate integration of feedbacks between processes. At the individual and family levels, these models have rarely employed economic models of optimizing behavior in the presence of constraints, though in principle they can incorporate elements of rational decision-making or individuals’ use of “rules of thumb” in some processes (e.g., retirement timing). These models contain no firms, though users can sometimes integrate firm behavior indirectly through links to macroeconomic models (see, for example, Wertheimer, Zedlewski, Anderson, and Moore, 1986).

An additional potential limitation of these models is their complexity. Many population microsimulation models produce a tremendous, indeed overwhelming, amount of output that needs to be organized and summarized in a clear manner. This places a heavy burden on the user to develop appropriate tables of cross-sectional and longitudinal outcome measures. More importantly, analysts may not have a clear idea why the model produces certain qualitative results, that is, what specific assumptions about processes are producing the outcomes that one observes.

Another difficulty with population microsimulation models is that, to date, there is no accepted method for describing the variability of the outcomes that they produce. This variability arises from a number of sources, including sampling variation in the starting population, errors from imputation of statuses to members of the population, error imposed though updating based on uncertain parameters (both micro-equation parameters, like regression...
coefficients and aggregate parameters, like alignment targets), and pure random noise (Monte Carlo variability). Some researchers have identified ways to capture different components of variability. Cohen (1991) focuses on general issues, while Betson (1990) describes the parameter variety.

Many microsimulators argue that differences between baselines and reform scenarios may not be sensitive to sampling errors in either parameters or starting populations, and that we can thus rely on these differences (or “deltas”) even when baselines are imperfect and/or subject to considerable variability. This argument is more compelling if one is modeling a reform that does not interact with whatever errors exist in the baseline. It is of course less of a problem if errors that are present in the simulation are randomly distributed (and thus cancel one another out) than if errors are systematic. This argument does not, however, diminish the concern that point estimates from a microsimulation may in isolation be misleading. Wolf’s (1995) on-going research on interval estimation in microsimulation modeling could help to address this concern. Unfortunately, consumers of microsimulation estimates often prefer point estimates despite their limitations (see, for example, comments of Myers in Social Security Advisory Board, 1997).

There is a great deal of heterogeneity in the models within this group. One can distinguish these models from one another based upon their characteristics along several core dimensions. These dimensions include: the definition of the model’s population (whether it is a single cohort, a series of cohorts, or a full population); the nature of the model's aging algorithms (rewriting, or static aging, as opposed to dynamic or semi-dynamic aging) and the modeler’s corresponding assumptions about the nature of core social and economic processes; the structure of the model's marriage market (and thus whether the population is “open” or “closed”); the processing order of the three model loops identified by Klevmarken (1997); the extent to which the model relies on alignment/calibration; how the model handles individual-level stochastic elements (implicitly or explicitly, with or without replication); when the model starts simulating outcomes (for example, in the recent past or the more distant past); and the time unit/dimension along which the model simulates (for example, a year as opposed to time to next event). We next discuss some of the technical dimensions associated with several of these distinguishing characteristics.

18 Full population models typically age individuals from a real sample unit, while cohort models create a population (often hypothetical) exposed to current (i.e., “birth”) year demographic and economic conditions (Falkingham, Harding, and Lessof, 1995: 64-65). Falkingham, Harding, and Lessof hail cohort models because every sample member has a complete life history, a feat that can be achieved for given cohorts within a population microsimulation model but never for the entire model population. Often, one chooses to construct a cohort model rather than a population model due to limited availability of data. Instead of requiring data to estimate transition parameters for dozens of cohorts, one needs only the relevant information for the single cohort.
Aging Algorithms.

By aging algorithms, we refer to the methods in which a modeler transforms a microsimulation sample population at time \( t \) into a sample population at time \( t+n \), where \( n \) is an arbitrary number of periods (later in time).

**Statically aging** a population entails reweighting it. Members of the population at time \( t \) serve to represent members of the population at time \( t+n \). The only difference is that persons with given characteristics, those by which one is reweighting, may occur with a different frequency (either lesser or greater) in the later period. When one employs static aging of a population, one thus implicitly assumes that members of the later population do not differ from members of the earlier population in any systematic, important ways along the characteristics by which one is not weighting at the two time periods. The advantage to using static aging is that it is straightforward and relatively easy and inexpensive, in computational terms, to implement. Where one expects feedbacks or one is modeling processes that concern accruals and changes to accruals over a lengthy period (say, half a lifetime), static aging may be limited relative to the alternatives. Describing when static microsimulation is an inappropriate technique, one analyst writes: “When medium-term or even long-term consequences also come to bear besides the effect of the first round, static microsimulation is not sufficient” (Helberger, 1986: 115).

**Dynamic aging**, in contrast, requires explicitly replicating the processes (e.g., birth, death, migration, marriage) that contribute to the change in a model population between \( t \) and \( t+n \). In dynamic aging, processes are frequently solved sequentially on the basis of whatever time unit the model simulates along, rather than on a lifetime basis. In such a case, all processes can depend on all other processes with a one period lag, and on processes which precede them in the processing loop contemporaneously. Dynamic aging is likely to be preferable to static aging, as noted above, when one expects that there are many interdependencies/feedbacks between processes, and less important when interdependencies/feedbacks are fewer and/or less significant.

Open Versus Closed Populations.

The distinction between open and closed models appears in some of the early literature on microsimulation modeling (see, for example, Orcutt, Caldwell, and Wertheimer, 1976: 355-356). (Other chapters of this report discuss open and closed models in different contexts.)

---

19. To illustrate, let’s assume a model with five aging processes (a through e, solved sequentially in alphabetical order) that simulates for five time periods (t1 through t5). A modeler can specify the outcome \( c \) in time t5 as a function of processes a and b in time periods t1 through t5, as well as of processes c, d, and e in time periods t1 through t4. He/she can specify outcome \( d \) in time t5 as a function of processes a, b, and c in time periods t1 through t5, as well as of processes d and e in time periods t1 through t4. Finally, one can express outcome \( e \) in time t5 as a function of processes a, b, c, and d, all in time periods t1 through t5, and of e in time periods t1 through t4.

20. In one context, “closed” refers to the fact that a model “endogenously [derives] the changes in individual
microsimulation literature, a closed model is one in which marrying members of the sample find spouses from within the model population. In an open model, spouses are imputed, either sampled from within a host population, with some spouses possibly re-selected, or generated synthetically. Wachter, Blackwell, and Hammel note that one “essential feature of closed models is that every partner has his or her own network of kin, so that kinship linkages are not truncated at marriage nodes, as they are in open models” (1998: 9).

One key advantage of a closed model is that it imposes some consistency on the population. At any point in time, for every marrying man in the population, there is a unique marrying woman. Another advantage is that individuals must cope with marriage market conditions, albeit in a rather stylized way. As in real life, people do not always get to marry a person with the characteristics that they most prefer (where preference is defined based upon historical covariation in the characteristics of mates) when a model is closed. They must choose their spouses based on who is available in the population, and therefore need to relax their criteria in situations where there is a shortage of appropriate mates. For example, if men prefer to marry women who are three years their junior, most can have their preference when cohorts are roughly similar in size, but many will need to adapt their preference if a large cohort is followed by a small one (for example, in the U.S. context, men born in 1964 and women born in 1967). The kinship linkages identified by Wachter and colleagues as integral to closed models could be useful for certain Social Security analyses, for example, any study of the lifetime impacts of child survivor benefits. A final advantage to closed models is that this consistency may, to a certain extent, make closed models easier to explain to lay audiences. To describe a spouse as an attribute of another person, rather than as a person him/herself, requires a good deal of sophistication in the model consumer.

Closed model marriage markets have disadvantages as well. When marriage markets are small, the user may be forced to make many “bad” or implausible matches. Even more fundamentally, there is no basis for empirically estimating preferences when market conditions are different than historical conditions (though this challenge applies to any matching model). Another concern with some closed models is that the mapping of kinship linkages can be somewhat contradictory.21 If the starting population from which one begins the simulations is a sample rather than a full population, then the model may not be balanced with respect to kinship linkages other than spouses (e.g., one's parents and siblings are not included in the base population). Closed models are also computationally very expensive, with the marriage portion of the model often accounting for a significant fraction of simulation time. Innovative work by Suen and Wu (1999) suggests one method for reducing the computational demands of mate matching.

---

21 We thank Douglas Wolf for bringing this point to our attention.
The corresponding advantages of constructing a microsimulation model with an open marriage market rest mainly in processing time (see, for example, discussion in Gribble, 1997: 2). Using a model with an open population, one can oversample populations in a more straightforward way than one can in a closed model, and one can run the model on smaller samples. Some models with closed markets simply cannot operate on small starting populations (of, for example, less than several thousand persons).

**Simulation Time Dimension.**

Both in estimating aging parameters and in simulating future trajectories, one may use either a discrete (for example, a month, a quarter, or a year) or continuous (instantaneous) time unit. Galler (1997) comprehensively discusses the respective merits of discrete- and continuous-time estimation and simulation approaches. Here, we summarize just a few points. Alison (1984) notes that, with respect to parameter estimation, where independent variables are time-varying, relative costs and conveniences of discrete- and continuous-time event history analysis tend to be quite close. The smaller the discrete interval, the closer the results that the two methods yield. Whether one chooses discrete or continuous time is likely to depend on a number of factors including, for example, how frequently an event tends to occur within alternative discrete units. Rendall (in Cohen et al., 1999: 24) points out that a discrete-time simulation approach tends to allow greater flexibility in specifying relationships in component regression equations in a microsimulation model as time-varying than does a continuous-time model. In models that employ calibration, a discrete-time formulation may have particular advantages, given that aggregate data on event counts in a given time unit are more readily available than data on spell lengths.

**Model Combinations.**

Models combine features of these different dimensions in many distinctive ways, and quite a few models are even "hybrids" within several categories (e.g., they use static aging for one process and dynamic aging for another, or they simulate one process on a yearly basis, while simulating other processes in continuous time on the basis of a competing risks formulation). We discuss several models and combinations of model characteristics below. Once again, while we present these models and types in a given order, roughly based on ascending levels of complexity, we do not intend to imply a necessary ranking among them. Simpler models might be cost-effective in some instances.

A first microsimulation model, the Social Security Administration’s Cohorts model (for a description, see, Pattison, 1999a, and the summary discussion in Chapter 2), begins with the fundamental assumption that researchers can use a simulation basis cohort over and over, with suitable adjustments, to represent a sequence of birth cohorts (Pattison, 1999c). The Projected Cohorts Model employs data on the nearly completed earnings and marital experiences of a recent
cohort (the 1930 birth cohort, roughly the 1992 Social Security entitlement cohort) as its current means for projecting trajectories for future birth and entitlement cohorts. To transform the cohort to represent later ones, the model’s developers scale and re-weight these historical data and, for groups experiencing the largest changes (e.g., women with respect to labor force participation), employ regression techniques. Chapter 2 provides an extensive analysis of this approach. While there may be, for example, more women with a post-college education in the 1960 cohort than in the 1930 cohort, one assumes that the earnings and marital trajectory of a women college graduate in the earlier cohort can be fairly easily modified to represent the earnings and marital trajectory of a woman in the later cohort. The population of the Cohorts model is open.

One central advantage to this approach is that correlations within and between processes for full lifetimes are “real,” rather than synthetic. A complex series of interacting processes, including labor force entries and exits, marriages and divorces, and onsets and terminations of disabilities, influences Social Security outcomes. In other types of microsimulation models, one needs to impose explicit assumptions about whether and how the systematic and random components of these many processes are related. This makes other microsimulation models more complex, and as consequence, typically more expensive to debug and maintain. Nonetheless, this reweighted cohort approach incorporates, essentially, equally as many implicit assumptions about these correlations between processes as do those models which explicitly represent them. Usually in this strategy, one assumes that the correlation patterns have remained constant across cohorts. Estimates of a model of this type are not subject to Monte Carlo variability. Data requirements for constructing a model of this type may be less significant than the requirements for other microsimulation types, as adjustments tend to be cross-sectional (e.g., scaling to account for wage growth), rather than longitudinal. Another advantage to this strategy is that simulation run times tend to be shorter, and this may facilitate certain modeling enhancements (e.g., stochastic replications).

The Cohorts model does not currently integrate forward-looking behavior, though developers report plans for future work in this area (see Pattison, 1999b). The model does not currently attempt to represent a full population (i.e., all cohorts alive in a given year), though developers report plans to change this and note that the desktop version of the model is designed to simulate full populations (see Pattison, 1999b and 1999c). The model does not incorporate new entrants to the population (e.g., immigrants), though developers suggest that the model may have some advantages for incorporating immigrants relative to other microsimulation models (Pattison, 1999c). Feedback may be difficult to integrate into a model of this sort, and Chapter 2 addresses this issue with specific respect to the Projected Cohorts Model. Because they do not explicitly replicate many of the outcome-generating processes about which one might wish to test sensitivity (e.g., marriage and divorce, pension and wealth accumulation), these models often have less extensive capability for sensitivity analyses than many cell-based or dynamic microsimulation models. This approach does, though, permit equally extensive sensitivity analyses for any of the processes that it does model. For example, as developers of the Cohorts model adjust individuals’ earnings and labor force participation rates and set mortality parameters, users can fairly readily
alter wage growth assumptions or the actuarial assumptions in present value calculations. One would, however, have difficulty adjusting other model assumptions (e.g., marriage or divorce rates or benefit take-up patterns).

The Social Security Administration's MINT model, a hybrid along many important dimensions, is a second microsimulation model that relies on a different approach. (For detail on the economic assumptions in the current version of MINT, see Toder et al., 1999. For detail on the model's demographics, see Panis and Lillard, 1999.) In “aging” its population, MINT starts from the assumption that filling in the partially completed life histories of cohorts currently in their prime working years is preferable to relying on the completed histories of members of earlier cohorts. Its aging functions rely on the fundamental presumption that demographic and economic events that a person has experienced in the past and is experiencing in the present influence both the chances that other decisions/events will occur and the ultimate future levels of outcomes for the person. MINT relies heavily on longitudinal data sources, for example the Panel Study of Income Dynamics and the Survey of Income and Program Participation, for parameters in its updating functions. To fill in the unobserved portions of life trajectories, MINT employs a combination of cross-sectional regression equations (functions in which attributes in the current period predict a given outcome in the current period) and processes that are more longitudinal in nature (functions in which attributes at the present time and at various points in the past, thus including one's duration in given states or one's risk status, predict a given outcome in the current period).

Using a competing risks formulation, MINT solves all demographic outcomes (marriage, divorce, and mortality; there is no treatment of fertility or immigration) for all sample members first and then solves each economic process (e.g., earnings, pensions, wealth) for all persons in the sample for a full lifetime before solving a new process. Events in prior demographic and economic processes in the sequence can thus influence outcomes in subsequent processes, but outcomes from later economic processes can influence earlier outcomes in only a general way (e.g., permanent income influences the stability of one's marriage, but income fluctuations do not). Individual stochastic elements are incorporated in the model based on single realizations

---

22 That is, the pure dynamic aging case that we describe above is made up of dynamically recursive blocks, while MINT (like many other microsimulation models) currently has a number of completely recursive blocks. (For more detail on this distinction, see the discussion in Sabelhaus, 1999).

23 Using the same notation as in footnote 16, in MINT one can specify economic outcome c in time t5 as a function of processes a and b in time periods t1 through t5 and of process c in time periods t1 through t4. One can specify outcome d in time t5 as a function of processes a, b, and c in time periods t1 through t5 and of process d in time periods t1 through t4. Finally, one can express outcome e in time t5 as a function of processes a, b, c, and d, all in time periods t1 through t5, as well as of process e in periods t1 through t4. Note that for the last process in the last time period, MINT permits an identical set of determinants to the dynamic model described above. One large difference is that outcome e effects all other outcomes in periods t2 through t4 in the dynamic case, but not in the case of MINT non-demographic elements.
of random draws. MINT’s marriage market, which is better described as an earnings imputation for the earnings of an unobserved spouse (i.e., a spouse who left a marriage to a sample member before the survey or entered into a marriage to a sample member after the survey), is open, in that a member of the sample can serve as a spouse to more than one person at a time.

MINT uses little calibration in its replication of micro-level processes. With calibration of certain of its aggregate outcomes, sensitivity testing for some processes could become more straightforward when using the model. While not employing calibration procedures globally, MINT specifically does employ parameter adjustment in its mortality module (see Panis and Lillard, 1999). To ensure that their PSID regression estimates, which are based on a relatively small number of deaths, would track national patterns, MINT’s developers estimated an additional regression on similarly structured spells from Vital Statistics. They then chose to use the difference between the micro-level and Vital Statistics estimates with just a few regressors (age slopes, calendar time trend, and race) as an additive adjustment to the micro-level estimates on the full model (which includes additional socio-economic variables, like education, marital status, and permanent income). To the extent that it can be generalized, this method is very promising for adjusting estimates in microsimulation population models of a variety of sorts. It differs from alignment/calibration in a number of other microsimulation models in that one adjusts several parameters in the regression equation (the intercept, age slopes, time, and race), as opposed to a single parameter (the intercept term), to ensure a better fit to aggregate data. Presumably, alignment at another level (for example to enable sensitivity testing) would not need to make as large adjustments given this first attempt at micro-level calibration.

MINT has great advantages for many types of Social Security analyses, especially in the relatively near term (for example, up to twenty years into the future), the period for which the model was designed. Because such large fractions of life trajectories are based on historical data for the exact person in the sample, such a strategy could have special advantages in comparing particular subgroups across cohorts. MINT has an extremely rich starting data base (the Survey of Income and Program Participation matched to administrative records) that validation studies have revealed accounts for a larger fraction of income and wealth than many of the prominent alternatives. MINT is somewhat more parsimonious than most dynamic microsimulation models, potentially allowing for somewhat easier use, debugging, and training of new users than more elaborate models.

Among the disadvantages of MINT are its current inability to integrate forward-looking behavior, the lack of behavioral relations in its updating algorithms, and its fairly lengthy simulation time. Because developers have restricted the analyses to certain cohorts and because the model does not represent all aspects of population growth and change, users cannot currently consider aggregates like Trust Fund balances using this model. Researchers could, though, compare costs of proposals that affect the age cohorts that are under study. MINT developers could in principle improve on these limitations in future work by adding institutionalized populations and fertility and immigration modules and, also, by extending the model’s projection
period (beyond 2020). Developers currently do plan to extend MINT considerably by integrating new modules of health, job demands, living arrangements, and the Supplemental Security Income program and by refining the model’s treatment of earnings, retirement, pensions, and non-pension wealth (Social Security Administration, 1999). These enhancements may include a change to a more dynamically recursive solution structure (from the current completely recursive structure) for the model’s economic processes.

A large group of microsimulation models, the most common for large-scale Social Security models in the U.S., has annual aging, a recursive structure, a closed marriage market, and heavy reliance on alignment. Because these models employ dynamic aging, they are frequently referred to as dynamic microsimulation models. Here we will refer to them as “DYNASIM-style” models, after the first large-scale model of this type.

Steven Caldwell's CORSIM, a DYNASIM derivative, the Canadian government's DYNACAN model, the Urban Institute's DYNASIM2, and the Lewin Group's PRISM are all examples of DYNASIM-style dynamic microsimulation.\footnote{Information on these models is available in, among other places, Caldwell (1996), Office of Superintendent of Financial Institutions (1997), Zedlewski (1990), and Kennell and Shiels (1990), respectively. Anderson (1999) provides an up-to-date, detailed, and comprehensive overview of each model as well. The Congressional Budget Office is currently in the process of developing a large-scale micro-level model that will, most likely, fit into this school as well (see Sabelhaus, 1999).} Each model starts with a different population sample and covers different processes (e.g., some models include modules of pension and/or wealth accruals while others do not). These four models simulate events in discrete (i.e., yearly) intervals, in contrast to some models in a closely related class (e.g., LIFEPATHS, described in Gribble, 1997, and DYNAMOD, described in Antcliff et al., 1996) that have employed continuous-time simulation techniques. Another example of dynamic microsimulation, though not a DYNASIM-style model, is the innovative microsimulation approach Wolf and colleagues (1995) employ. The model that Wolf et al. are building cannot currently be used for Social Security analyses, but this project is wrestling with many of the fundamental challenges for dynamic microsimulation, including treatment of the uncertainty that is inherent in dynamic microsimulation approaches.

To an even greater extent than the MINT model, DYNASIM-style models rely on the fundamental presumption that earlier demographic and economic events shape individuals' future trajectories. Economic and demographic characteristics influence the chance of experiencing an important life event, like getting married, going to college, starting work, or having a baby, and these life events trigger changes in behavior (e.g., married women with children have different labor market behavior than women without children). Like MINT, DYNASIM-style models rely heavily on longitudinal data files for the micro-dynamic aging algorithms that trigger these changes. They also tend to rely heavily on aggregate data to estimate alignment/calibration parameters, thus synthesizing aggregate and micro data.
As with the other microsimulation models we have discussed, DYNASIM-style models typically do not incorporate forward-looking behavior. The DYNASIM2 model's retirement timing algorithm is a noteworthy exception to this pattern, employing forward-looking behavior among prospective retirees who compare the benefits of retiring at time t to the benefits of retiring a year later. Sometimes models of this type rely on multiple realizations of chance elements, but more typically developers use a strategy of assuming that a single realization is adequately representative. Several of the models that use single replications (e.g., DYNASIM and CORSIM) rely heavily on variance reduction techniques (methods that ensure that the expected number of events for a process is realized within one event).

Dynamic microsimulation models have several key advantages compared with the prior three types of models and statically aged microsimulation models. One of the unique strengths of this class of models is the recognition of the ability of policy to influence demographic and economic behavior. Most of the other model types treat demographics as exogenous. In DYNASIM-style microsimulation, individuals and families can alter decisions about any life course process, including marriage, childbearing, and divorce, depending on what happens elsewhere in the model. For example, changes in the level and timing of women’s education and childbearing, similar to those we have witnessed in recent decades (see, for example, Rindfuss, Morgan, and Offutt, 1996), explicitly affect their earnings histories. (Note that the earnings projections in MINT, a less fully dynamic model, explicitly integrate educational differentials, but not fertility or marital status ones. Such differentials may be partially integrated in an implicit way through the models' use of a complex error structure and prior period residuals.) The reverse can also be true. If tax changes, for example, lead the user to forecast higher labor force participation among women, women’s fertility can change in later periods. More income for women can lead to higher divorce rates, and more income for men can lead to higher marriage rates. Whether the responses built into a DYNASIM-style model are appropriate depends on whether the developers' assumptions about the underlying processes are valid. Correlation in a past period cannot, of course, be viewed as documenting a causal relationship, but some variables (e.g., educational attainment) may serve as adequate proxies for the theoretical variables (e.g., human capital) in a

---

25Gribble (1997) has made important headway into estimation of Monte Carlo variability of estimates through the innovation of, for each piece of tabular output a user requests, dividing the sample into sixteen statistically independent subsamples. The model then performs the tabulations, enabling the analyst to compute standard errors for the cell using the sixteen observations, a feature that developers of other Monte Carlo models might wish to incorporate. Some models carrying around variables that enable similar estimates, but do not routinely conduct variability analyses (see, for example, Pattison 1999c on retention of the CPS rotation group number in the Cohorts model). According to Anderson (1999: 4-13) the PRISM model, by way of contrast, computes all outcomes twice. This does not enable the user to compute a standard error for various cells, but reduces the probability of an outlying distribution of random draws. While other models do not explicitly treat Monte Carlo variability, their users have often conducted multiple runs of baselines in order to estimate its importance (see, for example, Orcutt, Caldwell, and Wertheimer, 1976, Chapter 11, or Favreault, 1998, Chapter 5).
more elaborate model in which we have more confidence. Developers must take care to consider how compositional changes may translate into changes in the structures of underlying processes.

Another interesting and valuable characteristic of these models that the prior types do not share is that they often retain intergenerational linkages. For example, in certain models, an analyst can link an aging parent in certain cohorts with his/her adult children. These kinship links could be useful for an analysis of Social Security because of the relationships between caregiving, co-residence (e.g., returning to live with children in old age), and economic well-being. Rendall and Speare (1995), for example, find that co-residence reduces aged poverty significantly, especially for women. If a model retains kinship linkages, then its developers can more readily model such transitions as moving to a child's home.

Many analysts criticize the current group of DYNASIM-type microsimulation models for their lack of behavioral and theoretical content, particularly with respect to the macroeconomy, and also for their lengthy processing times and great development costs. While the first concern remains an important issue, the second has diminished considerably with the advent of much cheaper computing and the existence of many templates. Another criticism that has been levied against microsimulation models is that they do not have a “Social Security focus” and that they often have modules (e.g., state-to-state migration algorithms) that require staff time to maintain, but do not contribute significantly to Social Security analysis (Pattison, 1999a). This criticism is not necessarily valid, because such modules are not an inherent feature of the strategy, and there is no reason why modules not relevant to Social Security need to be developed or maintained. Again, though, we stress the view that because Social Security interacts with so many other dimensions of the economy, including the tax system, social assistance programs, and employer-provided pensions, in many instances elaborate models are necessary for exploring distributional aspects of reform.

DYNASIM-style microsimulation models rely on alignment/calibration to very different extents. Calibration takes very different forms in the models, with important implications for the ease with which users can integrate policy feedbacks. Two schools of thought dominate discussion of the use of alignment procedures in dynamic microsimulation models. Many contend that alignment is an essential way to control outcomes and thus a means for implementing assumptions, and that alignment factors are a legitimate set of parameters estimated from additional data (see, for example, Wilde, 1997). From this perspective, alignment is a less-computationally intensive alternative to full-information maximum likelihood estimation of the entire system of model equations under a set of constraints. Further, alignment can prevent small errors from snowballing into larger ones, because it ensures that successive processes have correct inputs. Other researchers believe that it is not legitimate to adjust micro-equation parameters (typically intercepts) to fit external targets. They believe that the presence of alignment factors signals a need instead to respecify the underlying outcome models (see, for example, Klevmarken, 1998). Any unexplained variation captured through alignment, critics argue, can become an even bigger problem with projection because the model is failing to capture
some part of the structure of a key process. These critics also point out that benchmarking targets are themselves estimates, subject to errors (though typically smaller ones than in the micro case).

Understanding how precisely different dynamic microsimulation models calibrate their outcomes and the magnitude of the adjustments that are made in simulations may shed some light on this debate. It may also shed light into the possibility and validity of linking micro- and macro-level models, because historically many modelers have instituted linkages from the “top-down” (see, for example, the discussion in Anderson, 1990). In the most recent version of the DYNASIM2 model, alignment factors, which include both linear and non-linear adjustments, are applied in any given year of the simulation based on the success of micro-equations in predicting a target (from historical data or an expectation about the future). If micro-dynamic equations predict too few (too many) events (for a categorical outcome) or low (high) means (for a continuous outcome), then probabilities or predicted values of outcomes are scaled upward (downward). These alignment adjustments are performed on a fairly small number of groups, often one to four age-sex or age-sex-race groups (though this is being reworked in the new version of DYNASIM currently under development). CORSIM, with its 1960 start time, relies even more heavily on alignment than other dynamic microsimulation models, including DYNASIM2. The model also performs alignment annually, and utilizes three separate alignment algorithms (one linear, one non-linear, and one based on sorting, depending on the character of the process), and aligns outcomes on various levels (subgroup and sometimes national aggregates) for relatively large numbers of groups (as many as eighty-eight age-race-sex groups, in mortality).

Magnitudes of alignment factors in both CORSIM and DYNASIM differ greatly across processes and across subgroups of the population. Where historical-period scaling factors are close to one (in the linear case), dynamic microsimulation modelers can have some confidence that micro-level equations are working reasonably well. Large alignment factors, in contrast, suggest the possibility of an underlying problem in specification, estimation, or data quality or consistency. In many cases, though, alignment factors are simply a way of imposing assumptions, like wage growth, rather than necessarily a corrective mechanism. When this is the case, the large size of an alignment may simply reflect the effect of compounding, rather than a poor fit in the micro-level process.

Assuming a model where all adjustments are small (one in the linear case) or non-existent is useful for illustrating one of the ways that a macro-micro link might be imposed in a microsimulation model with a dynamically aged population. Given an estimate from a macro-economic model of, for example, the change in the proportion of the population participating in

---

26Section IIIA.4 in Cohen et al. (1999) describes advantages and disadvantages of each procedure, and proposes alternatives.
the labor force due to a tax change, the microsimulation modeler can constrain the adjustment factor to replicate the new total. If the macro model says labor supply will be reduced by two percent, then scaling factors can be uniformly reduced from 1.0 to 0.98. If the macro model or prior research suggests that the change in participation will occur primarily among women, then the women's factor can be reduced by more and the men's factor reduced by less based on empirical elasticity estimates. Similarly, if economic stress caused by low investment leads to higher levels of unemployment in the society, users can adjust outcomes for groups traditionally hit hardest by unemployment (for example, young black males) more than outcomes for groups historically hit less hard. In either of these cases, the user can impose interaction with the macro model in a strictly top-down fashion, or both top-down and bottom-up through, for example, use of the overall employment level as a determinant in the micro-level equations. Presumably, the latter approach, if well specified and estimated, would require lower scaling factors in the top-down segment of calibration.

There are situations in which one microsimulation model type may be preferred over the others. Besides having differing structures and assumptions, the models have different populations (e.g., some include the institutionalized, though most do not), different program coverage (e.g., quite a few include SSI and DI, while others do not), and different levels of detail about economic well-being (e.g., some incorporate information about pension wealth and income, including detailed information about whether pension benefits are indexed and/or integrated to Social Security benefits, while others do not). When evaluating certain types of reforms we might have more confidence in a prediction from a model that is based primarily on historical trajectories, and for other types of analyses we might expect more favorable results from a model with more explicit structures for key outcome generating processes.

Where one believes that social processes (e.g., marital status, child bearing, and labor force participation) interact a lot and that there is significant, on-going change in population composition, one is likely to prefer a fully dynamic microsimulation model. Where one believes that interactions among processes are minimal, or that the links between these processes go for the most part in a single direction (e.g., from marital status to labor force participation, but not the reverse), one may prefer less elaborate aging mechanisms. There are conceptual and practical tradeoffs between all of the microsimulation strategies that we have outlined.

**Summary of Population Microsimulation Models.**

*Representativeness and completeness of starting data:* Variable, but generally very high; most start with exact match files, which have excellent detail on covered earnings; tradeoffs exist between having earnings records and having a fully representative population (including, for example, the institutionalized and those outside of Social Security covered work); in the case of the dynamic style models, the further into the future one simulates, the less important the starting point.
Cohort size (aside from stochastic replications): Usually at least hundreds, and often thousands, of persons and dozens or hundreds of family types.

Generalizability: Quite high, may encompass very detailed subgroups.

Ability to oversample parts of the population: In cohort simulation, fairly straightforward; in dynamic models with open marriage markets, need to make special provisions to exclude oversampled individuals from the marriage market; in any microsimulation type, need either to adjust tabulation weights or exclude oversampled populations from tabulations providing population-level estimates.

Theoretical and empirical validity of updating algorithms: Variable; at a minimum, assumes fixed structure from historical; may account for compositional changes in the population, but not structural changes in a process; generally processes have fixed structures with in practice few macro-level determinants; sometimes, though rarely, forward-looking aspects are integrated.

Autocorrelation/serial correlation: Variable; in regression-based updating, often incorporated through complex error structures (fixed or random effects models) within processes, may be less developed across processes (few jointly determined); other times, incorporated by using observed (or somewhat modified) historical trajectories, rather than explicit structures; alignment is a further means for correlating across individuals within periods (“spatially”).

Feedbacks: Variable; for models employing reweighting, difficult as currently structured, though issue explored more fully in Chapter 2; aligned dynamic aging models have greater potential than other model types for demographic feedbacks; economic feedbacks inadequately developed in most, but in principle possible/straightforward, especially if outcomes are aligned.

Outcomes: Extremely broad capability: full distributions of cross-section outcomes over simulation period; poverty analyses; with proper calibration, can provide detailed, or at least relative, cost estimates.

Programmatic detail: Typically very high, with nearly all aspects of OASDI; often includes detailed federal income tax and transfer estimates (e.g., SSI) as well.

Projection assumptions: Heavy burden on the user to be explicit about assumptions about future levels and differentials in processes; major aggregates (e.g., total fertility rate, life expectancy, or age-sex adjusted death rate) typically aligned to Trustees’ Report assumptions, although some models allow the user to “de-activate” alignment procedures (and thus to rely solely on predictions from the micro-level equations).

---

27 Burtless (1996: 259) describes this as an advantage of dynamic microsimulation in particular, since other strategies (e.g., cell-based modeling) make the same sorts of assumptions, only implicitly.
Sensitivity analyses: Extensive capability for multiple scenarios where alignment targets are integrated; can test sensitivity of both levels and differentials in social and economic processes; more difficult where outcomes are not calibrated or where reweighting is employed.

Stochastic replications per individual/cell: Ranges from one (most models) to two (PRISM), although some models use sampling procedures to estimate Monte Carlo variability of cells within a single simulation (see Gribble, 1997); several overviews of specific models report analyses of estimates' sensitivity to Monte Carlo variability; the need for replications declines with starting sample size; many models that do not explicitly perform replications employ variance reduction procedures.

Stochastic replications of aggregate demographics/economics: In practice, single or very small numbers of replications, though in theory could integrate stochastic elements; integration is more feasible where run time is faster (which tends to be the case with less complex models, though some complex models are quite efficient computationally); most models not likely to approach regular capacities for the number of replications easily integrated in aggregate models (e.g., models constructed by Anderson, Lee, and Tuljapurkar, which use around one thousand replications) soon.

Ease of use: Greatly improved over the past few years, but still a limitation of the strategy; nearly all major models examined now run on large populations for many decades in close to or less than an hour, permitting multiple runs per day; complexity of code quite variable; Social Security parameter changes straightforward to implement, changing code for parameters in underlying processes is often less so.

Accessibility: Variable; use of models that rely on administrative data for starting samples tends to be restricted, though their parameters are often in the public domain; several large models are held by for-profit organizations that restrict access to clients.

Development and maintenance time and cost: Substantial; among the more costly of the strategies profiled, but costs falling due to wider availability of longitudinal data, existence of sophisticated templates (CORSIM, DYNACAN, DYNASIM, PRISM), and rapidly declining costs of computer memory and processing time; costs for models relying on reweighting tend to be somewhat less than for dynamically aged microsimulation models.

5. Type 4: Computable General Equilibrium Models

Computable general equilibrium models, including overlapping generations models, have been receiving increasing attention in recent years. The Panel on Retirement Income Modeling (1995) describes them as one of the bright spots in a disappointing field of models that often lack theoretical content. While these models have firm theoretical grounding, relying on assumptions about utility maximization, and are often excellent at ensuring consistency among macro variables, such as employment rates, wage levels, and total output, few are sufficiently disaggregated to
capture both the richness and detail of policy reform options (e.g., policies that change the number of computation years or influence portfolio allocation) and also population diversity within cohorts.

Several implications of a GE structure frame its ability to inform debates about reform proposals. In these types of models, Social Security is represented as a component of the larger government sector. Developers generally set individuals' benefits at the mean for everyone in the cohort or cohort-sex group (though some recent models have more detailed representations of Social Security, including replacement rate based on lifetime income), and evaluate Social Security outcomes when the economic system is in long-term equilibrium (e.g., government spending and government revenues are equal, all labor and capital markets have cleared). While this approach is useful for considering long-run implications of the current system and various reforms, it is potentially difficult to use this type of model to reveal the distributional implications of the cross-sectional financing steps that lead to this eventual equilibrium.

Use of a general equilibrium model is possibly the only way that researchers can consider the long-run implications of integration of private accounts into Social Security for saving, capital accumulation, and wage growth. Such a model can formulate estimates of a group’s lifetime tax incidence assuming long-run government discipline, and is thus an excellent means for studying intergenerational equity issues with respect to Social Security and other taxes and transfers.

As critical as these interconnections may be for the study of more fundamental, structural sorts of reforms, at present many GE models are more useful for producing qualitative instead of quantitative predictions. Where a general equilibrium model provides cross-sectional estimates, these are not derived from shifts in the Social Security benefit formula as now structured. Because they do not produce distributions of outcomes, these models have little capacity to inform discussions about poverty and the status of at-risk populations. Likewise, while existing GE models can greatly inform discussions about the consequences of labor supply and saving responses to the changing incentives that arise when a government alters the relative sizes of social assistance and social insurance, they lack an ability to inform discussions of program interactions in very specific detail.

An example of this type of model is Auerbach and Kotlikoff’s (1987) model (the AK model) which uses fifty-five overlapping generations (adults, all married, ages twenty-one to seventy-five). This model, which has been expanded to include stylized lifetime income groups (Kotlikoff, Smetters, and Walliser, 1998, 1999), assumes forward-looking agents with access to complete information about future time paths of key variables who follow the life-cycle model in allocating consumption and leisure across time periods. The Fullerton and Rogers' tax incidence model (1993), which has a population that is disaggregated by income as well as by sex and cohort, also is based on the life-cycle model, but assumes myopic expectations of agents. The model employs a fuller, more detailed range of consumption choices than the AK model. The Aaron, Bosworth, and Burtless' Social Security model (1989; for a similar treatment, see
Bosworth and Burtless, 1997), with a population based on OASDI Trustees' projections, is a simpler example of a GE model. (See discussion of GE models in Chapter 4.)

A large class of models, numerical stochastic general equilibrium (SGE) models, are descendants of the deterministic overlapping generations models pioneered by Auerbach and Kotlikoff (1987), generalized to incorporate various forms of uncertainty. Within the last five years increasingly comprehensive versions of SGE models have been formulated that incorporate multiple decisions and types of uncertainty and fairly detailed representations of Social Security policy (see, e.g., Storesletten, Telmer and Yaron, 1999). The models that we describe in this section are among the most ambitious policy models that have been solved and simulated in academic research. Some of these models are sufficiently realistic to be taken seriously in policy analysis. They demonstrate that it is now feasible to provide detailed quantitative evaluations of both the short and long run general equilibrium feedbacks and welfare effects of a range of policy issues of current interest such as increasing the level of Social Security Trust Funds or privatizing Social Security via creation of individual accounts. These models provide both long–run steady–state forecasts as well as detailed calculations of the transition dynamics that move the economy from an existing steady state to a new steady state following a change in policy or in demographic variables such as birth rates or mortality rates (Huang, Imrohoroglu, and Sargent, 1997, De Nardi, Imrohoroglu, and Sargent, 1998).

The main drawback of SGE models is the computational burden involved in solving them. The computational problems arise from the need to re–solve individual agent DP problems repeatedly as an "inner loop" within a zero–finding "outer loop" algorithm that searches for values of prices, wages, and interest rates that clear the market, and values of tax rates and benefit levels that satisfy government budget constraints. Although it is now feasible to solve an SGE a relatively small number of times in order to evaluate various policy scenarios, it is infeasible to solve most of these models hundreds or thousands of times. This is what would be required to estimate econometrically the unknown parameters of SGE models, where a "third loop" would be added to solve the SGE model for various trial values of its unknown parameters until values were found that enable the predictions of the SGE model to "best fit" observed outcomes according to a well–defined metric. Instead, SGE models are calibrated via an unspecified informal procedure for selecting the unknown parameters so that the model's predictions "best fit" observed data according to an implicit but typically undefined metric. The current state of the art in the academic literature on calibration is an ad hoc combination of the somewhat in–bred process of "borrowing" most of the model parameters (e.g. coefficients of risk aversion, or the parameters of a Cobb–Douglas production function) from previous calibration studies, and then selecting a small number of remaining parameters in an attempt to match the model's predictions to reality. The criterion used to specify "best fit" is often unclear in calibrated models, and we currently lack a rigorous statistical theory that could be used to evaluate their goodness of fit and the degree of uncertainty in their predictions. This lacunae is due mostly to computational problems but also due to lack of an appropriate statistical theory. For example there are difficult unresolved issues about how to include unobservables in these models. Unobservables are needed
to avoid problems of “statistical degeneracy,” i.e., how to deal with observed outcomes that have probability zero of occurring in the SGE model. Even setting aside computational problems, there is no theory of econometric inference in general equilibrium models, and statistical methods for conducting estimation and inference are still largely undeveloped. For a more detailed description of the process of calibration and its problems see Hansen and Heckman (1998) and Srinivasan and Whalley (1998).

As a result of these inferential problems, it is not yet possible to test rigorously the validity of SGE models (i.e., the implicit rationality or equilibrium assumptions) and assess their predictive accuracy. Similar to DP models, the main way to assess the credibility of SGE models for use in policy forecasting is via informal “out-of-sample” predictive tests. Further, the “forecasts” of many SGE models are best viewed for their qualitative features rather than the realism of their quantitative predictions. Even among models that are intended to generate realistic quantitative forecasts, there are many difficult issues pertaining to the appropriate metric for ranking the accuracy of alternative SGE models that we are unable to resolve here. To help diversify “forecast risks” SSA might want to rely on a “portfolio” of forecasts from different models rather than attempting to identify a single “best” model. In spite of these difficulties, we believe that SGE models can be a very important tool for SSA, even if some models are best employed for qualitative predictions, and even if, among quantitatively oriented SGE models, it is not yet possible to quantify the uncertainty we have about their predictions scientifically and objectively.

The remaining limitations of SGE models from the standpoint of their use in policy forecasting at SSA concern their failure to incorporate various risks, aspects of reality, and features of the Social Security program. These limitations are largely related to the “curse of dimensionality” associated with solving the DP problem for more realistic formulations of individuals’ decision problem, as discussed in section 5.2 of Chapter 5. We simply list several of the most important limitations which are relevant to addressing policy issues of interest to SSA. One key limitation is the failure to account for stochastic macro shocks such as the effect of an oil shock or a stock market crash. The production technologies and rates of return on assets in current generation SGE models are deterministic. None of these models allows for the possibility of stock market “bubbles,” which is a focus of concerns about privatization and investing Trust Fund assets in equities. The failure to incorporate stochastic macro shocks limits the usefulness of these models for studying the role of Social Security as a device for intergenerational sharing of aggregate risks. This type of risk-sharing is tangentially addressed in some of the studies reviewed in Chapter 5 (e.g., the De Nardi et al. 1998 paper shows how different cohorts share the burden of a perfectly anticipated demographic shock) but at present there is no SGE model that can address the question of whether Social Security can provide a welfare enhancing transfer mechanism to help out a cohort that suffered the consequences of a sustained stock market crash or depression, which Blinder (1988) ascribed as one of the reasons motivating the introduction of Social Security in 1935.
A useful direction for work in the near term is to determine policy issues where it might be reasonable to abstract from general equilibrium feedbacks and thereby avoid the complexities inherent in solving SGE models. Our survey suggests that general equilibrium feedbacks cannot be safely ignored for fairly radical changes to the Social Security system such as privatization or changing from a pay-as-you-go system to a fully funded system. However it is likely that there are questions for which equilibrium feedbacks are second order. One example might be the issue of investing the Trust Fund in equities. If we believe that most individuals can undertake offsetting changes in their private portfolios and that the U.S. is increasingly a small player in a global equity market, then it seems that the impact on the equity premium of investing the Trust Fund in equities would be second order and might be safely ignored in policy analyses similar to what was done in the Feldstein, Rangelova and Samwick (1999) study that we will discuss in section IV.3 of Chapter 5. It is hard to determine in advance of actually solving an SGE model whether general equilibrium feedbacks are likely to be important, but common sense and intuition are likely to be good guides. This is standard practice in academic models where certain variables are assumed to follow exogenous stochastic processes even though it is possible to argue that on a grand scale “everything is endogenous.” We think it is eminently reasonable to try to reduce the computational burdens underlying SGE models by identifying certain processes such as equity returns, trends in fertility, mortality, marriage and divorce that can be reasonably treated as exogenous stochastic processes, at least to a first approximation.

Chapters 4 and 5 will expand on several of the issues associated with general equilibrium models that we have raised here, and introduce new issues as well.

**Summary of Computable General Equilibrium Models.**

*Representativeness and completeness of starting data:* Often like representative worker model; families highly stylized.

*Cohort size (aside from stochastic replications):* Variable; often just one or two workers per cohort, some have a few dozen workers, differentiated by lifetime income group, and some have large numbers of agents; usually just one or two family types.

*Generalizability:* To cohort-sex groups or cohort-sex-earnings groups, but usually very limited within generations.

*Ability to oversample parts of the population:* None as most models are currently constructed, but with weighting could employ oversampling for important populations.

*Theoretical and empirical validity of updating algorithms:* Consistent with rational expectations framework (either with complete foresight or myopic); parameters selected to be consistent with stylized facts, but often not empirically estimated.
Autocorrelation/serial correlation: Usually little heterogeneity, though some models introduce heterogeneity among agents using complex error structures.

Feedbacks: Can be extensive: in models where labor supply and saving are endogenous, takes into account both supply and demand sides of individual labor supply and consumption choices; in models where labor supply and saving are exogenous, does not incorporate incentive effects of policy changes.

Outcomes: Aggregate welfare effects; lifetime, but not usually cross-sectional, mean values for as many as several dozen hypothetical profiles; no distributions of targeted groups (e.g., no poverty incidence).

Programmatic detail: Typically low (mean or replacement rate approaches to benefit computation), related social programs not always disaggregated, but this is not inherent to the approach.

Projection assumptions: System is in long-term equilibrium, government budgets balance, exogenous targets.

Sensitivity analyses: Can test the sensitivity of the model to different key assumptions such as intertemporal substitution and labor/leisure substitutability.

Stochastic replications per individual/cell: See Chapters 4 and 5 for discussion of individual uncertainty in this modeling context.

Stochastic replications of aggregate demographics/economics: See Chapters 4 and 5 for discussion of aggregate uncertainty in this modeling context.

Ease of use: Variable, but generally requires extensive knowledge, training.

Accessibility: Variable.

Development and maintenance time and cost: Variable; templates exist, but detailed model could be as substantial an investment as a dynamic population microsimulation model, or even greater.

V. WHICH MODEL WHEN?

Our discussion of competing models has preceded at a fairly abstract level. Let us now briefly return to the earlier section in this chapter, where we discussed some specifics of proposed OASDI reforms, to become more concrete about situations in which one model type might be preferable over another. It is worth noting that in some instances an analyst may not need a full scale model of any of these types to answer important distributional questions about Social Security. Work by Johnson (1999) on COLA reductions, for example, suggests that simple
models based on recent empirical data can inform intuition about the relative effects of certain benefit-side reforms for people in fairly large groups. However, the number of situations in which this is all that one needs or seeks to know about a Social Security reform proposal are relatively limited.

Our earlier distinction between programmatic and structural reform is a useful starting point. For programmatic reforms in which feedbacks are likely to be relatively small, analysts can combine analyses from representative worker, cell-based models, and population microsimulation models to get a nearly complete picture of expected distributional effects of policy changes. The need for stochastic general equilibrium models is surely less acute in either of these types of situations than in circumstances in which incentive effects and behavioral responses are likely to be large.

Depending on the nature of a programmatic reform, a different model within each of these classes may be more useful. One is likely to choose a different model to evaluate proposals that change the treatment of earnings above the taxable maximum, that change levels of divorced spouse benefits before and after the death of a spouse, that change the number of computation years for determining Social Security benefits, or that change the benefit formula's bend points or bend percentages. In the first case, detail on distributions of earnings above the taxable maximum is essential, and several models that rely on administrative records for a starting data base do not currently account for these earnings. In the second case, forecasts of divorce patterns, based on up-to-date information, are necessary. At present, only dynamic microsimulation population models currently contain this sort of detail on marital status changes. In the third case, a model needs to have reliable patterns of year-to-year earnings variation. Some models have earnings functions that are geared toward replicating average lifetime earnings rather than toward predicting annual levels/fluctuations. The fourth case, in contrast, may be less demanding for earnings patterns. Reliable AIME predictions should be adequate.

For more structural reforms, or for programmatic reforms in which researchers anticipate large behavioral responses, analysts again are likely to need to use multiple models to inform their discussions. In the case of a structural reform like earnings sharing, population microsimulation models with explicit marital status projections are likely to be dominant, because researchers need full earnings and marital histories, information that is unavailable in a cell-based model and inadequate in general equilibrium or representative worker models. For individual account proposals, analysts again need population microsimulation models, especially when accounts are self-directed and individuals respond to them in heterogeneous ways. As noted earlier, analysts must use representative worker models quite carefully in this context, as highly stylized trajectories are likely to reveal significantly different relative advantages for defined benefit and defined contribution types of Social Security benefits than are more realistic examples (see, for example, Burtless, Bosworth, and Steuerle, 1999). When reform proposals are more fundamental/structural, consideration of feedbacks is more necessary. The more essential an
understanding of a proposal’s general equilibrium implications, the more necessary the reliance on general equilibrium models.

Linking microsimulation models to more aggregated models (like a computable general equilibrium model) may be a way to answer these most difficult questions in a unified way. Chapters 2 and 4 will discuss some of the options for micro-macro linkages in more detail.

VI. CONCLUSIONS

Our examination of the literature has revealed a number of positive developments in many arenas of individual and family-level modeling, although we have yet to see a model that fully integrates all of our concerns about relationships necessary for simulating the reform proposals advanced in recent months. Representative worker models may have much of the program detail that we want from a model, but they lack the necessary population detail. Cell-based models, while often effective for tracking certain aggregates like Trust Fund balances, do not typically generate individual earnings histories, rendering them inadequate for the study of distributional effects of many Social Security reform proposals. While general equilibrium models have the links to the larger economy (including production functions and demand for labor and capital) that are critical and also use a consistent framework that incorporates behavioral responses, they currently lack adequate distributional and program information. Reweighted microsimulation models can effectively describe short-run effects of changes to individual parameters, but they may not produce estimates that add up to national totals and are less persuasive in the face of major structural change or major cohort-to-cohort change in life trajectories. Dynamic microsimulation models can capture the interactions of multiple program parameters and social processes and are likely to be more effective when cohort-to-cohort changes in life trajectories are substantial, but, like their static counterparts, tend not to integrate behavioral relationships and macroeconomic constraints adequately.

The fact that no one model can do everything, however, does not imply that existing models cannot tell us a good deal at this juncture. Some models can greatly inform discussions of certain reforms while having limited capability to address others. Fundamental structural reforms place the greatest demands on models, and it is in evaluating prospects for distributional analyses of these proposals that we see the greatest need for both further model development and additional, basic research.

As Chapters 4 and 5 will discuss, one prospective approach to overcoming these existing limitations would be to develop dynamic, stochastic general equilibrium models which, like population microsimulation models, operate on large-scale data bases. Unfortunately, academic research and computational resources are not yet at the stage where this is a feasible short-term strategy. Linked micro and macro-level models with simpler behavioral relationships and less extensive feedbacks may be more feasible.
REFERENCES TO CHAPTER 1


CHAPTER 2

REVIEW AND ASSESSMENT OF THE COHORTS SIMULATION APPROACH

I. INTRODUCTION

To answer questions about the distributional effects of Social Security reform proposals, both in the long-term and over a transition period as those changes phase in, it is necessary to have some representation of the economic and demographic characteristics of the future population. The cohorts model under development in the Division of Economic Research (DER) is one method for making such projections.

This chapter provides a review and assessment of the cohorts model. The assessment of the cohorts model in this chapter is based on documentation provided by DER and focuses on the model as described in that documentation (see Pattison, 1998a, 1998b, 1998c, and 1999). Because the cohorts model is still under development, this assessment is meant to provide some guidance for future directions, as well as to suggest ways in which the model could be improved. The review begins with an overview of the model, listing the types of analyses that the model can currently perform, as well as those beyond the current scope of the model. It then provides a general assessment of the model as a tool for distributional analysis of Social Security reform options without incorporating individual level behavioral response or feedbacks from changes in the macro economy. This assessment focuses on issues and solutions for problems with missing information from data that are the basis of the model, and it provides an appraisal of DER’s actual and proposed methods for projecting future earnings histories. The next section considers whether the cohorts model provides any special advantages or weaknesses as a framework for incorporating individual behavioral responses and macroeconomic feedbacks relative to other microsimulation models. The final section summarizes the report.

In the current stage of development, the cohorts model projects earnings and Social Security benefits. In order to assess the overall well-being of future generations of retirees or to answer questions such as how many retirees are likely to have incomes near or below the poverty threshold, a simulation model must, in some form, also project future pension benefits, income from non-pension assets, and income from other sources including additional government transfers. Because the cohorts model at this stage in its development does not project these other sources of income, it cannot be used to answer these broader distributional questions. The issue of how other sources of income can be included in microsimulation models is discussed in Chapter 3.
A complete analysis of current Social Security reform proposals that include some type of individual saving component should account for the potential effects on future incomes of changes in aggregate savings and economic output. Because the cohorts model at this time does not capture the interaction between Social Security reform proposals and the economy, it cannot address these issues. A discussion of macroeconomic models and possible linkages to microsimulation models is provided in Chapter 4.

II. AN OVERVIEW OF THE COHORTS MODEL

The cohorts model is a microsimulation model designed to answer questions about the long-term distributional consequences of various Social Security reform proposals. The model uses a sample drawn from the 1994 Current Population Survey (CPS) of persons who became entitled to old-age benefits in 1992 and thus were born in 1930 or earlier. In addition to the demographic and economic information collected in the CPS, the data include matched earnings records and benefit entitlement data from SSA administrative records. The model uses a detailed algorithm to calculate expected lifetime Social Security benefits at the age of entitlement under current law and proposed alternatives.

The Historical Cohort Model (HCM) simulates the effects of proposed reforms on the 1930 cohort as if the reforms had been in place throughout the entire working lifetime of that cohort. The Projected Cohorts Model (PCM) takes the 1930 cohort as the starting point for simulating the effects on later cohorts. The model projects later cohorts by adjusting the earnings histories of the 1930 cohort to reflect the growth in average wages between the actual and projected years. The model makes further adjustments to labor force participation rates and the earnings of women. Using these projected cohorts, the model can simulate the effects of Social Security reform prospectively, showing how those reforms affect future cohorts during the transition phase and how they affect cohorts for whom the reforms have been in place throughout their working lifetimes.

1. Applicability of The Cohorts Model

Because the model uses actual data on earnings histories and benefit timing for a representative sample of the population, it is a useful tool for analyzing the near-term effects of Social Security reform proposals. The usefulness of the model for long-term projections will depend on how well the economic and demographic histories of the 1930 cohort represent the histories of future cohorts, *along dimensions that will matter to the analysis of reform proposals*, and, to the extent that they do not, whether implemented and planned adjustments to those histories will correct those shortcomings. The usefulness of the model will also depend on whether the model can incorporate behavioral responses to new Social Security policies where appropriate, although certain types of analyses, such as of winners and losers under different
proposals, are possible even in the absence of modeling behavioral changes. Behavioral responses are not a feature of the current model.

In its current form, the model can answer questions about the distributional effects of programmatic options for Social Security reform such as changes to the Social Security benefit formula, indexing of benefits, and the computation of average earnings (for example, reducing the number of dropout years). The model can compare the relationship between expected lifetime retirement and survivor benefits and lifetime OASI taxes for members of the 1930 cohort who reached the age of entitlement, although it will give an incomplete picture of aggregate lifetime taxes and benefits for that cohort unless imputations or adjustments are made for taxes paid by workers who die prior to reaching the age of benefit entitlement.

The model can also analyze certain structural reform options such as individual account proposals. The model simulates contributions to accounts from workers' earnings according to the specification of different plans (for example, contributions that are a fixed percentage of earnings), an allocation of investment in stocks, corporate bonds, and government bonds, and an expected value and variance of returns at retirement. Because the model does not currently simulate private saving or pension accumulation, however, it is not possible to model the effect substitution of private accounts for Social Security may have on other retirement saving.

2. Types of Analyses Beyond The Current Scope Of The Model

In its current stage of development, the model does not have the capacity to explore distributional consequences of Social Security reform outside the immediate confines of the Social Security program itself. That is not to say that DER could not expand the model to answer some of these questions. But in assessing the model, it is important to have in mind the types of analysis that are beyond the model in its current form. Briefly, these include the following:

Economic Well-Being of the Retired Population.

The PCM does not attempt to project income other than earnings for future cohorts. Thus, it is not possible to use the model in its present form to assess how changes in the Social Security program would affect the overall economic well-being of future retirement cohorts. It is possible to look at the distribution of total income for the 1930 cohort because information on other sources of income such as interest, dividends, proprietorship returns, and pension benefits are available for 1993 from the March 1994 CPS. DER has done some analyses with the model using this broader measure of income. While this provides a reasonable picture of the income of the 1930 retirement cohort at first entitlement to Social Security, it is not complete. The CPS has no information about assets. Returns to certain capital assets that are reported on the survey are significantly less than the amounts reported from other sources such as administrative tax return
Moreover, income from pensions and other assets for new retirees will overstate income later in retirement as these income flows tend to decrease in real terms over time.

The CPS information by itself is not adequate for projecting pension and asset income for future cohorts of retirees. While a simple extrapolation procedure, such as scaling reported pension and asset income by the same growth factors used to project future earnings, would provide an estimate of income other than earnings for future cohorts, it is not likely that the relative income from those sources received by today’s retirees would give an accurate picture of income for future retirees. There have been substantial changes in the institutional arrangements for retirement saving available to current workers including a shift in employer pensions away from defined benefit plans to defined contribution plans and a host of new tax incentives for individual saving. The Social Security program itself has greatly changed with increases in both the payroll tax base and the tax rate, and, despite recent cut-backs, increases in expected lifetime benefits. Moreover, the asset accumulation of the 1930 cohort took place under a particular set of economic conditions, such as the run-up in housing prices during the 1970s and equity prices during the 1980s and 1990s, that are not common to all cohorts. While some of these changes may be partially offsetting, it is unlikely that future retirees will have the same mix of Social Security, private pensions, and other saving as the current generation.

**Program Interaction.**

The PCM does not simulate Social Security disability benefits, SSI, or Medicare, although DER is exploring imputations of SSI incomes based on CPS information. Without modeling these additional benefits, the model cannot address how changes in the Social Security retirement program will affect payments under other federal programs, or how changes in other federal programs will affect Social Security.

**Post Retirement Earnings.**

Earnings data for the 1930 cohort are available only through age 63. With no post-entitlement earnings in the model, it is not possible to simulate the effects of Social Security reforms on work after retirement. The model can easily be extended to include full earnings histories for cohorts born prior to 1930, which would provide some information about post-retirement earnings. Because earnings histories extend only back to 1951, however, earlier cohorts lack information on early career earnings. Summary Earnings Records (SER) are currently available for individuals up to 1996. Matching the CPS with the most recent SER would add post retirement earnings for the 1930 cohort up to age 66.
Social Security Trust Funds Balances.

Because the model focuses on a single or a repeated series of cohorts as they reach retirement, the model does not simulate the entire population. Thus, at a point in time, say a projection to the year 2030, the model does not include younger workers who have not yet reached the age of retirement and older workers well beyond retirement age. As a result, the model cannot calculate total Social Security receipts and benefit payments, and cannot be used to estimate the effects of various Social Security reforms on trust fund balances. Such analysis would require that the model include all cohorts. There is nothing within the methodology of the model that prevents the inclusion of the full population. If projecting trust fund balances becomes a goal of the model, the omitted population can be added.

Macroeconomic Effects of Social Security Reforms.

Because there are no business, financial, international, or government sectors in the model, it is not possible to calculate the effect of Social Security reform on wages and interest rates, and thus on macroeconomic variables such as national saving, total employment, or domestic output. Future work with the PCM will focus on generating individual savings which will allow for some investigation of macroeconomic issues.

The inability of the PCM to perform some types of analyses is shared with other microsimulation models. While some of these analyses can be handled by other models such as DYNASIM, CORSIM, and MINT, others cannot. For example, none of the existing microsimulation models capture the macroeconomic effects of Social Security reforms. Some models, however, have a more developed representation of other sources of income, program interactions, and post-retirement earnings than the current version of the PCM.

III. KEY ISSUES FOR THE PROJECTED COHORTS MODEL: MISSING DATA AND EARNINGS PROJECTIONS

The first aim of the cohorts model is to move beyond representative worker models and base distributional analysis of Social Security reform options on more realistic and much more diverse lifetime earnings histories. A general assessment of the cohorts model suggests that, on those terms, it is indeed a useful tool, but there are areas where the model can be improved. This section discusses some problem areas for the model, many of which have been identified by the Division of Economic Research, and what could or should be done about them. It focuses on two specific areas: first, concerns about missing information in the CPS survey and SSA administrative data used in the model, and second, whether methods used to project future earnings histories capture important trends in the time-path and the intra- and inter-cohort distribution of earnings.
A second aim of the cohorts model is to provide a framework for more ambitious modeling of retirement issues including microeconomic feedback between policy variables and lifetime economic behavior and possible macroeconomic feedbacks. Those issues are addressed in the next section.

1. Starting Database

The model uses a sample of persons who first became entitled to Social Security old-age benefits in 1992, drawn from the March 1994 CPS, matched with earnings histories and entitlement information from SSA administrative records. There are some limitations to both the survey data from the CPS and the administrative records that may need to be addressed.

Incomplete and Underreported Survey Data From the 1994 CPS.

Although data from the CPS survey give a reasonably complete picture of the total income of the 1930 cohort at retirement, certain economic information is either underreported or missing from the CPS. In particular, income from assets is underreported relative to data from administrative tax records, and information on other resources, such as returns from the sale of capital assets, is not included in the survey. DER will need to consider adjustments or imputations from other data sources, such as the Internal Revenue Service Statistics of Income (SOI), the Survey of Consumer Finances (SCF), or the Health and Retirement Survey (HRS), if the model is used to look at the total economic resources of the current retired population.

Using data from the Survey of Income and Program Participation (SIPP) instead of CPS data would provide some advantages for measuring total economic resources. SIPP topical modules collect a wealth of information including data on financial assets, housing equity, and pension plan coverage. The 1990, 1991, 1992, and 1993 SIPP panels have been matched with Social Security Administration Master Beneficiary Records and Summary Earnings Records making this data readily available to SSA. Pension information is also available from the 1994 pension supplement to the CPS.

Data from the HRS and the accompanying study of Asset and Health Dynamics Among the Oldest Old (AHEAD) also provide extensive information on the income and assets of current retirees. HRS data have been matched with summary earnings records from the Social Security Administration and pension plan data from plan providers. The combination of information on income, assets, earnings histories, and pension plan characteristics is unique among recent surveys.

Neither the CPS, the SIPP, or the HRS includes the institutionalized population, although the HRS follows sample members who become institutionalized. If it is important to capture the entire Social Security population, for example to compute trust funds payments and receipts,
weight adjustments could be used. However, this must be done with some care, because the institutionalized have very different characteristics than the rest of the population.

**Missing SSA Administrative Data.**

The match rate for SSA administrative data with the CPS was about 82 percent for the 62 to 65 age group. DER has verified that the unmatched records do not differ from matched records along observable dimensions such as age, sex, marital status and income. Although unmatched records may still differ on other characteristics such as lifetime earnings, there appears to be little that can be done in this regard.

**Missing Earnings Information For Deceased Workers.**

The sample does not contain earnings records for workers who die before entitlement to retired worker benefits. If there is a survivor who is entitled to survivor benefits, the earnings records for the deceased person are kept with the survivor's information. For survivors entitled only to worker benefits, the earnings of the deceased spouse are missing. Dually entitled survivors have both their own and their deceased spouse’s earnings records. Because the probability of surviving to the age of entitlement will be higher for workers in later cohorts, some adjustment to correct for this selection problem is warranted. DER should consider imputing earnings records for deceased workers rather than simply reweighting the sample, as those earnings histories will likely differ from histories for the typical survivor. If the post 1993 date of death is known (from SSA numident files) for members of later cohorts on the matched file, those records can be used to provide proxy demographic information and earnings histories for the missing members of the 1930 cohort. Absent those data, information will have to come from longitudinal surveys such as the Panel Survey of Income Dynamics (PSID).

Missing information for cohort members who die before becoming entitled to retirement benefits is a particular problem for modeling disability recipients. Although the sample includes earnings histories for disability recipients from the 1930 cohort who survive until age 62, it does not have information on non-survivors. As the retirement age for full Social Security benefits increases, it is likely that a larger proportion of older workers will collect benefits under the Social Security disability program (Congressional Budget Office, 1999). Also, as female labor force participation increases, the risk of eligibility for disability benefits for these new workers also increases. Furthermore, certain Social Security reform options would increase the minimum age of eligibility for retirement benefits, leading to a likely increase in disability beneficiaries. A model that omits disabled workers would miss an increasingly important group of Social Security beneficiaries. Information for disability recipients in younger cohorts could substitute for the missing members of the 1930 cohort, with proper adjustments.
Missing Earnings Information For Former Spouses.

Earnings records of former spouses of divorced workers are not included in the sample. Although most divorced retirees receive benefits based on their own earnings rather than on the earnings record of their former spouses, there are beneficiaries for whom their divorced spouse’s benefits exceed their own worker benefits. This is also a problem for modeling certain types of reform proposals, such as earnings sharing, in which benefits would be based on joint earnings histories from current and previous marriages. Earnings histories for intact couples can be used to proxy the information for former spouses. Alternatively, DER could use a statistical matching algorithm to match divorced individuals to former mates. The latter method might better control for systematic differences in earnings histories of individuals who divorce compared with individuals who remain married.

Earnings Limited to Social Security Covered Earnings.

Earnings histories for the 1930 cohort show no earnings for years in which workers were in non-covered employment. The 1930 cohort began working in 1951 when Social Security coverage was much less universal than it is today. Although about 96 percent of civilian workers are covered by Social Security today (98 percent including those covered only under the Hospital Insurance program) the coverage rate was lower in earlier decades, growing from about 60 percent in 1950 to 86 percent in 1960 and 90 percent by 1970. Thus what might appear as years out of the labor force for members of the 1930 cohort may instead have been years in uncovered employment. The PCM currently imputes earnings to women to adjust for increases in labor force participation. DER should consider using similar methods for imputing non-zero earnings for men who were in non-covered employment early in their careers.

Censored Earnings.

Historical earnings on the SER are censored at the maximum taxable amount. Correcting the censored earnings is particularly important for men’s pre-1980 earnings, when the taxable maximum was lower and about 60 percent of men had censored earnings. DER uses a sophisticated econometric procedure to impute uncensored earnings.

The method used to estimate uncensored earnings first estimates age-earnings profiles by sex, controlling for period effects, using pooled earnings records from the SER for the entire sample, including earnings censored at the taxable maximum. Those estimates are then used to calculate an estimate of the correlation of individual earnings across different ages. An estimate of uncensored earnings for years in which an individual’s earnings were at the maximum is imputed based on the person’s uncensored earnings in other years and the estimated correlation structure of earnings.
The model uses Gibbs sampling to impute uncensored earnings (see Casella and George, 1992, Keane, 1993, and Gelfand and Smith, 1990). The Gibbs sampler is a technique for generating random variables from a distribution using a Markovian updating scheme. It iteratively replaces censored error terms with error terms generated from the estimated joint distribution of error terms across ages. Hajivassiliou and McFadden (1997) proved that the Gibbs sampler converges at a geometric rate to the true distribution.

It may be possible to improve the imputation method. First, as DER has noted, the approach does not use available information on the quarter in which a person’s earnings reached the taxable maximum. That information could be used to bound the imputed uncensored earnings. For example, if a person’s earnings reached the maximum in the third quarter of the year, then, assuming that earnings were the same throughout the year, the person’s uncensored earnings would be not more than twice nor less than 4/3 of the taxable maximum.

Second, it may be possible to improve the initial estimates of age-earnings profiles by controlling for education--running separate regressions for different education groups--and by including individual-specific fixed-effects (See, Honore and Kyriazidou, 1998). Controlling for differences in education is important because the shape of age-earnings profiles differs by years of education. A fixed-effects model would control for individual specific year-to-year persistence of above or below expected earnings for high-earners. Controlling for both these effects should yield more precise estimates of age-earnings profiles.

The replacement technique for censored earning does allow for individual-specific effects in estimating uncensored earnings. However, it relies on the year-to-year correlation of earnings across all sample members, which may not be the same for low and high earners. Based on our own tabulations of longitudinal earnings from the PSID, the variance in gender, age, and period adjusted earnings is lower for individuals who ever have earnings above the Social Security taxable maximum than for individuals who never have earnings at or above the maximum. This presumably is due to the relatively persistent nature of earnings for those individuals, usually men, who have higher earnings than individuals more marginally attached to the labor force. For example, if low-wage workers are more likely to experience time out of the labor force, the variance in their annual earnings will be higher than the variance in annual earnings for individuals in the labor force full-time full-year. The latter group are more likely to be those with earnings above the maximum. Applying the variance in observed earnings for the uncensored group in the SER sample to the censored group generally overstates the variance in projected earnings for the censored group. Education-specific tobit regressions may also improve the cross-age correlation structure that is the basis for the correction of censored earnings.
2. Calculation of Expected Lifetime Social Security Benefits

The calculation of Social Security retirement benefits follows program rules. The model calculates a retired worker benefit for each person, and potential spouse benefits for both members of a married couple. Retirees receive the higher of their worker or spouse benefit. The model calculates potential survivor benefits for both spouses. Surviving spouses receive the higher of their own retired worker benefit or survivor benefit. Expected lifetime benefits are the discounted stream of benefits weighted by survival probabilities at each age. Expected lifetime benefits for married couples reflect the joint survival probabilities of both spouses.

Differential Mortality.

The mortality probabilities for retirees are based on age and sex and are not adjusted for marital or economic status. Past studies have established strong socioeconomic differentials in mortality (Orcutt, 1989, Pappas, et. al., 1993, and Menchik, 1993), while other studies (Duggan, Gillingham, and Greenlees, 1992; Panis and Lillard, 1996) have found that using income-adjusted mortality rates does matter for calculating lifetime returns from Social Security. As DER notes, most of the income-based differences in mortality are for workers prior to the age of retirement, and that modeling these differences for retirees would overstate the effect of differential mortality on lifetime Social Security benefits. However, there is some evidence of differential mortality at older ages related to education and income (Manto, Stallard, and Corder, 1997; and McCoy, Iams, and Armstrong, 1994). Failure to use education and income related mortality probabilities will understate expected lifetime benefits for higher-income beneficiaries and overstate expected lifetime benefits for lower-income beneficiaries.

Taxation of Benefits.

The Social Security benefits calculated in the PCM are the full benefit amount adjusted for early and delayed retirement, but not for income taxation of benefits. Because the income tax thresholds for taxing benefits are not adjusted for inflation, a larger proportion of beneficiaries will pay income tax on their benefits in the future. In addition, certain reform proposals would increase the taxation of benefits. It is obviously not possible to model income taxes very well without some measure of retirees’ total income. However, a simple procedure would be to use effective income tax rates on benefits calculated for the 1930 retired cohort for future cohorts. The effective rates could be calculated using reduced thresholds that would reflect the decline in the tax thresholds relative to earnings for the particular future cohort of interest.

3. Labor-Force Participation and Earnings Projections

The heart of the PCM is the projection of earnings histories for future cohorts. The model relies on actual earnings histories for the 1930 cohort. The model projects earnings for future
cohorts by multiplying earnings in each year by the growth in average earnings between the actual and projected year. In addition, some relatively simple adjustments are made to earnings of women to reflect changes in earnings patterns that have occurred for more recent cohorts. More specifically,

1) Women’s labor force participation rates in the 1930 cohort are adjusted to match age-specific labor force participation rates in 1993, based on the SER. A probit model is used to estimate a latent participation variable for each woman in every year in the 1930 cohort. The adjustment increases the probability of participation for women out of the labor force at each age until participation rates by age match those in 1993.

2) Relative earnings for women are adjusted to reflect age-specific relative earnings from the SER for 1993.

3) Earnings in every year for all workers (men and women) are adjusted upward by the per capita growth in average wages between the observation year and the projection year. Thus, for example, to simulate the cohort born in 1960, annual earnings for each year from 1951 through 1992 for every member of the 1930 cohort is scaled upward to reflect a combination of actual and expected average wage growth over thirty years. This adjustment maintains the same relative earnings among sample members as in the original data, but changes the year to year pattern of wage growth by the same factor for all workers. Other cohorts are simulated in the same way. Thus, projected cohorts differ only in the scaling of average wages.

4. Changes in Labor Force Participation and Earnings Patterns

There have been several well documented changes in labor force participation and earnings patterns over the last half of the 20th century. Among these have been increases in labor force participation rates and relative earnings of women, earlier retirement of men, and increases in earnings inequality across a number of dimensions. These patterns will affect not only the separate earnings histories of men and women as they reach retirement age, but also the joint earnings histories of couples. In the following section we discuss these trends and how well they are captured in the cohorts model.

Increased Labor Force Participation of Women.

An important change for projecting future earnings histories has been the dramatic rise in labor force participation rates of women ages 20 to 64 over the last four decades. In 1950, when women in the 1930 cohort were 20 years old, the labor force participation rate for women ages 20 to 24 was 44 percent (see Table 2-1). By 1970, the participation rate for women of the same age
was up to 56 percent, rising to 72 percent by 1990. Similar patterns apply to other age groups. For example, in 1970, when women in the 1930 cohort were age 40, the labor force participation rate for women ages 40 to 44 was 52 percent, compared with a participation rate of 78 percent for women ages 40 to 44 in 1990. Clearly age-specific participation rates for women in the 1930 cohort do not give a correct picture of rates for women in later cohorts.

Table 2-1

Percent of Individuals in the Labor Force by Gender, Age, and Year

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>20-24</td>
<td>83%</td>
<td>86%</td>
<td>81%</td>
<td>83%</td>
<td>81%</td>
</tr>
<tr>
<td></td>
<td>25-29</td>
<td>91%</td>
<td>94%</td>
<td>93%</td>
<td>92%</td>
<td>91%</td>
</tr>
<tr>
<td></td>
<td>30-34</td>
<td>95%</td>
<td>96%</td>
<td>95%</td>
<td>94%</td>
<td>92%</td>
</tr>
<tr>
<td></td>
<td>35-39</td>
<td>96%</td>
<td>96%</td>
<td>95%</td>
<td>94%</td>
<td>92%</td>
</tr>
<tr>
<td></td>
<td>40-44</td>
<td>95%</td>
<td>95%</td>
<td>95%</td>
<td>94%</td>
<td>92%</td>
</tr>
<tr>
<td></td>
<td>45-49</td>
<td>94%</td>
<td>94%</td>
<td>93%</td>
<td>92%</td>
<td>91%</td>
</tr>
<tr>
<td></td>
<td>50-54</td>
<td>91%</td>
<td>92%</td>
<td>91%</td>
<td>88%</td>
<td>88%</td>
</tr>
<tr>
<td></td>
<td>55-59</td>
<td>87%</td>
<td>88%</td>
<td>87%</td>
<td>81%</td>
<td>79%</td>
</tr>
<tr>
<td></td>
<td>60-64</td>
<td>80%</td>
<td>78%</td>
<td>73%</td>
<td>60%</td>
<td>55%</td>
</tr>
<tr>
<td></td>
<td>65-69</td>
<td>60%</td>
<td>44%</td>
<td>39%</td>
<td>29%</td>
<td>28%</td>
</tr>
<tr>
<td></td>
<td>70-74</td>
<td>38%</td>
<td>29%</td>
<td>22%</td>
<td>18%</td>
<td>17%</td>
</tr>
<tr>
<td>Female</td>
<td>20-24</td>
<td>44%</td>
<td>45%</td>
<td>56%</td>
<td>68%</td>
<td>72%</td>
</tr>
<tr>
<td></td>
<td>25-29</td>
<td>32%</td>
<td>35%</td>
<td>45%</td>
<td>66%</td>
<td>74%</td>
</tr>
<tr>
<td></td>
<td>30-34</td>
<td>31%</td>
<td>35%</td>
<td>44%</td>
<td>63%</td>
<td>73%</td>
</tr>
<tr>
<td></td>
<td>35-39</td>
<td>33%</td>
<td>40%</td>
<td>48%</td>
<td>65%</td>
<td>76%</td>
</tr>
<tr>
<td></td>
<td>40-44</td>
<td>36%</td>
<td>45%</td>
<td>52%</td>
<td>65%</td>
<td>78%</td>
</tr>
<tr>
<td></td>
<td>45-49</td>
<td>35%</td>
<td>47%</td>
<td>53%</td>
<td>61%</td>
<td>74%</td>
</tr>
<tr>
<td></td>
<td>50-54</td>
<td>31%</td>
<td>46%</td>
<td>52%</td>
<td>56%</td>
<td>67%</td>
</tr>
<tr>
<td></td>
<td>55-59</td>
<td>26%</td>
<td>40%</td>
<td>48%</td>
<td>49%</td>
<td>56%</td>
</tr>
<tr>
<td></td>
<td>60-64</td>
<td>21%</td>
<td>29%</td>
<td>36%</td>
<td>34%</td>
<td>36%</td>
</tr>
<tr>
<td></td>
<td>65-69</td>
<td>13%</td>
<td>16%</td>
<td>17%</td>
<td>15%</td>
<td>17%</td>
</tr>
<tr>
<td></td>
<td>70-74</td>
<td>7%</td>
<td>10%</td>
<td>9%</td>
<td>8%</td>
<td>8%</td>
</tr>
</tbody>
</table>

Source: The Urban Institute tabulations of the Bureau of the Census Decennial Census.
Bolded numbers highlight the 1930 cohort through the decades.

The model adjusts female labor force participation rates by controlling to age-specific rates from the 1993 Summary Earnings Records. The adjustment method will overstate participation rates for cohorts born prior to 1970 and may understate future participation rates for cohorts born after 1970 if trends in female labor force participation continue. For cohorts born after 1930 but before 1970, the method disregards information on their partial actual earnings histories and uses instead cross-sectional data on mixed cohorts to make the adjustments.
tends to over adjust participation rates for earlier cohorts. For example, the DER labor force adjustment would assign positive labor force participation for 72 percent of 20- to 24-year-olds regardless of the simulation cohort. While 72 percent participation is a good estimate for the 1970 birth cohort, it overstates participation for the 1960 cohort by 4 percentage points, the 1950 cohort by 16 percentage points, and the 1930 cohort by 28 percentage points. This over adjustment is true for all age groups in all cohorts born before 1970. If DER is interested in cohorts born before 1970, then at a minimum, DER should align the labor force participation rates to year- and cohort-specific levels rather than 1993 cross-sectional levels.

Because the focus of the PCM is on the long-term, DER has a greater interest in projected labor force participation rates for cohorts just now entering the labor force and those that will enter in the future. The adjustment process for future female labor force participation assumes that age specific rates in place in 1993 will persist into the future, for example, that the labor force participation rate for women age 60 to 64 will be 36 percent in all future cohorts. The adjustment method likely understates future labor force participation rates for women at all ages. While it is impossible to know what will occur, it is probable that female labor force participation at younger ages will continue to increase, especially given current “make work pay” and “welfare to work” government policies. Even if the trend levels-off, it is reasonable to assume that the participation rate for older women will catch-up with the increases in rates at earlier ages as members of the 1940-1960 cohorts age.

DER should consider changing the adjustment method for female labor force participation rates to allow for increases in rates beyond those observed in 1993, at least for older women. The design of the PCM, however, allows users to substitute alternative adjustments to female labor force participation rates rather easily. Rather than basing its analysis on a single projection, DER should take advantage of this feature of the model by simulating Social Security proposals using different assumptions about future female participation rates.

**Decreased Labor Force Participation of Older Men.**

Over the last 4 decades, labor force participation rates for men age 55 and over have decreased markedly. In 1970, the rate for men age 60 to 64 was 73 percent. By 1990, labor force participation rates for men age 60 to 64 had dropped to only 55 percent. Of course, we do not know yet what will happen to labor force participation rates at older ages for future cohorts. There is some debate about whether the trend towards early retirement from the labor force will continue or reverse (Costa, 1999; Quinn, 1998, 1999). If labor force participation continues to decrease for older men, then the PCM will overstate earnings and Social Security benefits for future cohorts. If the trend reverses as the demand for older workers increase, then the model will understate earnings and future benefits. Given the uncertain future of male labor force participation, some sensitivity analysis seems warranted.
Over the last 4 decades, labor force participation rates for men under age 55 have also decreased, but less so than at older ages. In 1960, the labor force participation rate for men age 30 to 34 was 96 percent. In 1990, this percent had dropped to 92 percent. This difference is fairly persistent throughout men’s prime working ages. While not a significant decline, lower labor force participation rates for younger men may result in lower future Social Security benefits, especially under proposals that would increase the number of computation years for calculating average monthly earnings. The PCM makes no adjustments to male labor force participation of younger men in cohorts born after 1930, and thus may tend to overstate benefits for these later cohorts.

**Increased Relative Earnings of Women.**

Relative median earnings of women rose for all but the youngest and oldest age groups during the working lifetime of the 1930 cohort, with most of the increase occurring between 1970 and 1990 (see Table 2-2). In 1960, when members of the 1930 cohorts were 30 years old, median earnings for women age 30 to 34 were 62 percent of the overall median earnings, compared to 94 percent of the overall median by 1990. Relative median earnings for younger and middle-aged women in 1970 and 1980 were also lower than in 1990.

The PCM adjusts relative earnings for women in all cohorts to reflect age-specific relative earnings for 1993. As with labor force participation rates, adjusting relative earnings to age-specific rates in 1993 will tend to overstate relative earnings for cohorts born after 1930 and before 1970. DER should align the earnings of cohorts born after 1930 but before 1970 to year-and cohort-specific levels rather than 1993 cross-sectional levels. The assumption that age-specific relative earnings of women in future cohorts will remain fixed at 1993 levels seems less likely than a continued narrowing of relative differences. This seems particularly true for future relative earnings of older women as women in the 1940-1960 cohorts, who have had higher relative earnings throughout their working years, reach older ages.

**Decreased Relative Earnings of Younger Men.**

Over the last 4 decades, earnings of younger men have fallen behind the earnings of older men. In 1960, men ages 30 to 34 earned the 1.6 times the median wage (see Table 2-2). This was the same ratio to the median wage as the ratio for men ages 50 to 54. By 1990, men ages 30 to 34 earned only 1.5 times the median wage, while men ages 50 to 54 earned almost 1.9 times the median wage. Relative earnings growth also differs by cohorts. Men in the 1930 cohort
Table 2-2
Median Wage and Salary Income Divided by the Annual Median Wage by Gender, Age, and Year
For All Wage and Salary Workers

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-24</td>
<td>0.80</td>
<td>0.78</td>
<td>0.71</td>
<td>0.81</td>
<td>0.63</td>
</tr>
<tr>
<td>25-29</td>
<td>1.24</td>
<td>1.35</td>
<td>1.49</td>
<td>1.38</td>
<td>1.19</td>
</tr>
<tr>
<td>30-34</td>
<td>1.39</td>
<td>1.60</td>
<td>1.68</td>
<td>1.78</td>
<td>1.50</td>
</tr>
<tr>
<td>35-39</td>
<td>1.49</td>
<td>1.67</td>
<td>1.80</td>
<td>2.02</td>
<td>1.75</td>
</tr>
<tr>
<td>40-44</td>
<td>1.49</td>
<td>1.63</td>
<td>1.87</td>
<td>2.11</td>
<td>1.88</td>
</tr>
<tr>
<td>45-49</td>
<td>1.44</td>
<td>1.60</td>
<td>1.80</td>
<td>2.11</td>
<td>2.00</td>
</tr>
<tr>
<td>50-54</td>
<td>1.39</td>
<td>1.60</td>
<td>1.70</td>
<td>2.06</td>
<td>1.88</td>
</tr>
<tr>
<td>55-59</td>
<td>1.34</td>
<td>1.51</td>
<td>1.62</td>
<td>1.98</td>
<td>1.84</td>
</tr>
<tr>
<td>60-64</td>
<td>1.24</td>
<td>1.44</td>
<td>1.45</td>
<td>1.67</td>
<td>1.50</td>
</tr>
<tr>
<td>65-69</td>
<td>1.00</td>
<td>1.00</td>
<td>0.84</td>
<td>0.62</td>
<td>0.63</td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-24</td>
<td>0.61</td>
<td>0.52</td>
<td>0.53</td>
<td>0.56</td>
<td>0.47</td>
</tr>
<tr>
<td>25-29</td>
<td>0.71</td>
<td>0.62</td>
<td>0.73</td>
<td>0.87</td>
<td>0.88</td>
</tr>
<tr>
<td>30-34</td>
<td>0.66</td>
<td>0.62</td>
<td>0.63</td>
<td>0.87</td>
<td>0.94</td>
</tr>
<tr>
<td>35-39</td>
<td>0.66</td>
<td>0.65</td>
<td>0.65</td>
<td>0.81</td>
<td>0.97</td>
</tr>
<tr>
<td>40-44</td>
<td>0.76</td>
<td>0.65</td>
<td>0.73</td>
<td>0.84</td>
<td>1.00</td>
</tr>
<tr>
<td>45-49</td>
<td>0.76</td>
<td>0.68</td>
<td>0.77</td>
<td>0.87</td>
<td>1.00</td>
</tr>
<tr>
<td>50-54</td>
<td>0.71</td>
<td>0.75</td>
<td>0.77</td>
<td>0.88</td>
<td>0.94</td>
</tr>
<tr>
<td>55-59</td>
<td>0.61</td>
<td>0.71</td>
<td>0.79</td>
<td>0.89</td>
<td>0.88</td>
</tr>
<tr>
<td>60-64</td>
<td>0.56</td>
<td>0.65</td>
<td>0.73</td>
<td>0.78</td>
<td>0.75</td>
</tr>
<tr>
<td>65-69</td>
<td>0.46</td>
<td>0.37</td>
<td>0.36</td>
<td>0.42</td>
<td>0.44</td>
</tr>
</tbody>
</table>

Source: The Urban Institute tabulations of the Bureau of the Census Decennial Census.
Bolded numbers highlight the 1930 cohort through the decades.

experienced a 17 percent increase in median earnings between ages 30 and 40, while men in the 1950 cohort experienced only a 6 percent increase. Men in the 1930 cohort experienced a 10 percent increase in median earnings between ages 40 and 50, while men in the 1940 cohort experienced an 11 percent decrease in median earnings between ages 40 and 50. These changes reflect changes in the age-earnings profiles of men in the 1930 cohort compared with men in later cohorts. The PCM maintains the shape of the age-earnings profiles for men in the 1930 cohort and thus may misstate Social Security entitlement for future cohorts. While we do not know if the earnings profiles of later cohorts will catch up with earnings profiles of the 1930 cohort, we do know that they are starting off on a different profile.
**Increased Share of Wives Earnings.**

As women’s labor force participation and relative earnings increased, so has their share of earnings as part of married couples. In 1960, women ages 30 to 34 contributed 13 percent of married couples earnings, on average (see Table 2-3). By 1990, women in the same age group accounted for 30 percent of married couple earnings. In 1960, women ages 40 to 44 contributed 18 percent of married couples earnings, on average. By 1990, women in the same age group accounted for 31 percent of a couple’s earnings, on average.

**Increased Correlation of Spouses’ Earnings.**

The correlation between husbands' and wives' earnings has increased. Using earnings data from the March CPS for 1968 to 1994, Cancian and Reed (1999) report an increase in the correlation of spouses’ earnings beginning in the mid-1970s, although the correlation was still relatively low (approximately .25 at its peak in 1990). There has been a particularly large increase in the correlation between husbands’ and wives’ earnings for husbands with a college education. For marriages to college educated husbands in 1960, the correlation between earnings of husbands and wives was 0.04. By 1990, this correlation had increased to 0.13 (see Table 2-4), though most of that increase occurred between 1980 and 1990. Some of this increase in correlation is due to increasing labor force participation and rising relative earnings of women, and thus would be captured as part of the adjustment process in the cohorts model. However, the trend also reflects other influences such as the increased tendency for higher educated men to marry higher educated women (see Mare, 1991).

The percent of newlywed husbands with a college education who marry college educated women has increased over the last 4 decades (see Table 2-5). In 1960, when the men in the 1930 cohort were age 30, 33 percent of college educated newlywed men (13-15 years of education) married college educated women. By the mid-1980s, this percent had increased to 42 percent.

The increase in selective educational mating is even higher among college graduates and post graduates. In 1960, 39 percent of newlywed men with a college degree (16 or more years of education) married women with a college degree. By the mid-1980s, this percent had increased to 61 percent.
Table 2-3
Share of Couple's Earnings by Gender, Age, and Year
For Couples with Positive Wage and Salary Income

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>20-24</td>
<td>81%</td>
<td>75%</td>
<td>73%</td>
<td>69%</td>
</tr>
<tr>
<td></td>
<td>25-29</td>
<td>85%</td>
<td>81%</td>
<td>74%</td>
<td>69%</td>
</tr>
<tr>
<td></td>
<td>30-34</td>
<td>88%</td>
<td>85%</td>
<td>77%</td>
<td>71%</td>
</tr>
<tr>
<td></td>
<td>35-39</td>
<td>87%</td>
<td>85%</td>
<td>78%</td>
<td>71%</td>
</tr>
<tr>
<td></td>
<td>40-44</td>
<td>84%</td>
<td>83%</td>
<td>77%</td>
<td>70%</td>
</tr>
<tr>
<td></td>
<td>45-49</td>
<td>82%</td>
<td>80%</td>
<td>76%</td>
<td>70%</td>
</tr>
<tr>
<td></td>
<td>50-54</td>
<td>81%</td>
<td>79%</td>
<td>76%</td>
<td>70%</td>
</tr>
<tr>
<td></td>
<td>55-59</td>
<td>81%</td>
<td>78%</td>
<td>75%</td>
<td>69%</td>
</tr>
<tr>
<td></td>
<td>60-64</td>
<td>81%</td>
<td>77%</td>
<td>72%</td>
<td><strong>65%</strong></td>
</tr>
<tr>
<td></td>
<td>65-69</td>
<td>76%</td>
<td>68%</td>
<td>60%</td>
<td>57%</td>
</tr>
<tr>
<td>Female</td>
<td>20-24</td>
<td>18%</td>
<td>23%</td>
<td>27%</td>
<td>30%</td>
</tr>
<tr>
<td></td>
<td>25-29</td>
<td>13%</td>
<td>17%</td>
<td>26%</td>
<td>31%</td>
</tr>
<tr>
<td></td>
<td>30-34</td>
<td>13%</td>
<td>15%</td>
<td>23%</td>
<td>30%</td>
</tr>
<tr>
<td></td>
<td>35-39</td>
<td>15%</td>
<td>17%</td>
<td>23%</td>
<td>31%</td>
</tr>
<tr>
<td></td>
<td>40-44</td>
<td>18%</td>
<td>19%</td>
<td>25%</td>
<td>31%</td>
</tr>
<tr>
<td></td>
<td>45-49</td>
<td>20%</td>
<td>21%</td>
<td>26%</td>
<td>32%</td>
</tr>
<tr>
<td></td>
<td>50-54</td>
<td>22%</td>
<td>23%</td>
<td><strong>27%</strong></td>
<td>33%</td>
</tr>
<tr>
<td></td>
<td>55-59</td>
<td>21%</td>
<td>25%</td>
<td>29%</td>
<td>35%</td>
</tr>
<tr>
<td></td>
<td>60-64</td>
<td>19%</td>
<td>28%</td>
<td>32%</td>
<td><strong>38%</strong></td>
</tr>
<tr>
<td></td>
<td>65-69</td>
<td>16%</td>
<td>23%</td>
<td>27%</td>
<td>32%</td>
</tr>
</tbody>
</table>

Source: The Urban Institute tabulations of the Bureau of the Census Decennial Census.
Note: Universe includes all couples with positive couple earnings. Husbands and wives may be different ages so percents do not add up to 100.
### Table 2-4

**Correlation Between Husbands' and Wives' Earnings**

For Husbands and Wives with Positive Wage and Salary Income
by Educational Attainment of the Husband and Year

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td></td>
<td>0.175</td>
<td>0.128</td>
<td>0.124</td>
<td>0.205</td>
</tr>
<tr>
<td>High School Dropout</td>
<td></td>
<td>0.214</td>
<td>0.153</td>
<td>0.155</td>
<td>0.204</td>
</tr>
<tr>
<td>High School Graduate</td>
<td></td>
<td>0.107</td>
<td>0.088</td>
<td>0.103</td>
<td>0.163</td>
</tr>
<tr>
<td>College Graduate</td>
<td></td>
<td>0.044</td>
<td>0.045</td>
<td>0.039</td>
<td>0.126</td>
</tr>
</tbody>
</table>

Source: The Urban Institute tabulations of the Bureau of the Census Decennial Census.

Note: Universe includes all couples with positive husband and wife earnings.

### Table 2-5

**Percent of College Educated Newlywed Husbands with College Educated Wives**
by Year and Educational Attainment

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent of College Educated Newlywed Husbands with College Educated Wives(^1)</td>
<td>25%</td>
<td>33%</td>
<td>39%</td>
<td>42%</td>
<td>42%</td>
</tr>
<tr>
<td>Percent of College Graduate Newlywed Husbands with College Graduate Wives(^2)</td>
<td>32%</td>
<td>39%</td>
<td>48%</td>
<td>54%</td>
<td>61%</td>
</tr>
</tbody>
</table>

Source: The Urban Institute tabulations from Table 2 from Mare (1991).

1/ Thirteen to 15 years of education.
2/ Sixteen or more years of education.
While these changes are due in large part to an increase in the relative number of college educated women, this increase in education-selective mating will increase the inequality of income. Although the PCM adjusts female labor force participation rates and relative earnings of the 1930 cohort to 1993 levels, the adjustment process does not control for the characteristics of the husband. The relative earnings of wives of less educated men are adjusted by the same factors (controlling for age) as the relative earnings of wives of college educated men, and thus the adjustment does not capture this widening of the income distribution.

**Increased Earnings Inequality.**

Wage and earnings inequality increased during the 1980s, particularly for men, but also for women. Earnings of younger workers fell relative to the earnings of middle-aged and older workers. The gap between earnings of more and less educated workers widened. After controlling for age, education, and other factors, the within-group variance of earnings also increased (see Levy and Murnane, 1992; Bound and Johnson, 1995).

The gap between earnings of more educated workers and less educated workers widened. In 1960, men age 30 to 34 with less than a high school education were earning 62 percent of the median wage of men the same age but with a college degree. By 1990, that ratio had fallen to 42 percent (see Table 2-6). In 1960, men age 30 to 34 with a high school education were earning 80 percent of the median wage of college educated men the same age. By 1990, that ratio had fallen to 66 percent (see Table 2-7). Relative earnings for college educated women has also increased, with most of the gains coming between 1980 and 1990.

Some of the increase in the gap between earnings of college and high school-educated workers reflects shifts in relative supplies. During the 1970s the premium for college educated workers fell as the supply of recent college graduates grew much faster than the supply of younger workers with high school degrees. The college premium rose in the 1980s as the rate of growth in the supply of college educated workers slowed dramatically (Freeman, 1976; Welch, 1979; Levy and Murnane, 1992). An increase in the relative demand for college educated workers also contributed to the increase in the college premium. Non-neutral technological change led to an increase in the demand for college educated workers, while at the same time the demand for high-school educated workers declined as more manufacturing work formerly done by high-school educated workers in the U.S. shifted to overseas producers (Bartel and Lichtenberg, 1987; Blackburn, Bloom, and Freeman, 1990; Katz and Murphy, 1992, Levy and Murnane, 1992).

It is not clear whether these trends will continue into the future. Evidence suggests that the trend towards increased earnings inequality slowed in the early 1990s (Lerman, 1997), but recent individual income tax return data and W-2 forms show a widening income gap in 1995,
### Table 2-6
Median Wage and Salary Income of High School Dropouts Divided by the Median Wage and Salary Income of College Graduates by Gender, Age, and Year

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>20-24</td>
<td>1.35</td>
<td>0.96</td>
<td>1.09</td>
<td>0.96</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>25-29</td>
<td>0.73</td>
<td>0.70</td>
<td>0.72</td>
<td>0.64</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>30-34</td>
<td>0.66</td>
<td>0.62</td>
<td>0.59</td>
<td>0.57</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>35-39</td>
<td>0.57</td>
<td>0.57</td>
<td>0.54</td>
<td>0.50</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>40-44</td>
<td>0.58</td>
<td>0.56</td>
<td>0.51</td>
<td>0.50</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>45-49</td>
<td>0.58</td>
<td>0.55</td>
<td>0.49</td>
<td>0.52</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>50-54</td>
<td>0.54</td>
<td>0.55</td>
<td>0.50</td>
<td>0.54</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>55-59</td>
<td>0.60</td>
<td>0.53</td>
<td>0.52</td>
<td>0.54</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>60-64</td>
<td>0.60</td>
<td>0.56</td>
<td>0.50</td>
<td>0.51</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>65-69</td>
<td>0.55</td>
<td>0.47</td>
<td>0.34</td>
<td>0.33</td>
<td>0.33</td>
</tr>
<tr>
<td>Female</td>
<td>20-24</td>
<td>0.55</td>
<td>0.40</td>
<td>0.45</td>
<td>0.53</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>25-29</td>
<td>0.40</td>
<td>0.32</td>
<td>0.34</td>
<td>0.44</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>30-34</td>
<td>0.40</td>
<td>0.38</td>
<td>0.40</td>
<td>0.43</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>35-39</td>
<td>0.43</td>
<td>0.41</td>
<td>0.44</td>
<td>0.48</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>40-44</td>
<td>0.45</td>
<td>0.41</td>
<td>0.47</td>
<td>0.50</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>45-49</td>
<td>0.42</td>
<td>0.40</td>
<td>0.43</td>
<td>0.47</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>50-54</td>
<td>0.38</td>
<td>0.40</td>
<td>0.40</td>
<td>0.44</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>55-59</td>
<td>0.38</td>
<td>0.39</td>
<td>0.40</td>
<td>0.43</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>60-64</td>
<td>0.37</td>
<td>0.33</td>
<td>0.37</td>
<td>0.43</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>65-69</td>
<td>0.29</td>
<td>0.30</td>
<td>0.31</td>
<td>0.66</td>
<td>0.53</td>
</tr>
</tbody>
</table>

Source: The Urban Institute tabulations of the Bureau of the Census Decennial Census.
### Table 2-7
Median Wage and Salary Income of High School Graduates
Divided by the Median Wage and Salary Income of College Graduates
by Gender, Age, and Year

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>20-24</td>
<td>1.43</td>
<td>1.04</td>
<td>0.97</td>
<td>1.08</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>25-29</td>
<td>0.93</td>
<td>0.92</td>
<td>0.88</td>
<td>0.89</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>30-34</td>
<td>0.79</td>
<td>0.80</td>
<td>0.76</td>
<td>0.82</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>35-39</td>
<td>0.75</td>
<td>0.75</td>
<td>0.69</td>
<td>0.75</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>40-44</td>
<td>0.78</td>
<td>0.73</td>
<td>0.67</td>
<td>0.72</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>45-49</td>
<td>0.78</td>
<td>0.72</td>
<td>0.66</td>
<td>0.71</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>50-54</td>
<td>0.72</td>
<td>0.71</td>
<td>0.64</td>
<td>0.70</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>55-59</td>
<td>0.84</td>
<td>0.70</td>
<td>0.68</td>
<td>0.70</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>60-64</td>
<td>0.78</td>
<td>0.72</td>
<td>0.67</td>
<td>0.68</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>65-69</td>
<td>0.77</td>
<td>0.74</td>
<td>0.60</td>
<td>0.52</td>
<td>0.39</td>
</tr>
<tr>
<td>Female</td>
<td>20-24</td>
<td>1.00</td>
<td>0.91</td>
<td>0.80</td>
<td>0.85</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>25-29</td>
<td>0.81</td>
<td>0.63</td>
<td>0.59</td>
<td>0.70</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>30-34</td>
<td>0.74</td>
<td>0.63</td>
<td>0.54</td>
<td>0.64</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>35-39</td>
<td>0.76</td>
<td>0.65</td>
<td>0.57</td>
<td>0.62</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>40-44</td>
<td>0.73</td>
<td>0.63</td>
<td>0.59</td>
<td>0.64</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>45-49</td>
<td>0.71</td>
<td>0.61</td>
<td>0.57</td>
<td>0.62</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>50-54</td>
<td>0.64</td>
<td>0.66</td>
<td>0.55</td>
<td>0.59</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>55-59</td>
<td>0.67</td>
<td>0.62</td>
<td>0.56</td>
<td>0.61</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>60-64</td>
<td>0.69</td>
<td>0.60</td>
<td>0.57</td>
<td>0.62</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>65-69</td>
<td>0.49</td>
<td>0.55</td>
<td>0.50</td>
<td>0.89</td>
<td>0.67</td>
</tr>
</tbody>
</table>

Source: The Urban Institute tabulations of the Bureau of the Census Decennial Census.

1996, and 1997 (Utendorf, 1999, Cruiciano, 1998 and 1999). Even if the trend does not continue, however, lifetime earnings for cohorts whose peak earning years came during or after the 1980s will reflect the greater disparity in earnings that has already occurred. Projections based on the earnings histories of the 1930 cohort will understate the dispersion of lifetime earnings for future cohorts. Further, because lifetime earnings of projected future cohorts differ only in the scaling of average earnings, the projection method currently used in the PCM does not capture the important differences that exist among cohorts, such as between the early and late baby-boom generations.
Increased dispersion of lifetime earnings within and among cohorts will translate into increased dispersion of Social Security benefits and other income in retirement. If the increased dispersion in earnings only reflected changes in the upper tail of the earnings distribution there would be little impact on the dispersion of Social Security benefits, because of the annual maximum on earnings used to calculate benefits. However, the increased gap in earnings among older and younger men and between more and less educated workers has not been confined to high earners, and thus will result in relatively lower benefits for those in the bottom part of the distribution. Because of Social Security’s progressive benefit formula, Social Security benefits for lower income beneficiaries will be affected to a lesser degree than the differences in relative earnings, pensions, and saving, but the effect on these other income sources may be large.

Changing Macroeconomic Environment.

Members of the 1930 cohort came of working age in the 1950s and reached their peak earnings years in the 1980s. Future cohorts will experience a different set of economic conditions throughout their working lives. The cohorts model controls for the effect of a different macroeconomic environment by measuring all individual earnings relative to average earnings.

5. Alternative Methods For Projecting Earnings

The RFTOP suggested that an alternative method for modifying the cohort data would be to start with a later cohort and simulate the late-career earnings. This method has been successfully implemented for the Social Security Administration Model of Income in the Near Term (MINT). MINT projects future earnings for the 1931 to 1960 cohorts using separate fixed-effects regression models for men and women by educational attainment. The advantage of this method is that it preserves the actual earnings of all cohorts up until 1996, the last available SER observation. In its projections, it controls for individual-specific tendencies to be high, medium, or low earners based on observed earning histories. A disadvantage of this method is that it does not capture all of the heterogeneity that exists in individual earnings histories, particularly periods out of the labor force when earnings fall to zero.

Another option, not described in the RFTOP, would be to splice segments of observed work histories to construct a life-time work history for younger cohorts. For example, one could break work segments into 10-year age intervals (for example 20-30, 30-40, etc.) and normalize them by dividing by the year-specific average wage. Then for each cohort, one could use a statistical matching algorithm to find an individual from an earlier cohort with the same demographic characteristics (race, gender, marital status, educational attainment, number of children) and relative earnings over the 10-year intervals for which earnings are observed. This person’s earnings could then be spliced on to augment the observed earnings.
Splicing still relies on the observed late career earnings of older cohorts to project late career earnings for younger cohorts. For example, in projecting future earnings for women, it places the most weight on observations of women from older cohorts whose early and mid-career earnings patterns “best fit” the early and mid-career earnings patterns of women from younger cohorts. If earnings patterns are changing over the entire working career, then this method may not yield the best prediction of future earnings patterns.

The advantage of either of these methods (regression based projection or splicing) is that they use actual earnings for all cohorts up to the last observed period (1996) and then project future earnings. The observed earnings exactly measure relative changes in labor force participation, the distribution of earnings, earnings related to changes in educational attainment, marital status, and child rearing that have occurred over the last several decades. Using actual earnings would lessen the problem of censored earnings for younger cohorts. Both methods would retain the autocorrelation of individual earnings, and the correlation of earnings between spouses. Retention of these correlations is cited as one of the strengths of the current projected cohorts model, although it is not clear that either type of correlation in the 1930 cohort is correct for later cohorts.

Designing an appropriate test to compare alternative methods is difficult. One approach is to project near-term future earnings of cohorts where there are actual data, omitting the last few years of observed data. This test would likely find in favor of either the simulation or the splicing methods. Earnings of people from the 1960 cohort when they were 30 are probably a better predictor of earnings at age 33 than are earnings from the 1930 cohort for people age 33. The real test is which method gives a better prediction of earnings for people from the 1960 cohort when they are age 50 or 60, and that is an experiment that no one can perform.

Given the complex changes that have occurred in lifetime earnings over the last 4 decades, DER should perform some tests to validate the PCM projections. At a minimum, DER should compare projections for cohorts born after 1930 with the observed partial lifetime earnings histories of those cohorts. Significant variation between projected and actual lifetime earnings paths would suggest the types of modifications that might be needed to improve the projection method.

6. Demographic Projections

There have been important changes in demographic trends in the past decades. Most noticeably, there have been significant reductions in mortality and fertility, and changes in marital patterns that have increased the proportion of women who either never marry or divorce. The current version of the model makes no adjustments to the demographic characteristics of the 1930 cohort. Thus, the percentages of the population married, never married, widowed, or divorced at retirement for projected cohorts remain the same as they were in 1992 when the 1930 cohort
reached retirement age. Benefit entitlement for the projected cohorts in the PCM is based on the historical experience of the 1930 cohort. For later cohorts, more people will have experienced divorce and the percentage of beneficiaries receiving benefits as divorced spouses or surviving divorced spouses is likely to increase. Social Security retirement benefits also are affected by demographic changes after benefit entitlement such as divorce or remarriage, but the model does not allow for changes in marital status for beneficiaries.

Some of these changes can be captured by adjusting population weights for future retirement cohorts as DER plans to do. Other demographic adjustments may be more difficult. Adjusting survival probabilities for the 1930 cohort may require more than a simple reweighting of the sample. Workers from the 1930 cohort who die before reaching retirement age are not likely to have the same characteristics as those who survived to 1993.

Demographic changes have also changed earnings trajectories. For example, birth rates are much lower among more recent cohorts (see Figure 2-1). While modeling fertility itself is not important for the PCM (population growth can be captured by reweighting), decreases in fertility are associated with an increase in female labor force participation. The current adjustment to labor force participation rates may capture these effects. However, there have also been significant changes in the timing of births (see Figure 2-2). Delays in childbearing and increases in schooling are changing the pattern of years in which women have little or no earnings. They may also affect earnings growth patterns. Although the timing of years out of the labor force may not matter for some proposals, such as those that would increase the number of computation years, timing will matter for other options. For example, the returns from individual accounts will depend upon when investments in those accounts are made.
Figure 2-1
Central Birth Rates by Birth Cohort

Source: Vital Statistics Table 1-15

Figure 2-2
Central Birth Rates for First Live Births by Cohort
IV. THE PROJECTED COHORTS MODEL AS A FRAMEWORK FOR A MORE COMPREHENSIVE MODEL

A second aim of the Projected Cohorts Model is to provide a framework for more ambitious modeling of the distributional effects of Social Security reform proposals. These include incorporating behavioral feedbacks at the micro level and possible interactions with a macro model. DER should also consider how the PCM could be used to analyze the distributional consequences of reforms for total income and not just income from Social Security. If the model is to consider changes to the over-all economic well-being of retirees it must include saving and pension receipt. The PCM also should include a simple representation of the disability and SSI programs in order to capture important interactions with OASI.

1. Behavioral Response at the Micro Level

The behavioral responses that will likely have the biggest impact on the distributional effects of Social Security reform proposals are changes in the age of retirement and changes in saving behavior. The PCM does not currently model saving or retirement behavior.

The issues involved in modeling saving are discussed in detail in Chapter 3. Modeling saving in any microsimulation model is difficult, particularly if the goal is to capture the interaction between policy variables and saving at the household level. The cohort approach faces the same problems as other simulation models, but does provide some advantages for projecting saving in the context of a lifecycle model with perfect foresight, because the time path of future earnings is known for each worker. The PCM yields, for each individual, a lifetime series of earnings and Social Security benefits. Based on the known lifetime path of earnings and expected benefits for each worker, the model could solve for the allocation of consumption (saving) over time. A model that simulates earnings year-by-year would either need to simulate consumption based only on current and past income, or else project future expected earnings. Once the model departs from the assumption of perfect foresight, or allows the future time path of earnings to vary in response to changes in saving behavior, however, there is no longer any advantage to the cohort approach.

The current model does not simulate retirement behavior, but instead maintains the distribution of retirement ages observed in 1992. As DER notes, Social Security reforms that change the age of early entitlement, rather than the normal retirement age, will likely have the biggest impact on behavior, although changes in the normal retirement age could also delay retirement by reducing Social Security benefits at earlier ages. Some type of retirement behavior should be built into the model. Simple exogenous adjustments to retirement ages are straightforward to implement, such as a proportional shift in retirement ages to reflect changes in
Social Security retirement age rules. Modeling endogenous retirement behavior is much more complex, and should consider pensions and other savings as well as expected Social Security benefits. The cohorts approach provides no obvious advantage for modeling retirement compared to other microsimulation models such as MINT, DYNASIM, and CORSIM.

Some proposals that would radically alter retirement programs may also have an impact on pre-retirement earnings. The cohort approach does not have any advantage if labor supply is endogenous to policy changes. The time path of earnings needs to be recalculated each period as expected wages and lifetime income changed. Because the model does not have separate information on wages and hours worked, DER will need to consider some way to simulate a labor supply response that relies only on earnings.

A potentially important behavioral response to changes in Social Security rules will be the reaction by private pension providers. That type of model, however, is best left as a separate endeavor.

2. Macroeconomic Feedback Effects

Micro simulation models typically incorporate macroeconomic effects by calibration. The general approach is to run the model and compare its predictions against a set of macroeconomic targets. Incomes, participation rates, and other input variables are then adjusted until the model hits each target. Calibration, however, is not adequate in all circumstances and is not necessary for answering some questions.

Proper calibration is important when the distributional effects of a proposal are sensitive to changes in baseline macro variables. For example, a comparison of lifetime benefits across different cohorts will depend not only on the distribution of earnings within each cohort but also on the growth in average earnings over time. Proper calibration is less important, however, when the distributional effects of a proposal are not sensitive to changes in baseline macro variables. For example, the first-order distributional effects within a cohort of changes in the Social Security benefit formula would not depend on the baseline level of macroeconomic variables.

Moreover, proper calibration by itself does not ensure that the distributional effects of a policy proposal will be measured correctly. For example, if a policy changes relative factor prices significantly, one would have to take into account the resulting change in incomes to measure distributional effects properly. Models that do not incorporate feedback effects explicitly will not capture these distributional effects, even if they are initially calibrated to hit reasonable macro targets.
Baekgaard (1995) identifies four approaches for macro-micro linkages in simulation models:

1) The top-down approach in which a preliminary micro aggregate outcome is adjusted to match an externally supplied variable from the macro model.

2) The bottom-up approach in which aggregate totals from the micro model are used to adjust outputs from the macro model.

3) The recursive linkage in which the two models are linked through two-way lagged interaction over a short simulation period.

4) The iterative or simultaneous approach in which the outcomes of the micro and macro models are solved simultaneously within each period, generally through an iterative scheme.

A fifth type of linkage involves solving micro and macro models separately for the entire simulation process, calibrating the models, and then resolving until the models converge.

Chapter 4 presents a general discussion of possible macro-micro linkages. In its current form, the projected cohorts model presents limited possibilities for linkages to a macro model. Because the model does not cover the entire population, it cannot produce aggregate economic variables such as total labor supply and earnings, or program variables such as total Social Security taxes and benefits. These values are vital inputs to a macro model as part of a recursive or simultaneous solution with a macro model. This would leave a top-down approach as the only feasible linkage unless the PCM were extended to include the entire population. If the projected cohorts model were extended to include the entire population, any of the five macro-micro linkages are possible, although the structure of the model would lend itself to the fifth type of linkage.

V. SUMMARY/CONCLUSIONS

The cohorts model provides a way to analyze the distributional effects of certain types of Social Security reform proposals. Because it is based on actual earnings and demographic histories for a cohort that has reached the age of retirement, it provides much richer possibilities than typical representative worker models. There are problems with missing data and censored earnings histories in the CPS survey/SSA administrative data match, but these are not insurmountable. DER has developed a sophisticated method for correcting censored earnings. Imputations for missing earnings histories of deceased workers and former spouses, while not trivial, can be developed and implemented.
The strength of the cohorts model is that it is able to simulate an entire lifetime of earnings for each member of a cohort, using a simple computational process. Because of the simple process, the model can be readily used for sensitivity analysis by repeating simulations using a range of assumptions about future labor force participation rates and relative earnings. In the same way, the simplicity of the model allows for extremely rapid simulations of successive cohorts. The trade-off is that the simple adjustment used in the model will not adequately project the earnings histories of future cohorts if their lifetime earnings do not resemble earnings history data from the 1930 cohort. It is not clear that, with the adjustments currently implemented, the model captures important changes in earnings patterns for later cohorts. In particular, the model may miss changes in female labor force participation rates, the correlation of earnings between spouses, and increased earnings inequality. Because projected cohorts are scalar reflections of each other, the model misses important differences among future cohorts. Further, the model, at present, does not capture any changes in demographic trends. While some of those trends, such as the decrease in fertility, may not be as important for the intended uses of the model, others, such as the improvements in mortality and changing marital status, do have implications for the distribution of reform proposals.

The best way to address these problems would be by using information available in the earnings histories of later cohorts. Those data reflect economic and demographic trends that have already taken place. Not only would they improve the projection of complete lifetime histories for those cohorts, but they also would improve the projections for cohorts coming into the labor force now and in the future.

No microsimulation model exists that can answer all questions regarding the impact of policy change on individuals and the economy as a whole. The PCM, while not perfect, provides a useful basis for answering questions about how policy changes affect the distribution of Social Security benefits. In its current state, it cannot answer a wider range of distributional questions. The lack of pension, saving, and other income sources limits its usefulness for evaluating questions of income adequacy of future beneficiaries. This should be a priority area for future extensions of the model. The lack of the full population limits its usefulness for evaluating questions of trust fund balances and national saving, total employment, or domestic output. While these are all important issues, they are best addressed by models specifically designed for those purposes.

Further enhancements to the PCM might include modeling saving and retirement behavior and allowing for macroeconomic feedback effects. The cohort approach has both advantages and disadvantage for these enhancements, compared with other microsimulation models. The cohort approach has advantages for projecting savings in a way that depends on future or expected earnings because it simulates an entire earnings history for each worker. However, simulating other types of behavioral responses, such as changes in retirement age and pre-retirement
earnings, or allowing for macroeconomic feedback effects require more complex projection algorithms that depart from the simple adjustments that are the basis of the cohort model.
REFERENCES TO CHAPTER 2


CHAPTER 3
MODELING SAVING BEHAVIOR

I. INTRODUCTION

There are two key reasons for including saving and pension accumulation in a model designed to analyze the distributional effects of Social Security reform proposals: first to provide a complete picture of the economic resources of the retired population, and second, to capture important behavioral interactions that may affect some of the distributional consequences of changes in Social Security benefits.

Modeling saving by future cohorts of retirees is not an easy task, even if it is limited to projecting the amount of saving under current policies. One approach is simply to extrapolate estimates of saving by the current population to future cohorts. Even this is not straightforward. Good measures of current saving by households are hard to come by. Extrapolating forward to future cohorts requires separating out the effects of age, cohort, and time period in current estimates.

Another alternative is to use a model of economic behavior to project future saving. This has its own difficulties. First, the modeler must reconcile theoretical models of saving with seemingly inconsistent empirical evidence on wealth accumulation. Second, the modeler must account for different tastes for saving in behavior among individuals with similar socioeconomic attributes.

If the model is intended to capture important behavioral feedbacks on private and pension saving from changes in Social Security, the problem is even more complicated. It has proven difficult to measure the effects of the current Social Security program on private saving. Various theoretical models are at odds concerning the predicted outcome, and empirical estimates differ. Research has also come to differing conclusions concerning the effects of targeted retirement saving programs--a prominent feature of some reform proposals--on net saving.

This suggests that any attempt to model saving behavior -- over time or across cohorts -- will be a difficult challenge. While an eye should be kept on what can be learned from the vast literature on saving and consumption -- research that has grown rapidly in recent years - this literature is often inconclusive and does not offer a straightforward guide to a practical modeling strategy. Experience has also shown that attempts to model saving behavior by relying on simple and elegant theories will often lead to results that are at odds with empirical data. Nevertheless,
some guidance can be gained from the literature as to what variables are important to individual saving decisions and how they might enter into a microsimulation model of wealth accumulation.

A modeling effort should be guided by the types of questions one would like a model to answer and the types of analysis that will be required. This will, in turn, often indicate the types of data that will be needed to support model development, narrow the list of feasible approaches and suggest the structure of the model.

II. ISSUES IN MODELING SAVING BEHAVIOR

In the third quarter of 1999, the U.S. personal saving rate reached an all-time low of 2.1 percent. Viewed from an historical perspective, this can be seen as the continuation of a trend in falling saving rates that has persisted over the last twenty years. That this has occurred during the strongest and longest economic expansion in history has policymakers concerned about the trend and the income security of future retirees.28 Much of the current debate over how to reform the Social Security system focuses on how to raise this saving rate.

In a recent paper, Gale and Sabelhaus (1999) demonstrate that how one interprets recent evidence on trends in private saving rates in the United States depends, in part, on the particular definition of saving that is used. This has important implications for modeling saving behavior because a desirable goal of such a model is to be able to explain existing facts about saving and wealth accumulation. Most measures of private saving - including the measure used in the National Income and Product Accounts (NIPA), for example - exclude capital gains from the calculation. But since this appreciation is reflected in the asset values of retirees and, thus, available to finance future consumption, it should be included in a measure of saving designed to reflect existing resources. Choosing the appropriate definition of saving will be an important first step in the modeling effort. Additionally, the literature on saving and consumption – if it is to be used as a guide to the modeling - needs to be evaluated in terms of the definition of saving that is assumed.

1. What Questions Do We Want to Answer?

How one models saving behavior should be dictated by the policy questions that are likely to be considered. It is clear that there are at least two related issues that need to be addressed. The first issue is the income security of future retirees and how, under existing policies, they are likely to fare. This means that estimates of total financial resources available to finance consumption in retirement will be needed for future cohorts. These resources include Social Security wealth, pension wealth in the form of DB and DC plan accruals and personal saving including housing wealth. With these estimates in place, questions relating to poverty rates and income distribution across broad segments of the population can be addressed.

28 Recent declines in the saving rate have occurred during a period of falling government deficits. If Ricardian equivalence were true, then this would reduce the private saving rate.
A second related and essential question concerns how individuals are likely to respond to policy changes affecting saving. Lawmakers want to know if the policies they are considering are likely to increase the overall saving accumulated at retirement. This means that the modeling effort should be able to handle behavioral change in a consistent manner and in accordance with the theory.

2. What Definition of Saving Should We Use?

For the purpose of measuring household well-being, a definition of saving that reflects the accumulation of all household assets available to finance future consumption at retirement is needed. In principle, the measure should include all net additions to the stock of wealth including inflation-adjusted capital gains and losses on existing assets and the accrued value of pensions and Social Security. This suggests that a saving concept that is somewhat broader than the NIPA measure of private saving is required. The NIPA measure is meant to reflect only income and saving from current production, not changes in the market valuation of existing assets. Therefore, it includes capital gains of households only to the extent that they arise from retained earnings of corporations. To the extent that capital gains reflect improved expectations of future profits or increases in the value of household assets and business assets outside of the corporate sector, they are not counted in NIPA. In addition, NIPA does not include increases in accrued value of Social Security assets in household saving (or the associated increase in public sector liabilities as negative government saving.)

This definition of saving, while appropriate for measuring accumulation of real and financial resources at retirement, could be problematic when considered in the context of alignment with a macroeconomic model. This is because most macro models rely on NIPA measures of income and saving. In addition, we lack complete measures of household wealth changes. While the Federal Reserve Board publishes data on changes in total household wealth, the measures of housing wealth and assets in non-corporate business assets reflect changes in total investment less estimated depreciation. This measure is not necessarily equal to the change in the market valuation of those assets.

3. Stylized Facts With Respect to Saving

The last decade has seen an explosion in research on saving and retirement issues. No doubt, much of this relates to long-run policy concerns as baby-boomers approach retirement age. The availability of data sets such as HRS and AHEAD that allow for the testing and validation of richer models of individual behavior have enabled a new generation of studies. Also, new methodological approaches (e.g., stochastic dynamic programming) and the greater use of simulation methods have provided greater insight into the consequences and predictions of various theories of household saving and retirement decisions.

This section summarizes a number of empirical “facts” related to saving and retirement that have served as the impetus for much of the research in the last decade. Efforts to explain these phenomena in a unified and consistent theoretical setting have been a central focus of this research program. A brief review of these facts will serve two purposes. First, various theories of
saving and retirement can be evaluated in light of how predictions from the theory hold up to empirical scrutiny. And second, models used to project future retirement and saving decisions should, at a minimum, be consistent with observed data.

We summarize below five broad observations from historical and cross-section data on saving in the United States:

- The U.S. private saving rate has declined over the past four decades.
- Saving rates have declined in all age groups.
- High-income households on average save a much larger percentage of their income than low-income households.
- Many households reach retirement with little savings.
- Social Security and Pension wealth constitute most of assets of retirees.
- Older households eventually spend down their wealth, but at a rate much slower than would be predicted by a pure life cycle model.

*The U.S. Private Saving Rate Has Declined*

Table 3-1 documents the decline. Two aspects of this decline are worthy of mention. First, the most pronounced drop in the private saving rate began in the mid-1980’s, during a period of strong economic growth and relatively high real rates of return to saving. Second, this decline has continued throughout the recent economic boom.29

<table>
<thead>
<tr>
<th>Year</th>
<th>Private Saving Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960-1969 Average</td>
<td>8.9</td>
</tr>
<tr>
<td>1970-1979 Average</td>
<td>8.6</td>
</tr>
<tr>
<td>1980-1989 Average</td>
<td>7.1</td>
</tr>
<tr>
<td>1990-1994 Average</td>
<td>5.9</td>
</tr>
<tr>
<td>1995</td>
<td>6.2</td>
</tr>
<tr>
<td>1996</td>
<td>5.6</td>
</tr>
<tr>
<td>1997</td>
<td>5.2</td>
</tr>
<tr>
<td>1998</td>
<td>4.0</td>
</tr>
</tbody>
</table>

*Source: Gale and Sabelhaus (1999)*

---

29 Estimates of private saving do not include increases in the value of assets, so that capital gains on corporate equities, other than those attributable to corporate retained earnings, do not constitute saving under this definition.
Saving Rates Vary By Age, But Have Declined for All Groups

The decline in saving appears to have occurred across all age cohorts. Table 3-2, from Bosworth, Burtless and Sabelhaus (1991) shows net saving rates by age derived from the Consumer Expenditure Survey (CEX). (The concept of saving in the CEX is closer to personal saving than the NIPA concept of private saving because it does not include corporate saving.)

Table 3-2
Saving Rates by Age

<table>
<thead>
<tr>
<th>Age</th>
<th>25-34</th>
<th>35-44</th>
<th>45-54</th>
<th>55-64</th>
<th>65+</th>
</tr>
</thead>
<tbody>
<tr>
<td>1973/73 CES</td>
<td>9.5</td>
<td>12.1</td>
<td>16.8</td>
<td>22.9</td>
<td>14.9</td>
</tr>
<tr>
<td>1982/83 CES</td>
<td>9.6</td>
<td>8.6</td>
<td>10.5</td>
<td>15.8</td>
<td>11.5</td>
</tr>
</tbody>
</table>

Source: Bosworth, Burtless and Sabelhaus (1991)

Saving Rates are Higher for Higher Income Households

Saving rates are high for households with higher income. Table 3-3, also from Bosworth, Burtless and Sabelhaus (1991), compares estimates of saving rates across two different household surveys. The estimates of saving rates by income group differ dramatically between the two surveys, but both surveys show that the saving rate rises with income. Because the highest quintile receives such a large proportion of income, the figures suggest that most personal saving in the United States is done by a relatively small number of households. The figures also indicate that, on average, personal saving is negative for households in the bottom 40 percent of the income distribution.

Table 3-3
Saving Rates by Income Quintile

<table>
<thead>
<tr>
<th>Income Quintile</th>
<th>Saving Rate 1983-85 SCF</th>
<th>Saving Rate 1982-85 CEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>-2.4</td>
<td>-92.1</td>
</tr>
<tr>
<td>Second</td>
<td>-2.9</td>
<td>-10.3</td>
</tr>
<tr>
<td>Third</td>
<td>10.0</td>
<td>8.7</td>
</tr>
<tr>
<td>Fourth</td>
<td>9.9</td>
<td>16.7</td>
</tr>
<tr>
<td>Fifth</td>
<td>12.5</td>
<td>25.8</td>
</tr>
<tr>
<td>Total</td>
<td>9.5</td>
<td>10.8</td>
</tr>
</tbody>
</table>

Source: Bosworth, Burtless and Sabelhaus (1991)
Many Individuals Reach Retirement with Little Savings

Most individuals reach retirement age with little or no accumulated wealth outside of Social Security. Table 3-4 is from Poterba (1996) and shows median wealth holdings for those approaching retirement and those in the early stages of retirement. What is striking about these figures is that they indicate that at least half of households approaching retirement have no assets in targeted retirement accounts (e.g., IRAs and 401(k)s) and very little in other retirement saving. However, the figures do not include pension and Social Security wealth. Moore and Mitchell (1997) compare wealth holdings from the Health and Retirement Study (HRS), which allows for the calculation of Social Security and pension wealth. Their figures indicate that mean wealth holdings in the median 10 percent\(^{30}\) is about $325,000 when these two sources of retirement saving are included.

<table>
<thead>
<tr>
<th>Type of Wealth</th>
<th>Age 55-64</th>
<th>Age 65-69</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Financial Assets</td>
<td>8.3</td>
<td>14.0</td>
</tr>
<tr>
<td>Targeted Retirement</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Other Financial Assets</td>
<td>3.0</td>
<td>7.4</td>
</tr>
<tr>
<td>Home Equity</td>
<td>36.0</td>
<td>50.0</td>
</tr>
<tr>
<td>Equity in Other Property</td>
<td>8.2</td>
<td>6.0</td>
</tr>
</tbody>
</table>

Source: Poterba (1996)

Social Security and Pension Assets Constitute Most of Retiree Wealth

Financial assets alone comprise only a small fraction of total retiree wealth, especially among all but the wealthiest retirees. For most individuals, Social Security and pension wealth make up the bulk of retirement saving. Table 3-5 is from Moore and Mitchell (1997) and shows a breakdown of wealth, by wealth decile as reported in the Health and Retirement Study (HRS).

The Elderly Spend-Down Only Some of Their Wealth

One prediction of life-cycle models of wealth accumulation is that the elderly spend down their wealth in retirement. Early empirical work found that this did not appear to be the case and called into question the validity of the models. Subsequent research has tended to confirm that most aged households do begin to dissave at some point, but that this may not begin immediately upon retirement. Table 3-6 from Hurd (1990) shows how relative bequeathable wealth declines with age across four different surveys.

\(^{30}\) The median 10 percent is calculated as the mean wealth holdings in the 45\(^{th}\) to 55\(^{th}\) percentile.
Table 3-5
Fraction of Total Wealth, by Type of Asset and Wealth Decile

<table>
<thead>
<tr>
<th>Wealth Decile</th>
<th>Net Housing Wealth</th>
<th>Net Financial Wealth</th>
<th>Social Security Wealth</th>
<th>Pension Wealth</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-14%</td>
<td>4%</td>
<td>107%</td>
<td>3%</td>
</tr>
<tr>
<td>2</td>
<td>11%</td>
<td>11%</td>
<td>71%</td>
<td>7%</td>
</tr>
<tr>
<td>3</td>
<td>16%</td>
<td>12%</td>
<td>60%</td>
<td>12%</td>
</tr>
<tr>
<td>4</td>
<td>17%</td>
<td>15%</td>
<td>52%</td>
<td>16%</td>
</tr>
<tr>
<td>5</td>
<td>19%</td>
<td>19%</td>
<td>45%</td>
<td>18%</td>
</tr>
<tr>
<td>6</td>
<td>19%</td>
<td>21%</td>
<td>37%</td>
<td>23%</td>
</tr>
<tr>
<td>7</td>
<td>18%</td>
<td>24%</td>
<td>31%</td>
<td>27%</td>
</tr>
<tr>
<td>8</td>
<td>16%</td>
<td>27%</td>
<td>25%</td>
<td>32%</td>
</tr>
<tr>
<td>9</td>
<td>14%</td>
<td>33%</td>
<td>20%</td>
<td>33%</td>
</tr>
<tr>
<td>10</td>
<td>10%</td>
<td>58%</td>
<td>9%</td>
<td>22%</td>
</tr>
<tr>
<td>Median</td>
<td>14%</td>
<td>37%</td>
<td>25%</td>
<td>24%</td>
</tr>
</tbody>
</table>


Table 3-6
Relative Bequeathable Wealth by Age

<table>
<thead>
<tr>
<th>Age</th>
<th>1962 SCF</th>
<th>1979 ISDP</th>
<th>1983 SCF</th>
<th>1984 SIPP</th>
</tr>
</thead>
<tbody>
<tr>
<td>55-64</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>65-69</td>
<td>1.09</td>
<td>.85</td>
<td>1.27</td>
<td>.96</td>
</tr>
<tr>
<td>70-74</td>
<td>.96</td>
<td>.81</td>
<td>.84</td>
<td>.79</td>
</tr>
<tr>
<td>75-79</td>
<td>.89</td>
<td>.62*</td>
<td>.69</td>
<td>.69*</td>
</tr>
<tr>
<td>80+</td>
<td>.67</td>
<td>.62*</td>
<td>.52</td>
<td>.69*</td>
</tr>
</tbody>
</table>

Source: Hurd (1990) *Age 75 and over

Note: Bequeathable wealth includes financial assets, housing equity and the value of businesses, property and durable goods.

Interpreting cross-section data on age-wealth profiles needs to be done with care because, especially for older households, it is difficult to untangle life-cycle effects from differential mortality factors. Put differently, the perceived flatness of the age-wealth profile in old age is due partly to the fact that wealthier individuals tend to live longer, thereby increasing average wealth at greater ages.
III. OVERVIEW OF THE SAVING AND CONSUMPTION LITERATURE

The literature on saving and consumption is vast, complex, technically demanding and has been growing rapidly in recent years. In a comprehensive survey article, Browning and Lusardi (1996) cite over 150 articles related to household saving and over three-quarters were published in the past ten years. Since the influential work of Hall (1978), much of the empirical work on saving has followed two different but complementary tracks. One strand of research has focused on testing predictions from the utility-based, life-cycle model outlined in Hall's paper. Hall's finding that changes in annual consumption are independent of changes in annual income (i.e., consumption changes follow a "random walk") has been the subject of much scrutiny. Widely cited research that challenges this finding is Campbell and Mankiw (1989;1990).

A second strand of research has relied on simulation methods to investigate how enhancements to the simple, stripped-down life cycle model alter its predictions. Simulation methods are necessary because once the simple model is modified in certain ways, analytic solutions are no longer available and numerical methods must be relied upon. For example, how uncertainty over future earnings enters into the basic model can have important consequences for predicting saving behavior. When the assumption of quadratic utility is dropped and more general utility functions are allowed, this uncertainty translates into a precautionary saving motive. Without uncertainty, the life cycle model implies that younger individuals will borrow against future income to finance current consumption. But uncertainty over future labor income and the finite probability of negative wealth result in deviations from the standard, life-cycle path. The work of Carroll (1992), Carroll and Samwick (1994) and Deaton (1992) is in this vein.

When individuals face liquidity constraints, they are unable to smooth consumption over time and can exhibit behavior that resembles precautionary saving. Despite the large cost of increased complexity, simple enhancements to the life-cycle model result in a much richer depiction of individual saving behavior. The treatment of precautionary saving and liquidity constraints are perhaps the two most important innovations in the saving literature in the past ten years. Models of saving behavior that are enhanced in this manner are better able to explain important empirical "facts" about individual wealth accumulation. In particular, the models help explain why consumption tracks income quite closely over the life-cycle.

We focus on the theoretical aspects of saving to help guide the modeling. The life-cycle theory remains the central organizing principle because it captures a primary reason why individuals save -- a forward-looking desire to prepare for retirement. The recent theoretical developments noted above suggest how modifications to the simple model are likely to alter predictions from the pure life cycle model.

1. Life-Cycle Models

The starting point of most investigations of saving behavior is the Life-Cycle Hypothesis (LCH) of Modigliani and Brumberg (1954). Closely related to the LCH is the permanent income
theory of saving associated with Friedman (1955). Under the LCH, individuals save when they are young to finance consumption when they retire. A common, but by no means necessary, feature of the LCH is the familiar “hump-shaped” relationship between wealth and age. If individuals are assumed to maximize lifetime utility, straightforward application of dynamic optimization principles yield at least three predictions of the standard LCH model. First, individuals choose consumption and saving paths to keep the marginal utility of consumption constant over time. Second, saving and borrowing are the mechanisms through which this “smoothing” takes place. Third, because individuals attempt to smooth consumption, annual consumption is uncorrelated with predictable changes in annual income. The assumption of perfect markets is central to the predictions of the LCH and simple versions of the model assume no uncertainty.

A number of studies have attempted to test the validity of predictions from the standard LCH. Kotlikoff and Summers (1981), Ando and Kennickell (1987) and Carroll and Summers (1991) each found that individual age-wealth profiles did not exhibit the “hump-saving” features associated with the theory. Using aggregate data on consumption and saving, Campbell and Mankiw (1989; 1990) rejected the notion that consumption changes are uncorrelated with changes in income. Using micro data, Shea (1995) came to a similar conclusion: consumption and income appear to track one another quite closely. These findings have led researchers to conclude that the standard LCH needs to be modified in important ways.

The simple LCH results in strong predictions about lifetime wealth accumulation and has implications for modeling saving behavior. For example, if individuals are forward-looking in planning to save for retirement, then changes in future Social Security benefits or taxes will result in immediate changes in saving. In contrast, temporary tax cuts would have little or no effect on saving.

Another issue concerns substitution between different types of saving. The standard LCH suggests that different forms of saving are fungible and, therefore, perfect substitutes for one another. An implication is that if one form of saving is increased, it will be offset by reduced saving in other forms. This point is made by Feldstein (1974), who estimated that the Social Security system reduces private saving. This is consistent with the LCH because Social Security wealth replaces other forms of wealth accumulation. Barro (1978) showed, however, that Feldstein’s results were quite sensitive to the econometric specification used and how Social Security wealth was calculated. Under certain specifications, Social Security was shown to have a positive effect on saving. Barro’s results are consistent with the theory of Ricardian equivalence, which holds that individuals will alter their saving to offset government efforts to redistribute wealth between cohorts through increased government debt. By this theory, people will view the unfunded liability of a pay-as-you-go Social Security system as a reduction in private wealth, which will offset the increase in the present value of Social Security benefits.

A recent review by the Congressional Budget Office (CBO, 1998) summarizes empirical studies on the relationship between private saving and Social Security wealth. After analyzing a

---

31 The two models are nearly equivalent if permanent income is defined to be the annuity value of lifetime resources.
number of studies, it appears that the “center of gravity” of accepted research suggests this Social Security offset is between zero and 50 percent, with a mid-point estimate of 25 percent. Additionally, the size of this offset should also vary by cohort in order to capture differences in "money's worth" calculations. This will become important when assessing the impact of, for example, a private account mechanism as part of the Social Security system.

Another topic of much recent research has been the effect on private saving of targeted, tax-advantaged saving vehicles such as Individual Retirement Accounts (IRAs) and 401(k) pension arrangements. If individuals follow the pure life cycle model, then these arrangements should have no effect on lifetime wealth accumulation. Different studies have come to different conclusions on this issue, with estimates of the percentage of contributions to these accounts coming from other saving varying from close to zero to close to 100 percent. It is reasonable to conclude that these accounts may have some net positive effect on saving, especially for lower income and less wealthy individuals who may be liquidity constrained. But it would make sense to develop models that allow the user to experiment with a variety of parameters. Additionally, research suggests that the impact on saving behavior of these accounts is likely to be greater in the long run than in the short run.

Two features of the simple LCH that have been criticized as being unrealistic are assumptions of perfect markets and the absence of uncertainty. Much of the theoretical work in the last decade has explored the implications of relaxing these assumptions. In these models, uncertainty over future earnings and consumption result in predictions that can depart dramatically from the standard LCH and help explain why some individuals may choose to keep a “buffer-stock” amount of saving to draw on during times of unexpected shocks to earnings.

The manner in which these modifications to the standard LCH are implemented will likely have effects on how individuals are assumed to respond to various proposals related to Social Security reform. For example, if liquidity constraints are binding for certain population sub-groups, then setting up a system of mandatory private accounts is likely to have quite different effects on different segments of the population (Geanakoplos, Mitchell and Zeldes, 1998). It may be that wealthy individuals, who are already saving adequately for retirement and have no such constraints, treat these accounts as equivalent to other forms of saving and fully offset existing saving. But for those without access to credit markets, who are already consuming all or most of their income, such a system could result in new saving.

How our tax and transfer system interacts with individual saving decisions is likely to be an important modeling concern. Recent research suggests that our tax and transfer system can have important implications for saving and retirement decisions among certain segments of the population. The Hubbard, Skinner and Zeldes (1994) model indicates that not saving at all can be optimal behavior for many low income individuals. This is so because many transfer programs (e.g., food stamps and Medicaid) examine assets to determine program eligibility and, as such, impose a very high tax on saving. For an individual in poor health or whose income is uncertain, this creates an incentive to accumulate little or no assets.

---

2. Other Models of Saving

Rational models of saving for retirement implicitly assume that the individual will encounter no personal difficulty in implementing her lifetime saving/consumption plan. Some economists contend that this assumption may be unrealistic because carrying out such a plan requires the individual to choose to postpone consumption. This requires self-control, the ability to delay immediate gratification in order to achieve a longer run objective. There are two different models that rationalize the behavior of individuals who have difficulty exercising such self-control.

**Hyperbolic Discounting**

The life-cycle model assumes that people act as if their discount rate remains constant. For example, if an individual is willing to exchange $1 today for $1.05 in one year (implying a 5 percent discount rate), she would also be willing to exchange $1 to be received five years from now for $1.05 to be received six years into the future (Lowenstein and Thaler (1989); Ainslie (1991)).

But empirical evidence suggests that people act as if their subjective discount rate rises “as the future draws closer” so that the discount rate is higher in the short-run than it is in the long run. For example, from the vantage point of today, a person’s subjective discount rate between consumption to be received in year 20 and year 21 may be fairly low – say, 5 percent – but from the vantage point of year 20, the discount rate between year 20 and year 21 may be much higher – say 20 percent. This situation creates a conflict between preferences in the present and in the future. Viewed from today, the individual may, for example, plan to postpone $1000 of consumption in year 20, in order to have more consumption in year 21 because $1 consumption in year 21 is discounted at a rate of 5% relative to $1 of year 20 consumption. But, when year 20 arrives, the individual will decide to save less than $1000 (postpone less than $1000 of consumption), because the subjective discount rate is now 20 percent instead of 5%.

Such behavior may be contrasted with that which would result when the individual’s discount rate between year 20 and 21 is assumed to be the same – e.g. 5% – whether the viewpoint is today, or year 20. In that case, there would be no conflict between a plan formed in the present, and the action needed to realize that plan twenty years in the future. Because the relationship between the short- and long-run discount rates can be represented by a hyperbolic function in which events τ periods away are discounted with a factor $(1 + \alpha \tau)^{-\gamma}$, the behavior just described is called *hyperbolic discounting*.

If individuals employ hyperbolic discount rates, they will form lifetime saving plans that are dynamically inconsistent. Today the individual will plan to save more in the future than she will actually save. Put differently, in year 20 the person will not be able to exercise the self-control needed to implement the saving plan made in year 0 because tastes will have changed. The result is that people will save less than they would like to. This prediction is consistent with some recent survey evidence in which individuals report that their actual saving rate falls short of their self-reported target saving rate by 10 percentage points (Bernheim (1994)).
Such behavior might seem to imply that “forced saving schemes” – such as requiring individuals to deposit funds into individual Social Security accounts would boost saving. However, a theoretical model of saving behavior with hyperbolic discounting suggests that this need not be the case (Laibson (1996). The reason is that individuals can, in theory, “undo” some or all of the saving that is “forced” in such accounts by lowering their saving elsewhere.

The same theoretical model demonstrates that there are public policies that can boost saving when discount rates are hyperbolic, but these take the form of penalizing excess consumption and/or subsiding the rate of return to saving. Social Security reforms that do not have these features would not be predicted to raise saving, according to this model.

**Behavioral Life-Cycle Model**

Shefrin and Thaler (1988) argue that self-control is apt to be a problem even if tastes do not change over time (e.g. the discount rate remains constant). They argue that the life cycle model should recognize that exercising self-control is costly in psychic terms and should include three elements normally excluded from economic analysis: internal conflict, temptation, and willpower.

Incorporating these elements into the traditional life-cycle model results in what Thaler and Shefrin describe as a “behavioral life-cycle model” (BLC model). The key feature of this model is that individuals would like to allocate lifetime resources in a manner consistent with the life-cycle model, but they lack the willpower required to delay the amount of current consumption needed to fulfill the lifetime consumption plan. The result is that individuals act as if they have two sets of coexisting yet inconsistent preferences: one concerned with the long-run, the other with the short-run. The upshot of this dual preference structure is that the act of saving creates an internal conflict between the part of the individual which is forward-looking (the planner), and that part which cares only about the present (the doer).

Thus, the individual will strive to find ways of saving for the future which achieve his or her long-run goal, but which minimize the psychic cost, arising from the internal conflict between the planner and the doer, of achieving that goal. This in turn requires the development of an appropriate psychological technology to guide household saving behavior.

A key element of such a technology is the creation of mental accounts that separate income and wealth into different categories. These accounts are distinguished from each other by the degree to which it is psychologically permissible to use the account to finance current consumption. For example, the individual might be assumed to separate his or her finances in such a way that current consumption could be financed from discretionary cash income without guilt, from discretionary saving with somewhat greater guilt, and from saving mentally earned for retirement with considerably greater guilt. The purpose of creating such mental categories would be to permit increments to wealth to be labeled or framed as either being available for consumption, or earmarked for saving, thereby reducing the temptation of the doer simply to consume any increment to wealth.
Once such mental accounts are established, Shefrin and Thaler argue that individuals are in a position to rely on a mix of internal and external rules of behavior that enable them to save. Internal rules consist of behavioral rules of thumb, such as self-imposed prohibitions or borrowing to finance current consumption, or treating income received in lump-sum form, such as bonuses, as being more like saving than like current income, and hence less available for financing current consumption.

In addition to internal rules, the individual as planner may take steps to limit the amount of discretionary current income available for current consumption spending. This may be achieved by self-imposed external constraints or rules that ensure that some portion of income is deposited into one or more of the wealth accounts rather than into the current income account. Examples of such devices include contributing a percentage of earnings to a pension plan, systematically over-withholding income tax payments, and retention of earnings by corporations.

The BLC model has the attractive feature of being consistent with several different types of observed behavior that the traditional neoclassical model has difficulty explaining. For example, to the extent that individuals rely on mental accounts, it is not surprising that individuals would treat pension wealth as different from non-pension wealth. This implies that individuals would not respond to receipt of an additional $1 of pension wealth by reducing other wealth by an offsetting amount, a result that is consistent with some research findings, but not with the traditional life-cycle model.

Unlike the case of hyperbolic discounting, the behavioral life-cycle model has more straightforward implications for modeling the effects of Social Security reforms.

- To the extent that private Social Security accounts function like private pensions, the BLC model implies that each additional $1 of wealth accumulated in such accounts will displace less than $1 of other saving, so that creating such accounts would raise overall private saving.

- The BLC model predicts that individuals will be more likely to substitute saving deposited in external accounts for other saving as income rises. This follows from assuming that the marginal utility of consumption falls as income rises, which means that the marginal utility of giving in to temptation falls as income rises. Exercising self-control thus becomes less of a problem for higher-income savers. Such individuals would thus have less need to resort to mental accounts as a means of achieving desired levels of saving, and hence, would be more willing to treat amounts deposited into different accounts as fungible.

- Although the BLC model predicts that individuals will act as if saving held in different accounts is not perfectly fungible, it does not make such a prediction about different forms of saving held within the same type of mental account. An implication is that people would be more likely to substitute saving held in one type of external account for saving held in another. This suggests that people who already have access to external saving devices, such as pension plans, and especially 401(k) plans, might be
apt to substitute "added saving" in a Social Security account for such saving. Thus, added saving in Social Security accounts might displace other saving that was held in other external accounts.

From a modeling perspective, these observations suggest that Social Security reforms that create private accounts would increase the saving of some types of households more than others. Saving placed in such accounts would be least likely to displace other saving (most likely to increase total saving) among low-income households who did not participate in 401(k) plans. Conversely, such accounts would be most likely to displace other saving (least likely to increase total saving) among higher income households who did participate in such plans. For this latter group of households it would seem plausible to assume that saving deposited into Social Security accounts would simply displace saving done in other forms.

III. APPROACHES TO MODELING SAVING BEHAVIOR

It is our view that the treatment of saving in models for Social Security analysis should exploit the basic insight of the LCH that individuals are forward-looking in their decisions about how much to save. But models should also recognize that individuals sometime are unable to follow through on their plans or use a high discount rate when evaluating future alternatives. The combination of liquidity constraints and a high discount rate, for example, can result in a strong (positive) correlation between income and consumption for some households. The LCH can serve as a framework for organizing a strategy for modeling saving behavior by isolating those aspects of the theory that are at odds with observed data. Recent attempts to modify the standard model by making it more realistic (e.g., uncertainty, liquidity constraints) result in a much richer set of behavioral predictions. Just how one incorporates these variations into a microsimulation model is likely to involve some experimentation.

When considering how one might project saving and wealth for individuals forward to some year in the future, it is helpful to consider the following identity for the evolution of wealth. Here s represents the base year of the projection; t is the year the projection is to be made; \( W_t \) is wealth in year t; \( S_t \) is net saving in year t; \( r \) is the rate of return on wealth (assumed constant here); and \( c \) indexes each cohort.

\[
W_{c,t} = W_{c,s} (1 + r)^{t-s} + \sum_j S_{c,t-j} (1 + r)^j
\]

The identity captures several important features related to projecting wealth: (i) some base year estimate of wealth is needed; (ii) net saving, which can be negative, is an important component; (iii) the rate of return on existing wealth matters; and (iv) cohort differences in existing wealth and net saving can be central.

Several issues related to saving behavior are obscured by this identity and need to be considered as one attempts to implement a modeling strategy. For example, the equation assumes that there is only one type of wealth. If one believes that the composition of wealth affects saving or the rate of return, then some form of disaggregation is needed. In practice, it probably makes sense to separate housing wealth from other financial wealth since housing wealth is more illiquid.
and far less likely to be used to finance present consumption. Also, Social Security and pension wealth are important determinants of resources at retirement and should be treated separately. The assumption of a constant rate of return, though unrealistic, is probably an acceptable simplification, and the model omits any form of portfolio choice. Most importantly, though, the basic identity shown above by itself leaves unspecified how one might obtain estimates of net saving ($S_{ct}$). We turn to this issue in the discussion below.

Whatever approach is chosen to model saving behavior will likely involve some very strong assumptions on the part of the analyst about the knowledge and expectations of individuals. Uncertainty over the future course of income, health and medical needs, prices and rates of return on investments can often be handled explicitly in the modeling framework. It is frequently the case, however, that analysts make these assumptions implicit in the course of formulating and simulating saving behavior. In either case, the modeler should state up-front what, if any, assumptions are being made and allow for scenarios that test the robustness of the approach.

1. Dynamic Programming (DP) Methods

Recently, researchers have gained important insights into predictions from the basic life-cycle model through stochastic dynamic programming (DP) methods. In the standard DP model, households are assumed to choose a consumption path \{C_1, C_2, \ldots, C_N\} so as to maximize a (time-separable) utility function:

$$
E \left[ \sum \beta^t u(C_t, Z_t) + \beta^{N+1} V_{N+1}(W_{N+1}) \right]
$$

subject to the intertemporal budget constraint:

$$
W_{t+1} = (1 + r)(W_t + Y_t - C_t).
$$

In (3.1), the summation runs from 1 to N, which is the assumed life span of each individual; "E" is the expectation operator and reflects the fact that uncertainty causes the future values of certain variables to be unknown; $Z_t$ represents a set of characteristics (e.g., family size); $\beta$ is the discount rate, assumed to be less than one; $r$ is the rate of return on financial assets; and $V()$ is the value function placed on terminal wealth ($W_{N+1}$) and is usually meant to capture a bequest motive. Equation (3.2) shows the evolution of wealth where $Y_t$ is labor income in year $t$. A useful re-characterization of (3.2) defines a new variable $X_t$, called "cash-on-hand", as the sum of $W_t$ and $Y_t$. Making this substitution into (3.2) yields an alternative form of the budget constraint:

$$
X_{t+1} = (1 + r)(X_t - C_t) + Y_{t+1}
$$

It is standard in the literature to assume that the utility function be of the Constant Relative Risk Aversion (CRRA) class and be separable in the $Z's$:

$$
u(C,Z) = v(Z) C^{1-\rho} (1-\rho)^{-1}
$$

We thank Olivia Mitchell for pointing this out to us.
In (3.4), $1/\rho$ represents the intertemporal elasticity of substitution. Depending on the way
uncertainty enters (3.1), it is usually the case that an analytic solution to the household
optimization problem does not exist and numerical methods must be employed (see, for example,
Judd, 1998, Chapter 12). A solution is comprised as a sequence of consumption values that
achieve the maximum of (3.1) and satisfy the budget constraint (3.2) (or (3.3)).

In is unlikely that DP methods will be feasible for use in microsimulation models in the
near future because their computational burden is so immense. Nevertheless, much can be learned
by examining simulation results from the standard setup. This was the approach taken by Deaton
(1992) who was able to characterize the relationship between consumption and cash-on-hand by
solving a model like (3.1). His results showed that consumption tracks income quite closely when
resources (cash-on-hand) remain low. In this region, the precautionary saving motive dominates.
At a certain level of cash-on-hand, determined by the parameters of the problem, life-cycle saving
takes over and the household begins to accumulate substantial wealth. Liquidity constraints are
not explicitly introduced in the model, but the combination of uncertainty over future earnings and
the CRRA utility framework results in behavior that mimics that of a liquidity-restrained
household.

A similar exercise was performed by Gourinchas and Parker (1999). In their model,
consumer tracks income up to about age 45 when saving for retirement begins. In both
models, life-cycle saving begins when individuals enter their peak earning years as uncertainty
over future earnings prevent the young from borrowing. The models also help explain why
individuals with low incomes that are flat over the life-cycle end up in retirement with little saving:
cash-on-hand never reaches the critical value when saving for retirement begins. Insights gained
from DP models can be used as a guide in specifying how consumption and saving decisions
might enter less structured models.

2. Estimation and Projection of Wealth at Retirement

It may be the case that one is only interested in a point-in-time estimate of how certain
cohorts are affected by Social Security reforms. For example, one may want to examine how the
after-tax incomes of future retirees are affected by certain proposals. Alternatively, some other
measure of economic well-being, such as the overall poverty rate among retirees might be of
interest. If this is the case, a parsimonious and straightforward approach to calculating wealth
would be to use a projection equation to forecast future wealth given assumptions about the likely
course of earnings and demographic change.

Such a method benefits from not having to predict the year-by-year saving decisions faced
by households but has several limitations. First, because the accumulation equation would
presumably be estimated from existing data sources, cohort differences in saving patterns would
be difficult to capture. Second, certain events that affect saving decisions over the life course,
such as the birth of children and their eventual departure from the household, would have to be
ignored because such a model would not adequately handle changes in a household's status over
time. Third, it is difficult to see how a behavioral change to alternative policies could be
incorporated within the structure of the model, although off-model adjustments could be super-
imposed on model projections to simulate a policy change. Nevertheless, a simple equation for
projecting wealth at retirement is a good first approximation for estimating the financial resources
of future retirees.

A strategy would be to estimate an equation of the form:

$$W_{it} = X_{it} \beta + \mu_i + \gamma_t + \varepsilon_{it}$$

(3.5)

where $W_{it}$ represents wealth in year $t$ for household $i$; $X_{it}$ is a vector of explanatory variables; $\beta$ is
a vector of coefficients; $\mu_i$ represents an individual specific effect; $\gamma_t$ captures the time-specific
effect; and $\varepsilon_{it}$ is the usual residual term. An equation like (3.5) can be estimated from existing data
sets such as the PSID. Alternatively, cross-sectional estimates of annual wealth could be estimated
from the SIPP or the Survey of Consumer Finances (SCF). A synthetic panel of several years of
data would be necessary to estimate both age and period-specific effects with either the SIPP or
SCF. Presumably, variables in the X-array would include age and age-squared to capture life-
cycle effects of wealth accumulation; some measure of permanent income; and demographic
variables indicating marital status, race, education, family size and the like. Variables indicating
pension coverage and health status would also be important predictors of wealth. (Missing from
the specification is a parameter to capture a cohort effect because of the perfect collinearity of
age, time and cohort.)

When using an equation such as (3.5) to project wealth, a key consideration relates to how
demographic changes are to be incorporated. For example, certain variables that can affect saving
such as family size, marriage and divorce, health status and age of children certainly belong in any
equation that attempts to project age-wealth profiles. Whether or not these variables will be
available for use in the projection equation will depend on the type of underlying model that is
used and level of demographic detail that is forecasted.

Because it is not clear, in this approach, how to predict changes in saving in response to
alternative policy scenarios, a necessary second step would be to augment the projections to
account for changes in economic variables and behavioral responses. This may require performing
off-model adjustments to reflect changes in certain underlying parameters. For example, if after-
tax rates of return on saving are expected to be different under the alternative policy, then one
could estimate how higher returns would affect the growth in wealth between the base year and
retirement. A simple adjustment to the wealth projections could be incorporated to reflect this.
Likewise, if changes in future Social Security benefits alter permanent income for some future
cohorts, then the projections can be modified to reflect how lifetime saving behavior might
respond.

---

34 We have in mind here an equation for non-pension, financial wealth. Presumably, a similar equation could be used for
housing wealth. It is likely that a separate strategy for estimating Social Security and pension wealth would be necessary.
Social Security wealth depends on lifetime earnings and can be calculated reasonably well once earnings histories are
available. Projections of pension wealth should be tied closely to estimates of the changing structure of private pension
arrangements.
3. Projection of Wealth Based on Econometric Estimates of Annual Consumption

An alternative approach to a direct projection of retirement wealth is to build up wealth accumulation year-by-year by modeling the saving decision directly. This approach fits naturally into dynamic microsimulation models because these models simulate the behavior of individuals through time according to a set of descriptive equations, accounting identities and behavioral relations. Existing microsimulation models are discussed in Chapter 1 of this report.

Many existing dynamic microsimulation models do not estimate saving directly, but use imputation methods to predict wealth in a particular year. The predictive equations used in the models are generally reduced form models, estimated from some external data source such as the Survey of Consumer Finances (SCF) and relate wealth in year t ($W_t$) to a set of income and demographic variables. Such an approach was outlined in the previous section. These estimates of wealth and saving are only loosely tied to the individual life histories through the prediction equation. Of course, no structural relations are estimated.

An alternative approach would impose simple relationships on the underlying household data and estimate saving and consumption directly as part of a microsimulation model. Even if this falls short of a utility-based, dynamic structural model of consumption/saving, the approach outlined below is a first step in getting there. It also allows the modeler to make explicit choices about some of the micro behavioral issues discussed in this chapter, and opens the door for more sophisticated treatment of individual behavior.

Such an approach would track wealth accumulation by looking at annual changes in net worth in any year as a function of income, consumption and the return on saving. Wealth would evolve according to the equation:

$$\Delta W_t = r * W_{t-1} + Y_t + \text{Pen}_t + \text{Social Security}_t - C_t - T_t + B_t \quad (3.6)$$

This accounting identity says that the change in (financial) wealth between period t-1 and t is equal to income from all sources less consumption and taxes paid plus bequests. The first four terms on the right-hand side represent income from existing assets, earnings, pension distributions and Social Security, respectively; $C_t$ is consumption; $T$ represents taxes (income, payroll and other taxes); and $B$ is bequests. With the exception of $C_t$ and $B_t$, the other variables are available in one form or another in most existing models.\(^{35}\)

The approach has several advantages. First, consistency between wealth and important income variables is forced through the accounting identity. Second, estimates of saving are related to other important policy parameters through the tax function and rules for computing Social Security benefits. Third, macroeconomic alignment is facilitated because aggregate consumption can be summed over households and compared with macroeconomic projections. Finally, a

\(^{35}\)A more comprehensive approach could attempt to model the change in pension and Social Security wealth directly rather than the income flows from these assets.
framework is in place for more elaborate modeling of the interaction among the income components.

A critical factor in the above approach is obtaining an appropriate model for projecting $C_t$ and much of this chapter has emphasized how difficult this can be. (Most existing dynamic microsimulation models can, in principle at least, track intergenerational transfers in a straightforward way.) Below we outline three approaches to estimating the consumption/saving patterns of households that we believe could be implemented in this framework with little additional work.

**Consumption a Fixed Fraction of Current Income**

Perhaps the simplest way to derive estimates of annual expenditures is to assume that consumption is a constant fraction of annual income. To be consistent with observed data on consumer expenditures, it would make sense to allow this fraction to vary by income level. A further important refinement could allow this fraction to have a stochastic component that would be allowed to vary over time and across individuals. An advantage of this approach is that, by allowing the average propensity to consume (APC) to vary by income class, one is more likely to match cross-sectional patterns in consumption data. Among the main disadvantages of this approach is that there is no forward-looking component of the consumption/saving decision. Nevertheless, this approach takes an important step linking current saving decisions with future wealth accumulation.

This approach could be implemented by estimating $C_t$ as a function of income and demographic variables, using information from the Consumer Expenditure Survey (CEX). Income variables would include both earnings and income from investments. Demographic variables would include family size, age of the household head, ages of any children, indicators of health status (if available) and marital status. A stochastic, zero-mean residual could be added to predicted consumption to allow for variance in actual annual consumption spending across households with similar characteristics. A major drawback of this approach is that we probably have less reliable data on consumption than we do on wealth. It is likely that there is some measurement error in the CEX with respect to income or consumption (or both) making it difficult to obtain reliable estimates of saving. Nevertheless, the CEX is our only direct source of cross-section, household consumption data.

A possible refinement would be to estimate separate equations for consumption of durable goods and non-durable goods and services. One approach would use a Tobit regression procedure to estimate “lumpy” purchases such as automobiles and consumer durables. Spending on education might also be treated in this manner. In the next stage of the estimation, total spending on non-durable purchases would be estimated as a linear function of a similar set of predictors. It might also be useful to perform separate sets of estimates for different groups of consumers – homeowners and renters, for example. During the simulation, each year of consumption spending would be calculated based on existing earnings (including pensions and income form assets), age and other demographic variables.
This procedure has the advantage of tying household spending more closely to annual income and household demographic characteristics. Including age as a predictor would capture life cycle patterns in consumption/saving profiles. But without reliable information on future Social Security or pension benefits, it would be difficult to capture behavioral changes in this framework.

**Consumption as a Function of Permanent Income**

An approach closely linked to the theory would be to allow annual spending to be a function of permanent income. Alternatively, one could define permanent income as the annuitized stream of lifetime resources and set annual consumption equal to this amount. This is, of course, the basic result of the LCH and would be a first step in allowing the model to capture intertemporal tradeoffs. Presumably, permanent income would be equal to the present discounted value of lifetime earnings and future pension and Social Security benefits. Because this structure would be forward-looking, the components of permanent income would be forecasts.

As we discussed earlier, two significant criticisms of the LCH involve the absence of uncertainty and the assumption of perfect markets. Each of these could be added to the above framework in a straightforward manner. Uncertainty could be embedded into the earnings forecasts, allowing for the calculation of a transitory component of income. This was the approach taken by Paxon (1992) who included a measure of income uncertainty into her model and separated the saving decision into two components: saving out of permanent income and transitory income. Alternatively, an estimate of the variance of (lifetime) income could be incorporated directly into the consumption function. The assumption of perfect markets could be relaxed by allowing liquidity constraints to be binding for certain individuals. If these individuals lack access to credit markets, for example, then they could not borrow to smooth out consumption over their lifetime, causing income to track consumption very closely.

We believe this approach has significant advantages over other methods for estimating annual consumption spending. First, it is forward-looking and captures the essence of the LCH. Second, it links consumption – through permanent income – to important policy parameters such as tax rates and future Social Security benefits. This will allow, for example, induced changes in saving to occur as estimates of permanent income are altered. Third, uncertainty over the future course of income (and consumption) can be explicitly modeled.

A drawback of this approach is that good estimates of permanent income are difficult to obtain. Moreover, it is unlikely that the necessary parameters can be estimated from cross-section regressions using existing household-level data sets. Permanent income could be imputed by using regression methods to project future earnings and then estimating the value of Social Security wealth from the earnings projections. Presumably, the future value of pension wealth could be estimated in a similar manner. When a policy change is introduced, one could then calculate how the change would affect the estimates of permanent income. The consumption relationship would have to be estimated from macro time series data. This would mean that additional assumptions would be necessary to represent any diversity in consumption behavior among households.
**Consumption as a Function of Cash-on-Hand**

Another approach would calculate annual consumption as a function of cash-on-hand, defined as the sum of existing wealth and current-year earnings. This is similar to the approach taken by Deaton (1992, Chapter 6), who calculated optimal consumption paths in a utility-based, DP framework. Deaton’s results showed that under reasonable choices for the parameters of the problem, individuals consume most of their cash-in-hand, allowing for a small “buffer stock” of saving, until cash-in-hand reaches a certain threshold. Once the threshold is reached, usually when the growth in earnings begin to steepen, life cycle saving begins.

Deaton found that a simple function relating consumption to cash-on-hand fit his data quite well. A simple rule that says the marginal propensity to consume out of cash-on-hand is close to 1.0 for values below the threshold and something less than 1.0 (a value of .3 was chosen by Deaton) for any amount in excess of the threshold yielded a close approximation to the DP solution.

### 4. Issues That Models of Wealth Accumulation Must Address

**Quality of Data**

Unless one is willing to begin a simulation many years in the past, some estimate of existing wealth is necessary for initialization and alignment of the base year model data. If a data set such as the SIPP is used, then information on the Topical Modules can provide a good starting point. Urban Institute tabulations in a recent study (Urban Institute, 1999) suggest that wealth data on the SIPP compares quite favorably with data from other sources such as the PSID or the SCF. Differences in SIPP wealth distributions are only apparent above the 95th percentile and are due to the fact that the SIPP has fewer very rich households in the sample than the SCF, which over-samples wealthy households. This should not be a problem for projecting the retirement wealth of most people, including all those who depend on the Social Security program as a major income source. But adjustment of the underlying SIPP wealth data to derive better wealth estimates at the top will be necessary if the data are to be aligned with macro estimates of wealth because the top one percent of wealth holders account for about a third of all the assets. One approach would be to utilize recent HRS/AHEAD surveys to improve upon wealth estimates, especially at the top of the wealth distribution.

The quality of micro data on saving rates is considerably worse than wealth data. A related issue is that income, generally, is under-reported in the CEX. This means that saving propensities are understated when saving is measured as the difference between annual (disposable) income and expenditures. The discrepancy is most noticeable in the lowest income groups, where reported consumption is two to three times reported income.

**Treatment of Uncertainty**

How a model of saving behavior deals with uncertainty can have significant implications for the predictions of the model. At least four types of uncertainty come to mind:
• Uncertainty with regard to future *earnings*.
• Uncertainty with regard to the rate of *return* on saving.
• Uncertainty with regard to *health* status.
• Uncertainty with regard to the time of *death*.

The way that people react to these uncertainties can have huge effects on their propensity to save at various times during their lifetime. (How best to incorporate these types of uncertainty is discussed in Chapter 5 of this report.) A related issue for modeling saving behavior relates to how individuals perceive and/or respond to increased uncertainty with respect to permanent income. Some models suggest that people save more when their income becomes uncertain. This means, for example, that increased uncertainty over future Social Security benefits could increase saving.

### 5. Saving and Wealth Accumulation in Existing Simulation Models

This section focuses on how various models potentially appropriate for analysis of Social Security reform proposals currently treat saving and wealth accumulation and the retirement decision. We review a number of representative models.

All models must address two issues -- the unit of analysis and the data source(s) for estimating behavioral responses. The unit of analysis for models of saving needs to be a household rather than an individual. Household saving is generally measured as the difference between after-tax family income and family expenditures. As we have already discussed, good longitudinal expenditure data are rare. Data on stocks of wealth, though also incomplete, are more prevalent, allowing some researchers to model saving as the difference between wealth stocks in two periods less appreciation or depreciation of wealth over the interval. These two approaches--subtracting consumption expenditures from after-tax income and subtracting depreciation from wealth change--should in theory yield the same answer about household saving. Discrepancies are likely to arise, however, due to measurement error or definitional and conceptual differences when constructing different series.

To contend with limited data on consumption, saving, and wealth, researchers often combine cross-sectional data on individuals from the same cohort at different points in time rather than following actual individuals over time. This technique, known as a synthetic cohort approach, allows one to consider life-cycle, or longitudinal, patterns in these processes without having panel data. Using a synthetic cohort approach, however, one is more constrained in potential model specification than one is when using panel data. For example, employing panel data, one can incorporate elaborate lag structures into a model of a process and can consider change variables, whereas with synthetic cohort data one is typically constrained to using state and level variables because observations on individuals are only available for one time point.

Retirement, in contrast, needs to be treated as an individual-level rather than household event in these models. Characteristics of others in a household may influence an individual’s retirement choices, but individuals in the same household need not retire simultaneously.
The dozen or so models that the policy community has either used or evaluated as tools for analysis of Social Security reform differ considerably in their approaches to representing saving and wealth accumulation and the retirement decision. Each approach has strengths and weaknesses.

**MINT**

The Model of Income in the Near Term (MINT), developed at the Social Security Administration by Howard Iams and Steven Sandell and subsequently expanded by RAND, the Urban Institute, and the Brookings Institution, does not contain an explicit model of saving. Rather, it projects non-pension wealth stocks, including both housing and non-housing wealth, at ages 62 and 67 using two stages (for additional detail, see Toder et al., 1999). The first stage is a random-effects probit model which determines whether a household has positive wealth. The second stage is a random effects generalized least squares model which predicts the level of wealth for households with positive wealth. Parameters in both stages are estimated from the PSID. Determinants of whether a family has positive wealth include the age, race, and marital status of the family head, family size, and average earnings in the recent past. Determinants of the level of wealth are more numerous, and include age and cohort terms, marital status, family size, earnings measures, health measures, wealth indicators, and many interaction terms.

After a head reaches age 62 or 67 (whichever best describes the time of first Social Security benefit receipt), MINT uses OLS regression to capture changes in the family’s other financial assets (i.e., the decay of non-housing wealth). The estimate uses a synthetic panel approach, based on SIPP wealth topical module data, rather than actual panel data. Determinants of change in wealth holdings include Social Security and pension income, past earnings, education, race, sex, cohort effects, life expectancy, marital status, and numerous age interaction terms.

MINT also projects pension wealth at ages 62 and 67 using a series of transition probabilities and replacement rates, which are based on SIPP and BLS data. Pension income and wealth are updated after age 67 based upon whether pensions include cost-of-living adjustments.

As Chapter 1 describes, these MINT projections are made sequentially on a lifetime (rather than annual) basis for each member of the sample. Non-pension wealth projections follow the projections of demographics, earnings, and pensions in the overall model solution sequence. This implies that, as MINT is currently structured, wealth changes never affect demographic events or earnings. Such a structure limits researchers' opportunities for integrating feedbacks from policy changes into the MINT wealth accumulation process. Further, MINT functions do not consider prices (i.e., if the price of housing increases relative to the price of other financial assets, sample members will still make the same choices about their portfolios). A user who wishes to impose saving/wealth changes in response to policy reforms would thus need to do so from the “top-down” in MINT. That is, ex post adjustment to the predicted wealth variables would have to be incorporated.
MINT does not currently contain an explicit model of retirement, though it does have a forecast of take-up of Social Security benefits and an elaborate representation of the earnings of Social Security beneficiaries. Members of the panel of outside advisors to the MINT study have recommended developing an explicit model of retirement as an important step in making the model more reliable. The next stage in the development of the MINT model will address this problem. (See Social Security Administration, 1999)

**CORSIM**

The CORSIM model, developed at Cornell University by Steven Caldwell and associates, has separate saving and wealth modules. CORSIM generates a prediction for annual household saving by taking the residual between after-tax family income and consumption. After-tax family income is the sum of earnings, transfer income and asset income less federal, state, and local taxes. Transfer and tax functions in the model are typically rule-based, with participation/take-up probabilities estimated from sources like the PSID, SIPP, and NLS-Y. The model's earnings functions are estimated from PSID data, and then aligned to subnational totals from the CPS and National Income and Product Account aggregates. Because the model does not project pension income, its estimates of after-tax income are likely to understate total income of the aged considerably.

The CORSIM consumption regressions were estimated from the CEX and are not aligned to aggregate totals. The outcomes in the CORSIM consumption module are based primarily on family composition (determinants include age, education of head and spouse, family income, family size, race, labor force effort, and a variety of interaction terms) and are not sensitive to prices, thus limiting the ways in which saving/consumption could be linked to Social Security parameters in a more behavioral framework.

Because the consumption module is, however, located within the annual processing loop in CORSIM, changes in saving in response to policy changes could to some extent be integrated from the bottom up. For example, if family after-tax income decreased in response to tax changes, family consumption and thus saving would automatically react. Also, labor supply and/or demographics can in theory react to wealth changes, though they do not at present.

CORSIM also models changes in net worth over three-year intervals. This module uses a wide array of data sources, including the Survey of Consumer Finances (SCF) and the Survey of the Financial Characteristics of Consumers for imputation and for parameters in microdynamic aging equations; estate tax and flow of funds data are used for alignment of the micro-level predictions. CORSIM captures fourteen components of wealth in its net worth measure, and it models each wealth type independently. Separate equations determine the probability of ownership of a given wealth type and then the value of the wealth type (conditioned on ownership). Both types of equations differ for prior owners and prior non-owners. Determinants of the probability of ownership include age, education, family size, income, marital status, and race, and determinants of wealth value include age, lagged value of the asset or debt, education, family size, income, marital status and race.
Modeling wealth in such a disaggregated way has both strengths and weaknesses. One benefit is that the composition of wealth closely tracks historical data by type of wealth. Nonetheless, this approach is potentially problematic for trying to model reforms like individual accounts, because users would like to consider changes to individuals' portfolio allocation strategies in the wake of such a reform. The current wealth model does not imply a strategy, given a lack of interdependence in modeling ownership decisions.

As with the CORSIM consumption module, the wealth module does not use prices, so people make the same choices regardless of changes in values of different wealth components. Further, the heavy reliance on alignment in this section of the model implies that users would need to integrate any feedbacks that might arise from policy in a combination bottom-up and top-down fashion. In order to force consistency between wealth and consumption decisions, CORSIM implicitly assumes that any differences are a result of asset revaluations.

Like MINT, CORSIM lacks an explicit model of retirement, but it does have a forecast of first receipt of Social Security benefits. Unlike MINT, CORSIM generates its earnings predictions using multiple stages: the decision to work, the number of hours worked, and then the wage rate. This is likely to facilitate integrating CORSIM with a more elaborate retirement model.

**DYNASIM**

The Urban Institute's DYNASIM currently relies on a cross-sectional imputation of family financial assets, both housing and non-housing, based on 1984 SCF data. For non-housing wealth, the model relies on a tobit specification, and for homeownership wealth, the model relies on a probit approach. These wealth imputations occur outside of the loop in which the model simulates earnings and family demographic events. This leaves limited opportunities for including feedback between policy changes and wealth accumulation. The model is currently undergoing extensive revision, including implementation of a new starting data base. Saving and wealth components will be more closely integrated with other processes in the model revision.

Like MINT, DYNASIM projects pensions in a detailed manner. This portion of the model will also be updated in the on-going revision.

As Chapter 1 notes, DYNASIM does have an explicit model of retirement. This model has an number of appealing features. It considers forward-looking behavior, in that members of the sample compare the value of retiring at a given point in time with the value of waiting an additional year, and choose the better deal.

**PRISM**

Researchers at the Lewin Group are currently updating the PRISM model to employ a more recent starting database. The older version of the PRISM model used the 1983 SIPP to predict both family assets and family homeownership. PRISM appears to treat wealth in a manner that is more similar to MINT's than to CORSIM's algorithms. It too focuses on wealth at a
particular point in time and then models change to this stock. Anderson reports that PRISM developers “assign asset levels at age 65 and then specify saving rates, dissaving rates, and maintenance of wealth” (Anderson, 1999). Presumably, PRISM’s developers are updating these components of the model in their overall conversion to a new data base.

Not surprisingly given its name, PRISM has an extremely elaborate pension model. This model uses actual pension plan data rather than stylized plan descriptions.

PRISM explicitly models both age of first receipt of Social Security benefits and age of take-up of pension benefits.

**DYNACAN**

While DYNACAN is arguably the most extensively validated micro-level social insurance model in the world, with predictions that line up with and feed into the Canadian government’s actuarial model (ACTUCAN), it does not have elaborate information on wealth and saving. Members of the DYNACAN team report plans to add a wealth module to the model, perhaps based loosely on the CORSIM module, but this decision is still in process.

**EBRI/SSASIM2**

EBRI/SSASIM2, developed by Martin Holmer and several associates, differs significantly from the microdynamically-oriented models just described. Unlike the preceding models, EBRI/SSASIM2 does not operate on a representative population sample, but instead uses composite workers (18 per cohort) for different cohorts. Age-specific saving rates are assumed for each cohort. The model does, however, have important short-run capability for considering certain macroeconomic feedbacks. EBRI/SSASIM2 relies on a neoclassical growth model, based on a Cobb-Douglas production function which accounts for labor-augmenting technological change (Holmer, 1997). Any changes to the OASDI surplus (expressed as a fraction of GDP) affects not only the national saving rate, but also tax revenue and other retirement saving. The model can simulate individual account proposals with an exogenously set offset rate for other saving. EBRI/SSASIM2 is very much a top-down model. Individuals paths are shaped by stochastically determined macroeconomic trends, but do not contribute to them (Holmer, 1999: 32).

Given its focus on simulating Social Security individual account proposals, the model does not have any implicit or explicit treatment of pension income and coverage. Retirement decisions are handled implicitly through the application of labor force participation rates (rather than explicitly through decision rules) and specification of age of first Social Security receipt.

**Auerbach-Kotlikoff Overlapping Generations Model**

In the Auerbach-Kotlikoff (A-K) model, agents in 55 overlapping generations are forward-looking, consider factor prices (and thus labor-capital ratios), and have perfect foresight. Saving is explicitly modeled as part of the intertemporal allocation of consumption by utility-
maximizing, representative households in each cohort. The model includes 12 separate income groups within each cohort, and the developers are currently working on integrating differential fertility by income group. Intergenerational wealth transmission is modeled through a bequest motive which also differs by lifetime income group. While the model currently starts from a steady state, it will soon allow users to start from an arbitrary state, with the distribution of wealth providing the critical state variables (Kotlikoff, 1998). SCF data will be used to provide values for these values. Developers are also working on integrating dynamic programming into individuals’ decision-making about consumption, labor supply, and bequests, and on taking into account the borrowing constraints individuals face. (Chapter 4 contains additional discussion of the use of the A-K model and its successors to analyze the macroeconomic effects of Social Security reforms.)

The model does not contain explicit treatment of employer-provided pensions, but instead considers them as part (via a scalar adjustment, estimated from data) of earnings-ability profiles. In effect, pensions have no effect on net saving in their model. Individuals maximize an inter-temporal utility function. Saving depends only on an individual’s consumption tastes, the rate of return and the time path of earnings. It does not depend on the composition of assets. Because labor supply is endogenous in the model, retirement decisions are closely related to saving decisions.

IV. SAVING AND OTHER BEHAVIORAL DECISIONS

1. Retirement Decision

It is unlikely that individuals make saving decisions in isolation from the decision of when to retire. This is particularly true in the years leading up to (anticipated) retirement as workers become more aware of their retirement income needs and the adequacy of existing saving. This endogeneity of saving and retirement decisions complicates the task of the modeler and has consequences for analyzing the effects of certain policy options. Raising the retirement age, for example, is likely to cause delays in retirement for those with little or no saving and a corresponding loss in their well-being.

Demographic projections indicate a drastic change in the structure of the US population. In short, Americans are growing older. People aged 65 or older are projected to represent 20 percent of the population in 2040, compared with slightly more than 13 percent in the year 2000. The median age of the US population --now 34.9-- is projected to increase to 38.3 in 2040. The aging of the population will affect labor market opportunities for both younger and older workers. The economy may have to rely more and more on the elderly as a source of labor, which will, in turn, affect retirement decisions and retirement incomes.

The existing literature on modeling Social Security and the retirement decision reveals a trade-off between realism and tractability. Many different factors affect the timing of retirement, and incorporating all of them into one model is impossible given the current technology. Models are usually of two types. Ad hoc econometric models (see Chapter 4) specify equations to
describe the determinants of labor supply, typically expressed as a function of the wage rate, family structure and health status, among other variables, instead of deriving this relationship from first principles. Utility-based models represent individuals' behavior as if they were maximizing a utility function. Ad hoc models have the advantage of tractability, and can incorporate many different factors into the analysis (for example health status, family composition, and work characteristics together) but have two shortcomings. First, they are not very useful to analyze the welfare consequences of a policy change. Second, the estimated relationship from such models may not be stable if the policy regime changes. (See Chapter 4 for a discussion of this point, which is referred to as the Lucas Critique.) Utility-based models do allow for welfare analysis, but are often excessively complex and need to impose strong assumptions on functional relationships to be computationally tractable.

**Estimates of the Retirement Decision**

The literature is divided on the importance of Social Security benefits in explaining retirement behavior. Two recent studies consider how factors like Medicare and private pensions can affect the retirement decision. Rust and Phelan (1997) present a dynamic programming model of the retirement decision faced by workers who are not presently covered by an employer-provided pension plan. In a similar modeling framework, Lumsdaine, Stock, and Wise (1992) examine the choices faced by a particular group of workers nearing retirement in one firm who were covered by an employer-sponsored plan. They use data from one firm so it is hard to make general statements about their conclusions. They also have information about employees' retirement or separation from that specific firm, but they do not know if the individuals have left the labor market or if they have just moved to a different job. They found that retirement decisions were influenced by incentives in the firm’s pension plan. Rust and Phelan, on the other hand, do not have adequate information about pension plans so they use only a sample of people without private pensions. Not surprisingly the effect of Social Security and Medicare in this case is much more important than in the Lumsdaine, Stock, and Wise study, which examined people with private pension coverage.

Some empirical evidence suggests that retiree health insurance affects the timing of retirement (Johnson, et.al. (2000); Gruber and Madrian (1995); Karoly and Rogowski (1994)). Health insurance should probably be incorporated in a model of retirement with Social Security. Private pensions are also documented to be significant in the retirement decision (Madrian(1993) and Madrian and Beaulieu (1998))

No model has dealt with saving and retirement simultaneously, and there are many instances in which a feedback from saving to retirement (and vice versa) is present. Rust and Phelan (1997) have uncertainty in their model and imperfect markets, but no saving decision. Stock and Wise (1990) assume an individual consumes all of his or her income in every period.

**Joint Retirement Decisions of Spouses**

Until recent years, retirement was generally viewed as a transition made by men and unmarried women. Most married women in their fifties and sixties were not working for pay, and
those who were working generally did not have extensive employment histories. For most elderly women, Social Security benefits were based on their husbands’ employment, not their own, and their pension wealth was insignificant. As a result, household retirement wealth generally depended much more on the timing of the husband’s retirement decision than on the timing of the wife’s decision. Consequently, much of the retirement literature has modeled the retirement decision from the point of the view of the individual worker, and most empirical work has focused on the behavior of men.

Among members of the post-World War II birth cohort, however, married women have entered the labor force in large numbers. As these women approach retirement age, many have long employment histories and have accumulated substantial Social Security and private pension wealth. Because the retirement decisions of wives will have important effects on the total retirement wealth of the household for many couples, it is unlikely that most married women in coming years will continue to follow the lead of their husbands when deciding the age at which to retire. Instead, it is more likely that increasing numbers of husbands and wives will coordinate retirement plans with each other.

According to most economists, the primary motive for joint retirement decisions is the complementarity of leisure. For couples who wish to spend their leisure time together, leisure time becomes more valuable to one spouse as the leisure time of the other spouse increases. Couples may also want to relocate to different regions of the country as they approach later life, which would be easier to accomplish if neither spouse were working. Assortative mating may also lead husbands and wives to have similar retirement patterns; spouses with similar preferences over combinations of goods consumption and leisure are likely to retire at similar ages. Correlations in wages, pensions, and Social Security benefits between husbands and wives can also lead to similar retirement patterns.

Although joint retirement decisions of husbands and wives will probably become especially pronounced once Baby Boomers begin to retire, there is evidence that spouses from earlier cohorts coordinated their retirement decisions. For example, Blau (1998) examined retirement patterns among married couples between 1969 and 1979 with data from the Retirement History Survey (RHS). He found that wives were 63 percent more likely to stop working when the husband was not employed, and husbands were 53 percent more likely to stop working when the wife was not employed, than when the spouse was employed. Hurd (1990) found that about one quarter of married men who retired in 1980-81 left their jobs within 12 months of their spouse’s departure.

Several economists have recently developed retirement models that treat the couple, not the individual worker, as the unit of the analysis. Blau (1998) estimated a discrete-time, discrete-choice model of employment, in which the couple chooses the employment status of the husband and wife each period. Using a dynamic programming approach, Blau modeled transitions in the RHS among four possible states: both spouses employed, neither spouse employed, husband employed but wife was not, and wife employed but husband was not. Instead of specifying a bargaining model, he assumed that preferences could be expressed by a household utility function. His model included variables measuring education, age, experience, wages, Social Security
benefits, poor health, pension coverage (but not pension wealth), and race for each spouse, as well as the duration of spell in progress, type of prior spell, and the unemployment rate. The model was estimated using simulated maximum likelihood techniques.

Blau found that his estimated transition probabilities fit the actual data quite well, suggesting that retirement is appropriately modeled as a household decision, not an individual decision. He also found that the nonemployment of one spouse had large positive effects on the labor force exit rates of the other spouse, and large negative effects on the labor force entry rates of the other spouse. Mitchell and Pozzebon (1986), on the other hand, found very little effect on working wives of the husband’s retirement decision.

Gustman and Steinmeier (1994) estimated a structural model of retirement for married couples, using panel data from the National Longitudinal Survey of Married Women. They specified a standard utility function for the husband and wife, in which the well-being of each spouse was a function of their own goods consumption and leisure. In their framework, the labor supply of the spouse enters the maximization process in several ways. First, the spouse’s earnings can alter the budget set, thus affecting the level of goods consumption. Second, the retirement status of the spouse can affect the value of leisure. If leisure is complementary to the retirement status of the spouse, for example, it would be weighted more heavily in the utility function if the spouse were retired than if he or she were working full time. Finally, individual preferences for retirement may be correlated between husbands and wives.

Gustman and Steinmeier (1991) jointly estimated retirement behavior for husbands and wives under the assumption that the error terms followed a bivariate normal distribution function. The explanatory variables in the model were age, health, birth cohort, spouse’s retirement status, and the couple’s compensation stream, consisting of wages, pensions, and Social Security. The dependent variable was an indicator for leaving full-time work. When the correlation of the error terms for husbands and wives was omitted from the estimation, the husband’s retirement status had a small and insignificant effect on the wife’s retirement, but the wife’s retirement status had a significant positive effect on the husband’s probability of retirement. Simulations of the full model revealed sharp peaks in the predicted probability of retirement at the same age as the spouse, and these spikes appeared to arise from the effects of spousal retirement on the value of leisure, not from the correlation between husbands and wives of taste for leisure.

2. Bequests

Intergenerational transfers are estimated to account for a large fraction of wealth (e.g., see Kotlikoff and Summers (1981)). This result is inconsistent with the basic life-cycle model that assumes individuals reach the end of life with zero assets. Adding a bequest function onto the basic model, such as in (3.1), is a straightforward way to account for this result, but little information is available as to the form it would take. To the extent bequests are important to the intergenerational transmission of wealth and affect its distribution among cohorts, estimates of the wealth distribution of future cohorts that exclude them will be imprecise.
Huggett (1996) attempts to account for intergenerational transfers in his model and ends up with results that closely match the existing wealth data in the U.S. This is encouraging and suggests that future model development can be enhanced by a careful treatment of bequests. One promising direction might be to utilize the latest HRS/AHEAD datasets which inquire about planned bequests. Follow-up interviews attempt to trace how decedents dispose of their wealth.

3. **Other Sources of Income**

Most models of lifetime saving and consumption decisions assume that individuals receive income in the form of wages and earnings on existing assets. These models allow for the possibility that a bad income shock (or series of such shocks) could force consumption to zero. Of course, these models neglect the fact that there is often a government-sponsored, income floor below which individuals need not fall. Supplemental Security Income (SSI), disability payments, food stamps, Medicare and welfare benefits are examples of types of programs that ensure this minimum floor. The existence of these programs can be important factors in the income security of certain individuals and households.

The Hubbard, Skinner and Zeldes (1994) work mentioned earlier suggests that these programs can also have important effects on the saving patterns of some households. A model that allows for the smoothing of bad income shocks through an income floor is likely to provide further help in explaining the wealth accumulation of these households.

4. **Portfolio Choice**

Utility-based models of lifetime consumption and saving decisions that rely on the CRRA class of utility functions contain a parameter that governs the degree of risk an individual is willing to endure when implementing his optimal plan. It is also the case that this same parameter will affect asset allocation decisions over time and across assets. As individuals near retirement, they should become more risk-averse with respect to their portfolio composition because assets form a larger share of their total wealth. Clearly, how individuals balance their portfolio over time can have important consequences on lifetime wealth accumulation.

In the simple model (e.g., Samuelson (1969) and Merton (1969)), the individual owns two assets, a safe asset paying a riskless rate of return (a bond) and a risky security that fluctuates in value (a stock). The optimal strategy for the individual is to consume at a fixed rate from existing wealth at every point in time while keeping the portfolio constantly in balance according to a ratio that is determined by the parameters of the problem.

This model has the advantage of treating uncertainty over asset returns directly, albeit in a very simple two-asset world. Consumption shocks will occur as a consequence of changes in asset prices. (In the simple model, there is no labor income but wages can be added to the model with little change in results as long as they are known with certainty.)
An important recent extension of the basic model adds transaction costs (Davis and Norman (1990); Akian, Menaldi and Sulem (1996)). Here, the individual must pay a proportion fee or tax each time assets are sold. This results in a model where portfolio rebalancing occurs infrequently, resulting in “lumpy” transactions. This result seems to mimic what we observe in data on individual asset sales for large segments of the population.

V. CONCLUSIONS

Modeling saving is not an easy task. The economics profession has not reached a consensus on a theory of saving or agreement on how to interpret econometric results. Different approaches can yield very diverse predictions about how policy changes affect household saving.

The standard life cycle model (LCM) yields powerful predictions about saving behavior, but the predictions of the LCM have sometimes been inconsistent with empirical findings. Recent research, as discussed in this chapter, has improved our understanding of the factors and motivations that affect saving and suggested ways of modifying the simple LCM to make it more consistent with observed behavior.

The basic premise of the LCM is that individuals are forward looking in their approach to planning for the decline in earnings that accompanies old age. This means that they choose a consumption path over their lifetimes that balances the desire to consume now with the need to finance future spending when earnings will be lower. In the simple LCM, there is no uncertainty, utility is time-separable and based only on consumption and leisure, individuals have unlimited access to credit markets, and there are no bequests received or given. This model yields strong predictions about the path of consumption for any given pattern of lifetime earnings. A key implication is that consumption in any year depends only on lifetime resources and is unrelated to that year’s income. Another key implication is that consumption depends only on total wealth (including the present value of future earnings and Social Security benefits) and not on the form of wealth.

Enhancing the LCM to account for uncertainty, liquidity constraints, the effects of government transfer programs, and bequests makes the model more consistent with observed data. Modified approaches that account for these factors, and less traditional approaches such as the “behavioral” life cycle model, suggest that both current and lifetime income could affect current consumption and that the propensity to consume out of different forms of wealth may differ.

Existing large scale micro-simulation models typically contain very rudimentary treatments of household consumption. A standard approach is to project wealth of future retirees by a simple reduced form relationship that projects wealth at retirement based on wealth in some earlier period and other household characteristics. An alternative would be to introduce a consumption function directly into the model, compute saving as the difference between income and consumption, and build up household wealth year by year based on projected saving and assumed rates of return. Such an approach would have the advantage of allowing for direct simulation of changes in policies (such as individual accounts) on wealth accumulation. The
results from such simulations would, however, be highly sensitive to the exact specification of the consumption relationship. Because the current state of research still provides only limited guidance on how best to model consumption, it would make sense to test the implications of alternative specifications.
REFERENCES FOR CHAPTER 3


I. INTRODUCTION

Official estimates of the effects of proposed legislative changes on taxes and spending typically assume that the policy changes do not alter macroeconomic variables, although policy changes can alter micro variables. For example, an increase in the excise tax on cigarettes can reduce cigarette consumption, but is assumed not to change nominal GDP. Ignoring macro effects is defensible for short-run analysis of these policies because aggregate demand is mostly set by monetary policy—or at least, monetary policy can offset fiscal shifts. The aggregate capital stock changes little in the short run in response to many policy changes, as does, therefore, GDP.

Ignoring macro effects is untenable, however, when analyzing the long-run effects of Social Security reforms. These reforms can have major effects over time on both private saving and workers’ decisions when to retire. By doing so, they can substantially affect GDP, interest rates, wage rates, and other macroeconomic variables in the future. Moreover, it is important to know how different macroeconomic variables relate to one another. For example, changes in retirement behavior can materially affect the economy's capital intensity and, therefore, returns to saving. Even if researchers are primarily concerned with distributional outcomes, the macroeconomic effects of major policy changes cannot be ignored, because the price changes that are determined by macroeconomic forces feed back onto household incomes and behavior. If a policy change would significantly affect labor supply or capital accumulation, the macroeconomic effects on wages and rates of return need to be determined before we have an accurate story about what would happen to household incomes. For example, Kotlikoff, Smetters and Walliser (1998, 1999) show that greater funding of Social Security can increase the welfare of future generations of lifetime poor, even if benefits are no longer progressive, as might be the case in a system of individualized accounts. The reasons are macroeconomic, including the high opportunity cost of “investing” in Social Security instead of capital markets, as well as the increase in pre-tax wages as the capital stock grows with greater funding.

This chapter reviews the different types of models that may be used for the analysis of the macroeconomic effects of Social Security reforms and comments on their strengths and weaknesses. We examine the areas where the models are adequate and where further development is needed. We consider tradeoffs between adding detail to models and keeping the models sufficiently simple so that computational costs are reasonable and the determinants of findings are relatively transparent. This final version of our report clarifies the distinctions and
labels often used to differentiate the various classes of models and elaborates on the practical differences through some specific examples of policy simulations.

The “Request for Task Order Proposal” (RFTOP) described how this topic of macro models should place “particular emphasis on what is needed to extend such modeling to microsimulation modeling.” In this report, we use the term “microsimulation” to refer to the process of assigning (or “simulating”) socioeconomic circumstances to data on a household-by-household (i.e., “micro”) basis. These circumstances include aspects such as education, marriage, births, labor supply, asset accumulation, retirement, health status, and mortality. Microsimulation models might or might not contain behavioral responses in which household decisions, in reaction to economic variables such as prices and incomes, affect those simulated circumstances. These types of models are reviewed in Chapter 1 of this report.

Macroeconomic predictions can be integrated with a microsimulation approach, but they cannot be generated within the microsimulation framework alone. Microsimulation models focus on household-level decisions only; thus, even if these decisions could be added up to obtain predictions of consumer-side aggregates (such as aggregate consumption, saving, and labor supply), another type of model would be needed to specify the producer side of the economy (the demands for factors of production and supplies of goods) and to determine the prices that align the consumer and producer sides of the economy. Even with respect to modeling household behavior, a microsimulation model need not be used to determine effects on household aggregates in a “bottom-up” manner. It could be used in a “top-down” way instead, so that variables such as aggregate labor supply and saving are determined within a macroeconomic model and are simply divided up among households through the microsimulation model.

This chapter of the report surveys the different types of macro models that could be used in various ways alongside a microsimulation model. Macro models are distinct from household microsimulation models in that they must specify a production side to the economy, but among the great variety of macro models, most are very similar in how they actually model production. Most macroeconomic models characterize industry behavior by specifying explicit production functions and assuming profit-maximizing behavior. Beyond that, differences in production-side modeling may reflect different assumptions about industrial organization and different degrees of disaggregation.

Macroeconomic models are much more varied in their approaches to modeling household behavior. We make a broad distinction between “utility-based” or “optimizing” models where household-level equations are derived explicitly from the maximization of specific utility functions subject to certain assumptions about budget constraints, and “ad hoc” models that instead specify household behavior based on particular empirical relationships found in aggregate data.36 Most of

---

36The world of macro modeling unfortunately lacks a universal “glossary” of terms. In this report we have chosen to use the term “ad hoc” as synonymous with “non-utility-based.” The “ad hoc” approach refers to any type of macro modeling that does not derive household behavior as a solution to an explicit optimization problem. The label “ad hoc” is not meant as a pejorative term; it literally means “for a specific purpose, case, or situation,” which seems appropriate because these models are not tied to the values of the “deep,” invariant parameters of utility functions—such as the intertemporal elasticity of substitution—but instead focus on variables relevant to the particular economic effects.
the models discussed are distinct from microsimulation models in that they take a "representative agent" approach that distinguishes households by type--e.g., by age and income category--and summarizes everyone in each group with the behavior of one representative person. This is accomplished either through the characterization of a particular utility function and budget constraint, in the case of utility-based models, or through the imposition of a single set of parameters in aggregate household-side equations, in the case of ad hoc models. In the context of integrating microsimulation and macroeconomic modeling, however, a microsimulation, household-level approach could work in tandem with either a utility-based or an ad hoc macro model--at least in principle. If microsimulation is used in a "bottom-up" manner to predict household aggregates, these predictions can be generated either by simulating each individual's utility maximization problem or by simulating each individual's empirically-specified, ad hoc behavioral equations.

Besides understanding the basic differences between utility-based and ad hoc macro modeling approaches, a number of more specific characteristics need to be considered in recommending a particular approach for analyzing the macroeconomic effects of Social Security reform. These include:

1. the degree of disaggregation;
2. factor supply assumptions;
3. how models handle costs of adjustment in production;
4. specification of the government sector and fiscal policy assumptions;
5. measuring the effects of policy changes on welfare measures;
6. how to incorporate uncertainty of demographic and economic projections;
7. how best to include international capital flows; and
8. how to model non-standard saving behavior and non-rational expectations.

or policy problems considered. (In fact, this is the major criticism of the non-explicit, ad hoc approach; see later discussion on this point.) These models also tend to involve some degree of aggregation, where individual-level equations are collapsed, or "reduced," into single equations summarizing aggregate behavior. It is fairly common that these ad hoc equations are also purely reduced-form, in the econometric sense, but some of them are not. That is, some ad hoc specifications of consumer behavior allow feedback across equations, so that variables depend on other endogenous variables (not just exogenous ones).
A later section of this chapter considers a few specific policy examples—reducing Social Security benefits, establishing mandatory private accounts, and investing the trust fund in equities—and how different specifications of macro models might produce different assessments of these policies. A final section draws some general conclusions.

II. UTILILITY-BASED MODELS (EXPLICIT HOUSEHOLD OPTIMIZATION)

The standard types of models used in long-run macroeconomic analysis, including most computable general equilibrium models, are based on optimizing behavior by individuals and firms. In these models, agents maximize utility, which typically depends on consumption of goods and leisure in different time periods. The utility function generally incorporates an aversion to risk as well as a rate of time preference that weights future relative to current utility. Because such models are based on maximization of explicit utility functions, they can be used to measure the welfare effects of externally imposed policy changes.

Advanced utility-based models attempt to connect micro and macro variables explicitly. Aggregate variable levels, including the capital stock and labor supply, must be consistent with what households choose to supply on the micro level. What households choose to supply, in turn, depends on factor prices including wages and interest rates. Factor prices, in turn, depend on aggregate variable levels. An "equilibrium" requires that of all these variables are mutually consistent. An equilibrium also requires several other connections between micro and macro variables. In particular, fiscal policy variables must be consistent with household choices. For example, by definition Social Security payroll tax revenues (plus any intra-governmental transfers) collected in each year must equal benefits paid each year plus additions to the trust fund. Tax revenue in any year depends not only on the tax rate but also on how households change their labor supply in response to the tax. Similarly, non-Social Security federal, state, local and other taxes must also raise a certain level of revenue, consistent with household behavioral responses.

There are two main types of utility-based models. The first is the standard textbook Ramsey neoclassical model. The second model treats households as life-cycle consumers.

1. Ramsey Model

In the Ramsey model, an agent is modeled as having an infinite horizon. Although originally postulated for analytical convenience, infinite horizons may also reflect operative altruistic linkages between parents and their heirs (Barro, 1974). A parent who cares about his or her children who, in turn, care about their children ad infinitum effectively has an infinite horizon.

Despite being the standard textbook model, there are at least two problems with the Ramsey model. First, changes in the timing of taxes, including Social Security, are basically irrelevant in the Ramsey model. Social Security has no effect on household consumption, saving and labor decisions because an agent sees his or her children as a direct extension of himself or herself. Without Social Security, a parent decides how much wealth to leave his or her children. When Social Security attempts to redistribute resources from the children to the parent, the parent
simply undoes it privately by leaving a bigger bequest. This is known as "Ricardian equivalence." Although some researchers have argued that "Ricardian equivalence" might serve as a useful benchmark (Seater, 1993), the standard empirical tests seem to have very low power (Cardia, 1997). In the context of Social Security itself, the empirical evidence tends to contradict Ricardian equivalence and to support the notion that Social Security could play a potentially significant role in explaining the low saving rate in the United States (CBO, 1998).

Second, the standard Ramsey framework cannot accommodate heterogeneity among agents. Introducing multiple agents would admit a degenerate asymptotic solution: that is, either the most patient agent accumulates all the wealth in steady state or, if all agents are modeled with identical rates of time preference, the division of wealth is indeterminate. One option to deal with this problem is to assume that the rate used to discount the utility of future generations is increasing in consumption or wealth (Uzawa, 1968). Some have argued, however, that it is counterintuitive to assume that the rich are less patient than the poor (Blanchard and Fischer, 1989). The lack of heterogeneity is, of course, problematic per se (i.e., the world is very heterogenous). It also precludes multiple income classes, which prevents modeling the progressivity of Social Security, and it precludes modeling multiple countries.

2. The Life Cycle Overlapping-Generations Model

The second approach is to model households as finite-horizon life-cycle consumers. Agents face finite horizons in the sense that they care only about their own utility, but not directly about the utility of their children. Parents might leave a bequest, but the bequest has some basis other than the utility of their children. A common specification is that bequests reflect a "joy of giving." Agents are also "lifecyclers" in the sense that they care not only about how much consumption and leisure they enjoy today but also how much they will enjoy in the future. A person's productivity and wage is usually assumed to increase over much of the lifecycle and decline later toward retirement. Saving is motivated to pay for future consumption, especially after retirement. (Incorporating retirement and realistic lifetime wage profiles is a particularly difficult task in the Ramsey framework noted earlier.) The model also recognizes that generations overlap and, in particular, that children arrive before parents die. OLG models often explicitly require parents to spend resources on their non-adult children.

The OLG model does not share the problems of the Ramsey model. Social Security and other old-age programs (notably Medicare) have a profound impact on saving for retirement. Moreover, considerable heterogeneity along many dimensions (income and preferences) is possible in the OLG model.

3. Deterministic OLG Models

In a deterministic OLG model, a consistent equilibrium means that agents correctly forecast the future values of factor prices when making their consumption, saving and leisure decisions. While some view this perfect foresight as unrealistic, it does recognize that agents do try to behave in a rational manner. Moreover, perfect foresight models seem to generate a more
realistic—that is smaller—saving response to a change in interest rates than simpler myopic models where agents simply assume that any short run response in interest rates will last forever.

Despite their complexity, deterministic utility-based OLG models are becoming easier to solve. The most common solution technique is known as “Gauss-Siedel updating.” While slower than traditional Newton-Raphson methods, the Gauss-Siedel technique has excellent global properties and can be used to solve very large systems of equations. The Gauss-Siedel technique is typically employed by making an initial projection for current and future factor prices. Households then choose how much saving and labor to supply given this guess. Household saving and labor is then aggregated, as is earmarked payroll tax revenue and other tax revenue. The production function (e.g., Cobb Douglas) is then used to establish a new guess for factor prices. The government revenue requirements are used to make an update for the various tax rates. This process continues until the updates in factor prices become smaller and smaller. (Often, “relaxation” techniques are used which effectively dampen the speed in which factor prices are updated to avoid one part of the system of equations from converging faster than another part.) The resulting factor prices, household choices, aggregate variables, tax rates and revenue needs are then mutually consistent.

A large-scale lifecycle OLG model has been used by Kotlikoff, Smetters, and Walliser (KSW; 1998, 1999) to study various Social Security reforms. This model improves on its predecessor, the Auerbach and Kotlikoff (1987) model, by providing for intragenerational heterogeneity in addition to intergenerational heterogeneity. The new model also includes a detailed description of the U.S. Social Security, Medicare, and the federal, state and local tax systems. Utility is an explicit function of consumption and leisure in different time periods. The model calculates the effect of policy changes on the utility of separate cohorts, both by income class and across generations. The model, however, assumes a closed economy, although an option exists for simulating the U.S. economy as a small open economy.

These deterministic models, however, do not calculate explicitly the changes in utility associated with the changes in uncertainty of outcomes that might result from policies that alter the risk and return among investment portfolios (e.g., privatization or investment of a portion of the trust fund in equities). With perfect capital markets, however, some of these policies can be shown either to have no effect or to be economically equivalent at the margin to changes in payroll tax rates or benefits with no change in investment policy (Smetters (1997)). So under those conditions, the problems arising from the absence of explicit modeling of the disutility of risk can be overcome. However, other and larger policy changes will affect the size of the equity premium, the historic difference between stock and bond yields. A deterministic model, in contrast, incorporates only a single riskless asset. Fully-deterministic OLG models also fail to capture the effects of individual earnings uncertainty or uncertainty about future government policies.

4. Uncertainty in OLG Models

There are two types of uncertainty that have been modeled in the OLG framework. The first type, known as idiosyncratic earnings uncertainty, assumes that aggregate variables are
deterministic (as in the deterministic OLG framework discussed above) but that individual households face uncertain wages. In particular, the aggregate productivity of the economy is known but the productivity of each individual household is unknown. Like the fully-deterministic framework discussed above, this type of model incorporates only a single riskless asset. But the idiosyncratic earnings model does generate precautionary saving, as a means to weather downturns in earnings, in addition to saving for retirement. The dual nature of saving dampens the saving response to Social Security reform. Examples of these types of models include Huggett (1996), Huang, Imrohoroglu and Sargent (1997), and Huggett and Ventura (1999). The household maximization problem is typically solved by backward induction, i.e., “dynamic programming”. More advanced models iterate on numerous variables, as described earlier, to make sure that aggregate variables are consistent with household supplies.

Idiosyncratic earnings models have their drawbacks as well. One potential complication is that they might generate too much precautionary saving even without Social Security and Medicare in place to reduce the demand for retirement assets. Empirically, a vast majority of all wealth is held by the top few percent of wealth holders, who presumably are not motivated much by precaution. These models are unable to generate this degree of skewness in the wealth distribution. However, Hubbard, Skinner and Zeldes (1995) show that social insurance programs can be potentially quite powerful in explaining the low savings among lower income groups. A second problem is that, in practice, these models are not being used to solve for consistent equilibrium transition paths, except under very specialized and debatable utility functions. (Sometimes myopic transition paths are employed in which agents do not correctly forecast factor prices.) Instead, the models, at best, solve for consistent steady states. This limits their usefulness for analyzing the impact of policy changes over the next few decades. There is no reason, however, why a model cannot be built that includes both earnings uncertainty and correct forecasts of future factor prices. The absence of aggregate uncertainty implies that future factor prices can be perfectly forecasted, as in the deterministic models considered above, and can be incorporated into household decisions.

Another type of uncertainty in OLG models is aggregate uncertainty, which is the uncertainty surrounding economy-wide shocks such as those caused by oil price shocks, innovations, political events, wars, and weather conditions. In this case, both household-level and aggregate-level variables are stochastic. Such a model can incorporate a demand for both stocks and bonds and, hence, an equity premium. Models with production uncertainty in which agents live for many periods are computationally impossible to solve without some form of approximation (due to the well-known “curse of dimensionality”). As a result, mean-variance analysis is often employed. Mean-variance analysis, which assumes that capital returns are normally distributed, can severely underestimate the level of risk being taken because the theory implies agents might take risks that lead to negative consumption (in exchange for higher expected returns), even though most realistic utility functions are not defined over negative consumption. Storesletten, Telmer and Yaron (1998) used a stochastic OLG production economy to focus on the lack of market instruments available to hedge idiosyncratic shocks over the lifecycle. Their model focused on steady states. The strength of their analysis is in demonstrating that the stochastic lifecycle framework can yield very different predictions than the more conventional infinite-horizons specification. But they note that their model cannot replicate
the level of the equity premium. Generating a realistic equity premium, although not essential for their paper, might become more important for studying the non-marginal effects of some Social Security reforms including a proposal to invest a portion of the Social Security trust fund in equities or privatization with a government-backed benefit guarantee.

The aggregate uncertainty OLG model can also be used to analyze market incompleteness between generations that exist as current generations cannot negotiate contracts with future generations. This has been examined in Bohn (1998) and Smetters (1999) using two-period models. A two-period model can generate a realistic equity premium and can be solved without resorting to approximation techniques that typically eliminate important risk properties. While two-period models often yield quantitative predictions similar to their larger-scale counterparts, they are unable to produce higher resolution year-by-year numbers.

Ideally, the “best” OLG model would include aggregate uncertainty and many periods. But the “curse of dimensionality” prevents such a model from being solved correctly. Approximations are unlikely to be very reliable because they severely underplay the importance of risk. These tradeoffs suggest that a dual approach might be used by the SSA. A large scale model without aggregate uncertainty might be useful for analyzing many policy reforms such as changes in the level of benefits or even privatization without a pay-as-you-go benefit guarantee. (The value of a benefit guarantee, though, could be approximated using option pricing techniques and converted into an explicit amount of debt.) The model could include realistic demographics. Sensitivity analysis would be used to test various assumptions of paths for demographic variables (including mortality) and aggregate productivity. A smaller scale two-period model might be used to analyze the non-marginal effects of various investment policies, including investing a portion of the trust fund in equities, as well as benefit guarantees.

5. Making Utility-Based Models More Realistic

Utility-based models may restrict the behavioral response of agents too much. One concern of utility-based models is that some agents might be myopic. In particular, these agents might discount their future by more than is “good for them.” The empirical work by Samwick (1998), for example, suggests that a large variation in discounting might exist among US households. Utility-based models rarely incorporate this type of heterogeneity, but they easily could. A related concern of utility-based models is the standard assumption that agents do not face any liquidity (or borrowing) constraints. Empirically, as much as 20 percent of households might be borrowing constrained (Zeldes, 1989). Borrowing constraints can be incorporated into utility-based models; incorporating borrowing constraints into models solved via dynamic programming is, in fact, very easy.

Incorporating myopia and borrowing constraints is likely to improve the predictive power of simulation models, especially for welfare calculations related to the lifetime poor. Borrowing constraints are likely to be especially important for the lifetime poor for two reasons. First, the lifetime poor might have higher discount rates (or may be more “myopic”) than the lifetime rich; higher discounting might, in fact, be the reason why the lifetime poor are poor, i.e., higher discounting compels them to invest little in human capital. Second, as noted above, social
insurance programs themselves might encourage the poor to hold little private saving, possibly driving them to be borrowing constrained. Incorporating myopia and borrowing constraints is unlikely to be important for calculating the effect of social security reform on macroeconomic variables. This is because, for given wage profiles (i.e., given investments in human capital), the saving by the lifetime poor is likely to compose a very small fraction of aggregate saving, even if the lifetime poor had the same discount rates as the lifetime rich and social insurance programs did not exist. However, incorporating these features might have a more material impact on calculations describing the impact of changes to Social Security on welfare of the lifetime poor.

Another concern with utility-based models is whether the functional form selected for the utility function is the appropriate one and whether good choices are made for the deep parameters associated with the selected utility function. Even if we were to agree on the form of the utility function, various deep parameters, including the intertemporal substitution elasticity, have not been precisely estimated in the past; see Browning, Hansen and Heckman (1998), Mulligan (1998), and Ziliak and Kniesner (1999) for recent discussions. Unfortunately, one cannot discriminate between different choices for the utility function and deep parameters based simply on how well the model is calibrated to the actual economy. In particular, while sufficiently detailed utility-based optimizing models can typically be calibrated to observable economic variables, they might generate very different predictions about the effects of policy changes on macroeconomic variables. For example, Engen, Gravelle, and Smetters (1998) calibrate several different models and sets of model parameters – deterministic OLG, stochastic OLG, infinite horizons, and ad-hoc, each with several parameter set choices – to the exact same initial economy and show that they produce very different predictions of the effects of fundamental tax reform. The utility-based optimizing models must therefore be distinguished not on their calibration abilities but on their ability to replicate historical economic changes after a policy reform (what Joel Slemrod has sometimes referred to as “back casting.”).

III. **AD HOC MODELS (NON-EXPLICIT SPECIFICATIONS OF HOUSEHOLD BEHAVIOR)**

As explained earlier, we use the term “ad hoc” to refer to modeling approaches that do not use explicit household utility functions and hence do not compute household-level optimizations within their own structure. But this is not a particularly descriptive label with clearly-defined lines, as some ad hoc specifications are quite similar to explicitly-derived, utility-maximizing equations in their characterizations of household behavior and their attention to heterogeneity among households. While some ad hoc models use a small number of equations to summarize systems of equations, and where parameters are estimated from very aggregate data, others use separately-estimated demand and supply equations where parameters are derived from specific consumer and producer optimization problems and assumptions about the mix of different agents in the population.

1. **Advantages and Disadvantages of Ad Hoc Specifications**

By construction, the utility-based optimizing models will lead to predictions consistent with the particular economic theories defining them, while ad hoc specifications only attempt to
represent the empirical outcomes of undefined (and perhaps numerous) theories. But researchers will often confront a tradeoff between analytical rigor and empirical practicality; for example, the requirements of an explicit overlapping generations (OLG) framework can become especially daunting when adding features such as openness of the economy (goods and factor flows), monetary policy and price level effects, and various types of uncertainty (e.g., in health/longevity, wages, and rates of return) facing consumers. Hence, reducing the optimization problems of consumers and firms to a smaller number of equations—for example the utility maximization of many consumers to a single aggregate and “ad hoc” consumption function—can be considered a way of taking “shortcuts.”

The possible advantage of using “shortcut” approaches to capture the distributional effects on macroeconomic variables are: (i) as mentioned above, it allows richer modeling of other aspects, and so (ii) ad-hoc equations enforce less structure on the model. When considerable detail is desired, an ad hoc approach will be much easier, and potentially quite accurate for analyzing the types of reforms that have been seen in the past.

Ad hoc models, however, can have their own serious problems. The more fundamental limitations are largely unavoidable and due to the fact that these specifications are not based on explicit utility functions. In contrast, other problems are just typical in these models, but could potentially be alleviated by increasing the degree of disaggregation and/or improving the specification of variables.

First, the estimated econometric relationships used in these models, because they are typically not “deep,” invariant relationships contained within explicit utility functions, are unlikely to provide accurate predictions of changes not seen in past history. When the values of parameters within ad hoc specifications are derived within a particular economic and policy environment, one cannot assume that these estimated parameters would remain the same under a different environment. This is the so-called “Lucas Critique.” This criticism is presumed to apply more to ad hoc models than to utility-based models, because utility parameters such as the intertemporal elasticity of substitution are believed to be more fundamental and stable over time and within different policy environments than are parameters such as the interest elasticity of saving. But confirmations of the Lucas Critique in practice are surprisingly rare, and some economists consider the critique a “possibility theorem” more than an empirical truth.

Second, without explicit utility functions, models using ad hoc specifications are unable to measure exact welfare effects, and instead have to approximate those effects by looking at changes in incomes, consumption, leisure, etc. Because the behavioral equations in these models

---

37 A survey of modeling approaches by Ralph Bryant and Warwick McKibbin (1998) focuses on the use of such analytical “shortcuts” to capture demographics and compares this strategy to the use of OLG models that explicitly keep track of different cohorts and their economic decisions.

38 See Ericsson and Irons (1995) who also point out that many economists inappropriately use the term “Lucas Critique” in an indiscriminate manner when referring to any type of econometric misspecification, including the “more mundane” omitted variable bias.
are not derived explicitly from utility maximization, the estimates using them do not generate measures of the welfare changes resulting from changes in leisure, risk-bearing, or the intertemporal pattern of consumption. Distributional estimates generated from ad hoc models must therefore be interpreted with great caution, because changes in after-tax income generally do not correspond to changes in welfare, either in size or even in direction, when work effort, intertemporal discount rates, and the uncertainty of income are changing. For example, because after-tax income does not incorporate the value of leisure, a fiscal policy that increases after-tax income at the cost of a reduction in leisure could actually lead to lower utility. But under certain circumstances, ad hoc models can supply useful insights about winners and losers from policy changes, even without an explicit utility function. In particular, if the relevant behavioral responses on the part of households--their labor supply responses, timing/intertemporal decisions, or reactions to income uncertainty--are believed to be small, then changes in after-tax income may provide a sufficient approximation of welfare change.

Finally, ad hoc models typically involve more aggregation than utility-based models, although this need not be the case. The problem with using aggregated data to estimate parameters is that it can obscure the underlying distribution of resources and decisionmaking occurring at the household and firm level. For example, an empirically-estimated equation relating consumption to conventionally-measured wealth that is estimated over data generated during a time of tax reforms can often have the opposite sign as a relationship estimated over a time of an expansion in Social Security. This is because an aggregate specification of consumers would fail to take account of redistribution among generations. Even if one were to try to account for all households over all generations in computing aggregates, it is inherently difficult to come up with aggregate measures of certain crucial variables, such as an aggregate measure of Social Security wealth. For example, Steve Goss (1999) has estimated Social Security's "shut down" liability equal to about $9 trillion dollars with its "on going" liability closer to $3 trillion. The Pension Benefit Guarantee Corporation now accepts about a dozen or so different formulas. Unfortunately, these individual measures are not necessarily highly correlated because they focus on different variables. Thus, correctly measuring the aggregate and/or accounting for the disaggregate patterns of Social Security wealth is crucial to evaluating the impact of Social Security reforms on saving and other household decisions. This can be done rather naturally within the context of an overlapping generations, utility-based model, but could potentially be accomplished within an ad hoc model that differentiates among different cohorts.

2. Overview of Ad Hoc Modeling Approaches and Examples

All of the ad hoc models discussed in this section share a fundamental characteristic with the utility-based OLG models discussed earlier: macroeconomic effects are ultimately driven by a convergence to a "neoclassical" equilibrium in the long run. Various equations determine the supply and demand of factors and goods, with emphasis on the production functions and their technological parameters linking factor inputs to potential output. Those supply and demand "equations" can vary in the extent to which they are endogenous with respect to other variables in the system; in the simplest cases, many of the variables are set exogenously based on information from outside the model. The models solve for the real wages and rates of return to capital that will equilibrate markets in the long run. While in utility-based computable general equilibrium
models only relative prices matter, some ad hoc models solve for nominal prices. Typically, in the long run, the supply of output (as determined by production) determines the amount of output, and demand determines the nominal price level.

The simplest neoclassical growth models are versions with very aggregate equations, a limited treatment of the international economy, and little within-model feedback of prices onto consumer or producer behavior. (Exogenous variables can be adjusted according to microeconomic behavioral responses determined from outside the model.) Examples of this type of model include a model developed at Brookings (Aaron, Bosworth, and Burtless, 1989 and Bosworth and Burtless, 1997), and the long-term model used at the Congressional Budget Office (CBO, 1997). More recently, Rogers, Toder, and Jones (1999) use a more disaggregated version of this type of model in order to capture the heterogeneity across households of different ages and the effects of demographic change.

The more complicated models incorporating long-run neoclassical behavior are often referred to as “multi-region, general-equilibrium models.” Examples of these are models by the OECD (“Minilink”), IMF (“MULTIMOD”), McKibbin and associates (“MSG2” and “G-Cubed”), and the Federal Reserve Board (“FRB/Global”). These models solve for the prices (real interest rates and exchange rates) and patterns of resource allocation consistent with the movement of goods and factors across worldwide regions. In these models, “general equilibrium" often means that many interactions across markets are captured, but not that the economy is in a full market-clearing equilibrium at every point in time.

The Fed model is an example of how the lines between forecasting and general-equilibrium models are getting blurred; such models started out in the 70's as shorter-run macroeconomic forecasting models, but over time (largely as a result of the Lucas critique) have evolved into models with significant attention to the formation of expectations and the satisfying of longer-run neoclassical constraints. Private-sector forecasting models such as the DRI model and the Washington University Macro Model (WUMM) fall into this category of models that now have capabilities for both shorter-run forecasting of disequilibrium economies and longer-term policy analysis of equilibrium economies.

Yet most of these larger macroeconomic forecasting models are not typically used for the analysis of the long run. Many of these forecasting models have very sophisticated short-run properties that, however useful and important in certain contexts, might not be necessary to capture the essential macroeconomic effects associated with longer-run demographic and fiscal-policy changes. For example, to address its shorter-run forecasting requirements, the Fed's model

---


40 See Brayton et al. (1997).
contains detailed specifications of expectations formation, exchange rates, monetary policy, and price level effects. It is also based on a quarterly cycle. The modeling of many different regions—as opposed to just the “rest of the world”—in a global economy typically requires giving up some disaggregation of domestic consumers and producers. Moreover, having to specify stable policy benchmarks for a rational-expectations, general-equilibrium framework may imply that the models are not able to forecast movements from a truly unsustainable policy path. (This point is elaborated later.) Finally, any behavioral feedbacks in the models must be characterized according to aggregate equations in ways that must align with the rest of the large model, and such constraints may weaken any connection those aggregate responses have to the true underlying, microeconomic decisions. Hence, the smaller-scale neoclassical growth models mentioned earlier might have some appeal in providing a cleaner story focused on the longer-run effects on the domestic economy.

3. Macroeconomic Modeling with Microdata

The above ad hoc models are intended to reflect underlying consumer optimization theory to varying degrees. Yet these models still seem “macro” in that they characterize the household sector in a very aggregate way. Another approach—still qualifying as “ad hoc” perhaps, but also representing somewhat of a compromise—might be to use more disaggregate, partial-equilibrium modeling of household behavior and characteristics as inputs into a simplified general-equilibrium, neoclassical growth framework. Microdata could be used to parameterize or simulate the behavioral responses of individuals and households, and this might provide more realistic forecasts of responses to long-run demographic and policy changes than aggregate ad hoc specifications can. Some examples of the type of microsimulation models that could be used are described in Topic 1 of this report and include the “Dynasim” model developed at the Urban Institute, and the “CORSIM” model developed at Cornell. Coopers & Lybrand constructed an integrated model that combines a microsimulation tax model with a macro forecasting model. It uses household-level microdata to characterize household types and their distributions of income, which in turn determine aggregate consumption demand. The household side is then interacted with the production side of the economy within a general-equilibrium model. But the model is only partially a “bottom-up” one in that labor supply and saving decisions are still characterized by parameters applied to aggregate equations in the macro model.41

IV. HOW MODELS ADDRESS PARTICULAR ISSUES

Existing macro models take a variety of approaches in handling some of the more specific issues relevant to Social Security.

41 Brief descriptions of the Coopers & Lybrand model can be found in a volume compiled by the Joint Committee on Taxation (1997) on tax modeling issues. See the introductory sections and paper by John G. Wilkins.
1. **Degree of Disaggregation: Household-Level**

One can think of a spectrum of household-level disaggregation in macroeconomic models. At one extreme are the ad hoc models that summarize consumers with an aggregate consumption function that does not account for any heterogeneity in demographic characteristics. In the middle are the ad hoc models where consumption functions and sometimes labor supply functions depend on some variables that capture the demographic mix of the population, or models where there are different types of consumers (e.g., some life-cyclers, others rule-of-thumb). Utility-based OLG models also fall in the middle, because heterogeneity in the population is summarized by the optimization problems of representative agents (e.g., one per age and income category). Potentially at the other extreme would be models using household-level microdata and microsimulation models as inputs into a model of the full economy, although the extent to which disaggregation of microdata is preserved when interacting it with a macro model could vary.

2. **Degree of Disaggregation: Producer-Level**

Many of the macro models--whether utility-based or ad hoc in their characterization of households--specify a single, typically Cobb-Douglas, production function that captures capital and labor’s shares of aggregate value added. For many fiscal policy questions, industry disaggregation provides no additional information (see Gravelle and Kotlikoff, 1989). Some disaggregation might be useful, however, in modeling the macroeconomic effects associated with changing demographics because the demands for health care, education, and housing have strong life-cycle components, and these different consumer goods might be produced in very different ways, suggesting subsequent changes in the markets for capital and labor (see Bryant and McKibbin, 1998 for this argument). Examples of models with disaggregated production sectors include MSG2 and G-Cubed. The CBO’s model includes 5 production sectors. Although the Auerbach-Kotlikoff OLG model specifies a single industry, such disaggregation is possible within structural models (see Fullerton and Rogers, 1993, for example).

Some macro models that have only one production sector do disaggregate by capital type (examples are Bosworth and Burtless, CBO, and the Fed’s models), so that the mix of investment can vary over time, whether in an exogenous manner or in response to changes in costs of capital.

3. **Consumption, Saving, and Wealth Accumulation**

The issue of how to model household saving behavior is a critical one in the choice between utility-based and ad hoc models, especially in the context of analyzing the effects of aging populations and Social Security reform. Economists are far from consensus on a behavioral model that best describes how and why people save, and letting the microdata do the talking has not helped. This is bad news when one has to model behavior, whether through an explicit utility structure or through ad hoc equations. On the one hand, with very structured models (such as the
OLG models), the modeler commits to a whole, intricate story that could be incorrect. On the other hand, with ad hoc specifications, the estimation of equations based on aggregate consumption data could mask the true behavioral effects that occur at the individual level, especially if different people facing different circumstances—e.g., different patterns and levels of incomes and wealth—save according to different models.

The alternative models of saving in turn affect predictions of the economic effects of the current Social Security program and reforms to the program. This has the troubling implication that the choice of the behavioral model—which must be somewhat arbitrary given the lack of consensus in the economics profession—can affect the size and even direction of estimated effects of policy reforms. For example, under the simple life-cycle model of consumption with fully rational, forward-looking individuals, perfect capital markets (no borrowing constraints or other capital-market imperfections), and no uncertainty, the current PAYGO Social Security system should reduce private and national saving by transferring resources to the elderly, who have higher marginal propensities to consume than younger workers. Moving away from the “pure” form of the life-cycle model either by adding borrowing constraints, or by adding uncertainty in earnings, can dampen somewhat the negative effect of the current Social Security program on saving and hence dampen the positive effect of moving from PAYGO to a fully-funded system. In the Kotlikoff, Smetters and Walliser (1998, 1999) model, however, only those in the bottom 10 percent of lifetime income would want to borrow when young, and so adding borrowing constraints would have little impact on their results. But other optimization models suggest very different results. At one extreme, infinite-horizon models suggest that private bequest behavior would completely offset public intergenerational transfers so that Social Security and its funding status do not affect national saving. At the opposite extreme are some models of “seemingly irrational” behavior; an example is the “goal gradients” theory which suggests that the current Social Security program encourages people to save more than they otherwise would by enticing them to save more once they see the initial stock that Social Security puts in place. 42

The utility-based models, discussed earlier in Section II of this chapter, assume consumption and saving behavior basically follows the standard life-cycle hypothesis where consumers maximize lifetime utility subject to lifetime budget constraints, resulting in smooth paths of consumption over their lifetimes. With hump-shaped earnings profiles, this implies that people tend to borrow when young, save when middle-aged, and dissave (draw down assets) when retired, when the agent’s personal rate of time preference is close to market interest rates. When market interest rates exceed the rate of time preference (as, e.g., in the KSW model), the borrowing by the young might be minimal or even non-existent.

It is sometimes argued that adopting the standard life-cycle model—with forward-looking, fully rational households operating in perfect capital markets—is problematic “philosophically” in the context of Social Security reform, because the model cannot justify the existence of the current program. In order for Social Security to have some purpose, some part of this model must give: for example, at least some people have to be near-sighted, have incomplete access to

---

42 More detailed descriptions of alternative models of saving behavior can be found in Engen and Gale (1997) and in Chapter 3 of this report.
capital markets, or face certain types of uncertainty that Social Security can help insure against. This argument, however, assumes that Social Security was created paternalistically, to address myopia or deal with some market inefficiency. This might not be the case at all. Instead, the current Social Security system, being that it is not fully funded, could represent a redistribution mechanism from young to old that is supported by the median voter in the lifecycle model. Moreover, the current Social Security system is not necessarily very effective at dealing with risk sharing either within or between generations. Within generations, the progressive benefit formula could be replaced with an Earned Income Tax Credit that can replicate Social Security's existing progressivity and might even be Pareto improving if some people are borrowing constrained (Smetters, 1997b). Between generations, the wage-indexation of benefits means that Social Security, in the steady state, does not share productivity risk between generations in the same way that, for example, government debt does (debt payments are not wage indexed).

Another potential problem with adopting the life-cycle approach--within either a utility-based or an ad hoc specification--is that the empirical evidence on the validity of the model in the context of demographics is inconclusive. This is in part due to the fact that lifecycle models with realistic demographics have not been constructed in the past. (The newest version of the Kotlikoff, Smetters, and Walliser model, though, includes more realistic demographics.) With respect to aggregate relationships in the data, the life-cycle theory predicts that demographic change should affect aggregate saving rates. But there are various aspects of demographic change that affect the sign of this relationship in different ways, so that any particular estimate of the empirical relationship cannot prove or disprove the life-cycle model. On the one hand, a rise in life expectancy without adjusting retirement age would lead individuals to save more in working years to fund a longer retirement period, suggesting a positive relationship between the dependency ratio and the saving rate. On the other hand, decreasing fertility would not affect individual behavior, but would tend to lower aggregate saving as the elderly, with high propensities to consume, make up a larger fraction of the total population; this would suggest a negative sign. Fewer children means less saving for college, but it also means less required consumption by children while they are at home, and fewer young adults with relatively high propensities to consume (see Rogers, Toder, and Jones (1999)). The surveys by Meredith (1995) and Turner et al. (1998) find that most aggregate cross-section and time-series studies find a negative relationship between dependency ratios and saving, but span a broad range. In contrast, estimates based on disaggregate household survey data suggest either a much smaller negative, or even positive, relationship. In the United States, Bosworth and Burtless (1997) explain that the movement of the baby boom generation into middle age--peak life-cycle saving years, coupled with increasing retirement periods--should have caused a noticeable increase in aggregate private saving after the early 1980's, but instead private saving fell. More formal econometric tests of the life-cycle hypothesis have also failed to support the pure version of the model (see Deaton, 1992). However, more recent work shows that the lifecycle model can be quite consistent with the empirical evidence once proper controls are used (e.g., Attanasio and Browning, 1995). Gale and Sabelhaus (1999) show that the most recent drop in personal saving is tempered considerably if saving is defined as a change in wealth, including capital gains.

Despite the lack of a clear empirical relationship between demographic composition and aggregate saving, some ad hoc macroeconomic models attempt to capture life-cycle features by
including demographic variables, such as a dependency ratio or the labor-force participation rate, as determinants in an ad hoc, aggregate consumption function. These models also deviate from the pure life-cycle theory, however, by mimicking a certain fraction of their aggregate consumption as being determined by liquidity-constrained consumers.\footnote{\textsuperscript{43}} For example, The Fed's domestic model, FRB/US, specifies consumption as depending on financial wealth and the present discounted value of future labor income (a la the pure life-cycle model), but also depending on current and past income (a la the liquidity-constrained version). The consumption function of the OECD's Minilink model contains similar determinants, as do the MSG2/G-Cubed models. The Fed's global model, however, does not include wealth variables in the consumption function, so permanent income and the life-cycle approach are captured solely through past income. The model used by Bosworth and Burtless, in contrast, specifies consumption and saving functions that do not depend on life-cycle aspects and instead are fixed fractions of national income. CBO's long-term model also ignores life-cycle aspects but adds a Ricardian offset (ala the infinite-horizon model) to the exogenous component so that private saving partially offsets changes in public saving.

Summarizing the consumption and saving behavior of individuals through an aggregate consumption function can be problematic: basically, we cannot aggregate up without losing some of what's really going on. For example, a variable such as a dependency ratio might not be refined enough to pick up the differential effects of varying marginal propensities to consume versus changes in retirement periods. Aggregate measures of income and wealth might not accurately capture differences in the true budget constraints facing individuals (see Bryant and McKibbin). Keeping track of explicit cohorts and modeling lifetime budget constraints and behavioral responses--either through a utility-based OLG model or an ad hoc microsimulation approach--would capture important effects that the aggregate consumption approach would miss.

4. **Investment Adjustment Costs**

The "adjustment-cost" model of investment implies that firms find it costly to increase investment rapidly; thus, firms prefer to have less volatile investment. But in a closed economy, McKibbin and Wilcoxen (1998) show that this objective cannot be achieved simultaneously with the objective of life-cycle households to smooth consumption as income changes. Adjustment costs cause asset prices to be more volatile than would be expected by long-run economic fundamentals. Other formulations of adjustment costs, however, suggest that they might not be of a first-order importance (e.g., Barro and Sala-I-Martin, 1995). Whether we need to be concerned about this in a model where capital goods producers have some foresight is not entirely clear.

Because adjustment costs are primarily a shorter-run issue concerned with volatility in asset prices, most long-run OLG and macroeconomic growth models do not model adjustment costs, while the larger macroeconomic forecasting models capable of analyzing shorter-run issues

\footnote{\textsuperscript{43} This is motivated by econometric results from Campbell and Mankiw (1989).}
(such as those of the Fed) do. One exception is the Kotlikoff-Smetters-Walliser model referenced earlier. This model allows for a detailed specification of adjustment costs. Recent simulations reported in Altig, Auerbach, Kotlikoff, Smetters and Walliser (1999) of various tax reforms suggest that large adjustment costs can play a non-trivial role in slowing the rate at which capital responds to fiscal policy changes.

5. **International Capital Flows**

In terms of the relationship between saving and investment, the issue of international capital flows is probably of greater concern, and many of the models discussed in this report—both utility-based and ad hoc—attempt to capture these flows. As discussed by Bosworth and Burtless (1997), while the slowdown in population growth suggests that investment needs will decline, increases in longevity and retirement periods suggest the need for increased saving. An open economy might be able to reconcile these apparently conflicting needs by allowing the additional saving in the U.S. to be invested abroad. This possibility also depends, however, on differences across countries in their patterns of demographic change. Moving from a closed-economy view to an open-economy view would likely have implications for the rate of return to saving (allows it to be maintained despite increases in the capital-labor ratio) and for distributional outcomes (gains would go to savers instead of workers because the domestic marginal product of labor is kept from rising). It thus seems crucial that potential international capital flows be considered by a long-term model, and in turn, that such a model accurately capture any differences in demographic trends across countries.

The models discussed in this report take varying approaches to modeling the rest of the world. Some models are altogether closed, some interpret the U.S. as a small, open economy so that factor prices are exogenous, some model the U.S. versus the "rest of the world" or a fraction of the rest of the world (see Cutler et al. (1990)), and others—not surprisingly, models used by the Fed, OECD, and IMF—model several different countries.

Currently, there is not a large-scale lifecycle OLG model that allows for a realistic treatment of the open economy and, in particular, recognizes the US as a large open economy. Although one hypothetically could be built, it would represent a considerable undertaking. This is especially if important realistic empirical rigidities are to be incorporated. These rigidities include the high domestic correlation between saving and investment as well as the general failure of the factor price equalization theorem.

---

44While Bosworth and Burtless argue that the rest of the world will be able to support increased U.S. saving because other countries, particularly developing countries, are still growing rapidly, Cutler et al. (1990) argue in favor of reduced U.S. saving being optimal based on an analysis accounting for only OECD countries (the rest of whom are aging even faster than the U.S.).

45These rigidities are well-established empirical facts in the international economics literature. See, for example, Krugman and Obstfeld (1997).
Correctly modeling international capital flows—and the impact on factor prices—is potentially quite important. For example, Kotlikoff, Smetters, and Walliser (1999) show that, relative to the standard closed-economy assumption, the long-run increase in the American-owned capital stock, here and abroad, from prefunding Social Security is almost twice as large when the United States is assumed to be a small open economy. The differences between the two cases reflect the fact that more saving reduces the interest rate and increases wages in the closed-economy case, but not in the open-economy case. A higher choice for the intertemporal substitution elasticity relative to their benchmark setting tends to enhance the difference between the two models since the general-equilibrium feedback effects become stronger in the closed-economy version of the model (i.e., total saving does not increase as much). Including the demographic changes and saving responses of other countries is likely to be very important as well.

A fully consistent open-economy model would have to take account of offsetting effects on the trade flow and terms of trade of international capital movements. An example is a recent model by Gravelle and Smetters (1998) to analyze the incidence of a corporate income tax in an open economy. Their analysis shows that imperfect substitutability between domestic and foreign products can play an important role in limiting the flow of capital goods. Their model, however, does not focus on intertemporal decision making.

Because developing a large-scale intertemporal optimization model that is true to a theoretical foundation for international trade in goods and capital is quite difficult, economists have traditionally simplified in several different ways. We discuss each in turn.

**Reduced-Form Equations.**

Some researchers use simpler, reduced-form equations that bypass the theoretical problems noted above (e.g., Aaron, Bosworth, and Burtless, 1989, CBO, 1997, 1998). CBO (1997, 1998) uses a reduced-form elasticity based on the correlation between domestic saving and investment first estimated by Feldstein and Horioka (1980) and updated in later empirical work. In a similar vein, McKibbin and Sachs (1991) have also developed an international macroeconomic model (the MSG model) that explicitly represents capital and trade flows. Other models in use at the International Monetary Fund and the Federal Reserve Board may address international capital flow issues. The MSG, IMF and Board models, however, rely, in part, on “add factors” and other empirically-based adjustments as part of their calibration routine, thereby requiring some subjective decisions to be made.

**Small Open Economy.**

Researchers have sometimes modeled the United States as a small open economy in which factor prices are exogenous. This has the advantage of simplifying the computational requirements of the model. The model by Kotlikoff, Smetters, and Walliser (1999) allows for a small open economy option. The US capital stock is about 30 percent of the world's capital stock, however, so whether it is appropriate to assume that international flows of US capital would have no effect on global rates of return is debatable.
Ignore Capital Flows Altogether.

Some models examine only a closed economy. This is the simplest approach and might be justified by pointing to some evidence that US households appear to be substantially under-diversified in foreign asset holdings (Gordon, Roger H. and A. Lans Bovenberg, 1996).

6. Labor Supply

The models discussed in this report vary in the degree to which labor supply responds to changes in economic and demographic variables. Most OLG models do allow labor supply to respond to changes in the net wage rate and incomes. Retirement occurs when agents desire to supply a non-positive amount of labor. But these models often fail to represent demographic choices such as education, marriage, and childbirth, and hence cannot capture the likely influence of those choices on labor supply. To the extent that the United States will experience large demographic changes in the long run (regardless of policy), models lacking demographic influences might fail to get baseline predictions of aggregate labor supply correct. This in turn has implications for baseline saving predictions, because retirement plans and asset accumulation are no doubt intimately related. OLG models also assume that agents smooth leisure across time in the same way in which consumption is smoothed across time. The actual evidence on intertemporal smoothing of labor is mixed, although more recent evidence (reviewed in Altig, et al, 1999) suggests that this type of smoothing seems to take place.

Existing ad hoc growth models are even less satisfying in their modeling of labor supply responses to economic variables. A typical assumption is that labor supply is completely fixed with respect to prices and incomes, so that an exogenously-specified labor supply pins down the equilibrium quantity of labor, and the demand for labor determines the wage rate. Often the exogenous projections of labor supply, however, come from the SSA forecasts. To the extent that forecasts are based on microdata and microsimulation and can be adjusted, the non-explicit, ad hoc approaches might do better than the structured, utility-based models that are more constrained in how labor supply can be modeled.

Related to the issue of international capital mobility is the question of whether mobility of labor ought to be considered and modeled as well. The typical assumption in global models has been that capital is perfectly mobile and labor perfectly immobile. But the real world does not fall neatly into that simplified characterization. Given the differences in demographic trends discussed earlier, Bryant and McKibbin (1998) and Helliwell (1998) speculate that the pressures to permit greater amounts of immigration into the U.S. could intensify, with potentially significant effects on the long-term macroeconomic outlook. It is possible, for example, that increased immigration-together with increases in female labor force participation and retirement ages-- could at least partly substitute for desired increases in capital accumulation.

46See Toder and Solanki (1999).

Models differ significantly in the extent to which they model the government sector.

**Degree of Details.**

Macro models can have different levels of detail in representing government spending and tax policies. For example, the Congressional Budget Office (1997) includes in its long-term reduced-form model fairly detailed projections for spending on Social Security and health programs that account for demographics and projected growth in unit costs of health care. The disaggregation of government spending by major programs, especially those for which the demographic mix of the economy matters, is important in order to get accurate forecasts of long-run levels of aggregate government spending. Most of the macro models described in this report are quite naive, however, in their modeling of how government spends money.

Demographic change could also affect the mix of tax revenues and how total revenues will grow over time. Tax policies can also be represented in varying levels of detail. Most ad hoc aggregate models specify extremely simple tax systems (a single tax rate on labor income, for example), but utility-based models often do better. A typical method is to represent a progressive individual income tax by two parameters—an exemption amount and an average marginal tax rate, which may vary by income group. Some models, like the KSW model, model the tax system more precisely. In the Social Security context, including a detailed tax system is likely to be important for at least two reasons. First, changes in Social Security can induce general-equilibrium outcomes that affect the rest of the government's budget constraint. Second, and most importantly, KSW show that the choice of the tax base used to finance certain major reforms can have the dominating effect on the simulation results. Using a consumption tax to finance the transition, for example, can lead to quicker gains in the capital stock and output than using an income tax.

**The Government's Budget Constraint.**

A key issue in modeling the public sector is how to represent government fiscal constraints. Some ad hoc models have very rigid specifications; for example, the Bosworth and Burtless model assumes that government spending is a fixed, exogenous fraction of GNP. At the other extreme, models in the OLG tradition often assume an intergenerational budget constraint, in which the present value of non-interest spending over an infinite horizon must equal the present value of taxes. This constraint, however, prohibits examining the solvency of baseline policy because an insolvent baseline is mathematically impossible.

An alternative approach is that taken recently by the CBO. The CBO model essentially imposes no government budget constraint and examines the consequences of an aging population on government expenditures and cash-flow budget deficits. It projects a long-run disaster scenario in which deficits initially caused by rising spending on entitlement programs for retirees
raise interest rates and reduce investment and economic growth, which in turn raise the deficit even more. This approach is quite useful for examining solvency but might be less useful for examining policy reforms.

Another approach might assume that the federal government adheres to some short-run (e.g., annual) cash-flow budget constraint, or some target debt-to-income ratio. Such a model would have to state explicitly how that constraint is achieved. One could adjust discretionary spending, income tax rates, payroll tax rates, or other entitlement programs according to assumed formulas to meet specified short-run fiscal constraints. (The simplest specifications with a single overall tax rate--such as in the Fed model--typically allow that tax rate to be endogenous in the long run to achieve a long-run debt target.) These adjustments should, of course, produce further feedback effects on private sector activity, so that the model would have to solve simultaneously for macro variables and government policy parameters. Still, another approach might be to compare policy changes relative to alternative long-run baselines that are solvent. In the newest version of the KSW model, the authors compare policy changes relative to two solvent baselines: maintenance of present-law benefits with an immediate increase in taxes, and maintenance of present-law taxes with an immediate decrease in benefits.

8. **Incorporating Uncertainty in Exogenous Variables**

Parameters that are exogenous to the model are typically not known with certainty. These include preference parameters such as the level of risk aversion, the rate of time preference, the intratemporal substitution elasticity between consumption and leisure within a given time period, and the intertemporal substitution elasticity of consumption and leisure over time. Technology parameters such as the rate of exogenous technological change, the factor substitution elasticity and the capital share are not known with perfect certainty. The same applies to demographic parameters such as fertility, mortality, and immigration levels. A standard technique in analyzing policy is to consider a reasonable range of parameter choices. The most important parameters for analyzing Social Security reforms are the rate of technological change and the demographic parameters. Unfortunately, most OLG models abstract from important demographics by simply assuming a constant fertility rate with a fixed maximum lifespan and fixed mortality risk. The newest version of the KSW, however, allows for more realistic demographics, although it does not yet include immigration.

9. **Measuring Welfare and Efficiency Gains**

As noted above, ad hoc models are unable to translate economic effects into welfare effects. Moreover, there is no direct mapping between aggregate variables and welfare because, as noted earlier, a policy that increases GDP could be welfare decreasing if it leads to less leisure being consumed. One possible approach would be to approximate welfare effects by assuming particular social welfare functions that would take the levels of aggregate variables, such as consumption and leisure, into account. Within an ad hoc framework such an approach would be somewhat arbitrary, however, because the social welfare function used would not necessarily have
any connection with the implicit utility functions driving the household sector's behavior. Thus, at a minimum, a variety of functional forms should be considered.

OLG models can be used to calculate welfare. In the large-scale lifecycle OLG model used by Kotlikoff, Smetters, and Walliser (KSW; 1998, 1999) to study various Social Security reforms, utility changes associated with varying levels of leisure and consumption are calculated across income classes and for each generation. Changes in remaining lifetime utility are calculated using the endogenous path of the after-tax interest rate (possibly non-constant) as the discount factor. An equivalent variation measure is reported for each agent in each generation and each income class.

Calculating (Pareto) efficiency gains—as opposed to a welfare measure for each person—is harder. This is because efficiency gains represent a decrease in deadweight loss and are typically hypothetical in nature (i.e., what is left over after all the losers are fully compensated). One cannot directly sum lifetime utilities across agents (or, as is often sometimes done, sum lifetime budget constraints across agents) to get a measure of efficiency gains. Summing utilities across generations would require a subjective choice for a social welfare function. (A policy that increases a particular social welfare function may not be Pareto improving, although the converse is generally not true.) Instead, efficiency gains are calculated using a formal lump-sum compensation mechanism that fully compensates losers to bring their utility back up to a level equal to that before the policy change. Auerbach and Kotlikoff (1987) introduce the idea of the "lump-sum redistribution authority" that uses lump-sum transfers to compensate agents alive at the time of the reform so that their lifetime utility remains unchanged. All gains/losses are given to new cohorts.

V. SOME POLICY APPLICATIONS

This section briefly compares and contrasts how some different types of utility-based and ad hoc models would evaluate three Social Security policy options. As explained earlier in this chapter, many of the practical differences between current utility-based models and current ad hoc models are not inherent differences, but rather are coincident differences. Existing utility-based models tend to use more disaggregate representations of the household sector in which redistribution across households can be captured, and in which detailed aspects of individual budget constraints are measured. Existing ad hoc models tend to use more aggregate equations and data for the household side of the economy, and also tend to use highly simplified, abstract, consumption functions. But these differences should not be taken as generalizations, because it is possible to design a very aggregate and yet utility-based model (such as an infinite-horizon model with a single, representative agent), or a very disaggregate, yet ad hoc model (such as one that disaggregates by cohort and keeps track of individual budget constraints). As such, one cannot really talk about how the two broad classes of models in general would differ in their policy predictions, but rather how specific examples of models might produce very different stories, whether between the two classes or within them.
1. **Cut in Social Security Benefits**

Consider cutting a particular generation's future pay-as-you-go Social Security benefits with no (or little) change in the payroll tax rate they face. This experiment might correspond, for example, to the transition toward a funded pension system; i.e., this is the “transition” generation that pays but does not receive. In the lifecycle model, private saving by this generation would immediately increase -- potentially significantly depending on the size of the reform -- in order to offset the loss in Social Security wealth. Consumption, therefore, would decline immediately.

In the simplest type of ad hoc specification, where consumption is simply a function of current-period disposable income, there would be no change in consumption and saving, because expectations of future economic variables are immaterial. The same story would hold true for consumption functions that include conventional measures of wealth but fall short of explicitly calculating the lifetime budget constraints facing individuals. But it is conceivable that ad hoc consumption functions could be disaggregated by cohort and specified to include broader and longer-run definitions of wealth (including the present value of Social Security benefits), in which case the ad hoc model could simulate a decline in consumption similar to that generated within a utility-based, life cycle model.

2. **Mandatory Private Accounts**

Consider introducing a new additional mandatory private account with no change in Social Security benefits or future payroll tax rates under the baseline. In the standard lifecycle model, where there are no borrowing constraints and the borrowing and lending rates are identical, this policy change would be neutral: it would have no effect on consumption or saving. Households that are already engaging in retirement saving at or above the level required by the new accounts would simply reduce their saving outside the new accounts dollar for dollar, because they view the money in the new accounts as perfectly substitutable for their other retirement saving. Other households with little saving would simply borrow against the money in their new mandatory private accounts, thereby returning to their previous net saving level. Introducing imperfect capital markets -- e.g., borrowing constraints -- suggests that this policy might lead to a small increase in national saving by forcing those with little saving to save more, but probably not a big increase because the vast majority of saving is already done by households who are less affected by borrowing constraints and hence less likely to be induced to engage in new saving.

The simplest ad hoc specification, however, would suggest a potentially large decrease in consumption and increase in saving, because consumption is simply a function of current disposable income, which would decrease by the required contribution to the private account (the contribution is like a tax). Again, the disparity is caused by the fact that expectations of future economic variables do not affect consumption in the simplest specification. But one could include measures of future income or wealth in an ad hoc consumption function (such as what is done within the Fed's domestic model), in which case the mandatory accounts are likely to have a smaller effect on consumption because decreases in some forms of income or wealth would be offset by the new form of wealth in the mandatory account.
3. **Trust Fund Investment**

Consider investing a portion of the Social Security trust fund in equities. The riskiness of equities implies that either payroll tax rates or benefit levels must adjust to stock returns. Suppose that Social Security maintains its strict "defined benefit" nature which is often interpreted to mean that future tax rates would adjust. With Arrow-Debreu complete markets both within and over time, the lifecycle model recognizes that the risk passed to future generations is exactly equal in value to the increased expected return that future generations receive in the form of lower payroll taxes stemming from the expected greater return to equities.

Suppose it is Social Security benefits that then become risky. Then this policy change is economically neutral in the lifecycle model: agents simply undo a shift in the public portion of their portfolios by investing less in stocks and more in bonds (possibly shorting) in their private portfolios. If, instead, Social Security taxes are risky, then trust fund investment will impact prices. It might also impact corporate debt-equity ratios and the allocation of capital, as debt-financing contracts and equity financing expands. In this non-neutral case, if trust fund investment is being used to afford a level of benefits that are higher than would otherwise be given, trust fund investment would increase the unfunded liabilities placed on future generations (Smetters, 1997a). In sum, trust fund investment either leaves unfunded liabilities unchanged, because it involves only a switch in portfolio choices rather than a real move toward prefunding that increases national saving, or trust fund investment increases unfunded liabilities.

In a simple ad hoc model, risk is not properly accounted for. Instead, the modeler typically looks at expected returns and so the higher expected revenue to the government would appear as a "free lunch" and could be passed along to individuals as either a benefit increase or a tax decrease, in either case stimulating consumption. But in many circumstances, risk could be accounted for in the ad-hoc framework by using a risk-adjusted rate of return. For example, to the extent that the equity premium reflects risk, a bond rate would be taken as the return to a trust fund invested in equities, thereby eliminating the appearance of a free lunch. A generation-indexed wealth effect would also have to be incorporated, along with using a risk-adjusted return, if trust fund investment allocates the risk to future generations while allocating the expected return to earlier generations (Smetters, 1997a). This asymmetric allocation appears in many plans to invest a portion of the trust fund in equities. On several occasions the Social Security Administration has faced the issue of choosing an appropriate rate of return for trust fund assets that are invested in equities. In many of these cases, the Social Security Administration reported a range of projections using different rates of return. However, the range was limited to increasing the value of stocks at either a government bond rate or the higher historical average return to equities. Scenarios with returns below the risk-free rate were not considered. (See Bazelon and Smetters, 1999, for a review). This creates the false impression that the risk-free rate is a lower bound rate of return instead of a risk-adjusted rate.

VI. **SUMMARY/CONCLUSIONS**

Microsimulation analysis of the effects of demographic change and Social Security reform requires using long horizons and acknowledging that major reforms will likely have significant
effects on the macroeconomy. Microsimulation approaches on their own, however, cannot account for these long-run macroeconomic effects, because micro models lack the production side of the economy and the calculation of price adjustments that result from the interaction of supplies with demands. Thus, this report has reviewed some of the general strategies taken in existing macroeconomic models—strategies that might be used in conjunction with microsimulation in analyzing Social Security issues.

Empirical macroeconomic models can be classified in various ways, just one of them being whether or not household utility functions and optimization problems are explicitly modeled. The utility-based models with explicit optimizations have an advantage in theoretical rigor and the ability to measure the effects of policy changes on economic well-being. The ad hoc models may have an advantage in allowing for more sectoral detail and can often be tailored to modeling policy reforms that have been seen in the past. Although the highly aggregate nature of most ad hoc specifications may obscure important differences in the underlying decisions taken at the individual household level, this criticism can be at least partially addressed by further disaggregation and is not inherent to the ad hoc approach. Choosing a particular macro modeling strategy will inevitably involve tradeoffs.

In carrying out a microsimulation analysis (such as the Dynasim effort at the Urban Institute) that acknowledges macroeconomic effects, the degree and nature of the computational interaction between the micro and macro models could vary considerably. At one extreme, the two types of models could be integrated into one. In this case, the micro and macro models would communicate to each other in an iterative manner, where decisions of individual households in the micro model are added up to determine aggregate factor supplies and aggregate demands for goods, which are then fed into a production-side macro model which would simply provide the production equations and the price changes suggested by differences between quantities supplied and quantities demanded. Price changes computed in the macro model would then be passed back to the micro model, where individual households would react to the new prices, those new decisions would be added up again to determine household-level aggregates, and so on. With this approach of relying on the microsimulation model for the household side of the macroeconomy, the implications of using different types of macro models become less severe, because most of the key differences across macro models concern household-level behavior. At the other extreme, the microsimulation and macro models could work more independently. In that case the macro model would have its own consumer side as well as producer side, and the microsimulation model would be used to predict distributional outcomes among households but not behavioral outcomes. The microsimulation model would then simply divide up the household-side variables predicted from a macro model, whether those variables coming out of the macro model were computed in a highly aggregate manner or disaggregated by broad groups. Following this approach, the alternative types of macro models could provide substantially different stories about the consumer side of the economy.
REFERENCES TO CHAPTER 4


CHAPTER 5

STRATEGIES FOR INCORPORATING RISK, UNCERTAINTY, AND PRIVATE INSURANCE MECHANISMS IN MODELS OF SOCIAL INSURANCE

I. INTRODUCTION

This chapter evaluates the current state of the art in social insurance modeling, with particular emphasis on models that account for multiple types of uncertainty and private mechanisms for insuring against these risks. We examine two main strategies for incorporating various forms of risk and uncertainty into policy models used by the Social Security Administration (SSA). The first, or “short cut” approach, is to perform stochastic simulations of reduced-form econometric models and microsimulation models that do not attempt to derive endogenous behavioral and general equilibrium responses to policy changes. The advantage of this approach is that it can be implemented relatively quickly and inexpensively. The second is the DP/SGE approach, which involves substantial investments to develop credible individual agent dynamic programming (DP) models and multi-agent stochastic overlapping generations (SGE) models. The advantage of this approach is that provides internally consistent models that endogenously derive behavioral and general equilibrium responses to changes in Social Security policies.

Many of the investments necessary to implement the second approach have already been made. Improvements in computer hardware and numerical solution algorithms have enabled economists to develop increasingly realistic models that have succeeded in incorporating many relevant forms of uncertainty and features of the Social Security program. Some of these models are now sufficiently realistic to be useful to SSA in forecasting and policy analysis.

We provide a broad overview of recent developments in the rapidly growing recent literature on DP/SGE models. The models differ in a number of ways, including the methods used to estimate their unknown parameters. Many DP models are estimated by econometric methods such as maximum likelihood, using panel data sets with linked administrative data records, including Social Security earnings histories and benefit records. These data sets include the Health and Retirement Study (HRS) and the Survey of Income and Program Participation (SIPP). The unknown parameters of SGE models are typically estimated by calibration methods, with some of their parameters taken from previous individual-level micro studies and remaining

---

47 This chapter was prepared by John Rust of Yale University, under a subcontract to The Urban Institute. It reflects comments on earlier drafts by Gary Burtless, Jonathan Gruber, Eric Hanushek, David Pattison, Eugene Steuerle, and Eric Toder and editorial input from Urban Institute staff.
parameters informally chosen so that predicted outcomes of SGE models “best fit” corresponding observed quantities according to some metric that is typically not explicitly specified.

There are many issues about the best way to estimate and test the empirical validity of DP/SGE models that we do not cover here. However, it is important to note that these models depend on strong \textit{a priori} behavioral and expectational assumptions that may not be satisfied in the real world. Among the key assumptions that are typically employed are: 1) individuals are forward-looking utility maximizers, 2) individuals have rational expectations about uncertain future events and are fully informed about Social Security programs and private mechanisms and opportunities for insuring against risks, and 3) wages and interest rates are specified by relatively simplistic equilibrium models in which, for example, these quantities are set equal to the marginal products of labor and capital using a hypothetical Cobb-Douglas aggregate production function. These and other strong simplifying assumptions are imposed in order to yield DP/SGE models that are computationally tractable and also to identify some of the unknown parameters entering individuals’ utility functions and beliefs about uncertain events. If these underlying maintained assumptions are not satisfied in the real world, DP/SGE models could provide poor approximations to the behavior of real individuals and their forecasts could be seriously misleading.

SSA cannot wait indefinitely for the creation of the ideal model for Social Security policy analysis, if indeed one could be created. It needs models now that can estimate the behavioral and distributional impacts of a variety of relatively radical reform proposals, including the creation of individual accounts and investing a share of the Trust Fund in equities. In recognition that SSA may view existing “state-of-the-art” approaches as too difficult to implement or too experimental to use in the short term, we also examine promising short-cut approaches. The main objective is to account for key risks using stochastic simulations of relatively simple reduced-form econometric models or cell-based “accounting models” instead of more complex models that endogenously derive individual behavioral responses and general equilibrium impacts of policy changes. These simpler models use “judgemental” forecasting methods to represent key behavioral relationships needed to assess the effects of policy changes. The welfare and distributional impacts of policy changes can be assessed using various summary statistics such as “money’s worth” measures, or in some cases, expected utilities. We discuss some of the pros and cons associated with the short-cut approach, providing a number of examples that illustrate its limitations and its potential, in some circumstances, for yielding misleading policy conclusions.

Section 2 of this chapter motivates the need for next-generation models that are capable of accounting for risk and insurance. Section 3 lists the most important types of risks that individuals face in old age that Social Security needs to consider incorporating in its policy models. It emphasizes the importance of developing policy models that account for the key parts of the Social Security program in an integrated fashion and that consider actions individuals can take to insure against risks, including private insurance plans, pensions, other savings, and family transfers. Section 4 discusses “short cut” approaches, and Section 5 discusses the DP/SGE approach, including a summary of the current state-of-the-art in academic modeling of Social Security. Section 6 summarizes some of the substantive conclusions about whether the benefits due to improved risk sharing from social insurance programs such as Social Security outweigh the costs of providing them. These costs include reduced output, capital accumulation, and labor
supply resulting from the general equilibrium impacts of operating a “social safety net.” We summarize the pros and cons of the short cut and DP/SGE approaches to evaluating these benefits and offer some suggestions of areas where future modeling improvements could have the biggest payoffs from the standpoint of policy modeling at SSA.

II. THE IMPORTANCE OF ACCOUNTING FOR RISK AND BEHAVIORAL/EQUILIBRIUM FEEDBACKS

There is increasing awareness of the limitations of current generation policy simulation models, especially with regard to their treatment of risk and uncertainty. These problems are not unique to SSA. A recent paper by Bazelon and Smetters (1999), for example, argues that “a significant amount of policy analysis conducted inside the Washington DC Beltway is potentially very flawed due to the improper treatment of risks associated with future costs and benefits of many government projects and programs.” (p. 28). In the context of Social Security, the limitations of current policy models have been highlighted most recently by the dearth of models that allow meaningful distributional and welfare analyses of various proposals to “privatize” Social Security, including far less radical proposals such as the Administration’s plan to invest a portion of the Social Security Trust Fund in equities. It has long been known, however, that existing models are deficient in their treatment of many other risks, including risks related to labor market opportunities, health status, and health care expenditures. This limits their usefulness for analyzing a wide range of policies, including proposed changes to the early and normal retirement ages, changes in the Medicare eligibility age, and changes in the process by which disability awards are made.

Most of the policy and forecasting models currently used by SSA do not explicitly account for uncertainty. These models include the Projected Cohorts Model (PCM) developed in the Division of Economic Research, the Model of Income in the Near Term (MINT) recently developed by the Division of Policy Evaluation, and the collection of models used by Office of the Actuary for the long term projections in the annual Trustees Reports. In addition, none of SSA’s models account for endogenous behavioral or general equilibrium responses to policy changes. Finally, none of these models provide an integrated treatment of the different forms of risk for which the current U.S. social safety net provides insurance.

In this section, we provide examples that illustrate how accounting for risk, endogenous behavioral changes, and general equilibrium responses can have a big impact on forecasting and evaluation of the impact of alternative Social Security policies. We then present examples that illustrate the need to analyze multiple risks simultaneously. Before presenting these examples, however, we address two preliminary concerns: the technical difference between risk and uncertainty and the relative importance of accounting for risk and allowing for behavioral and general equilibrium feedbacks in Social Security policy analysis.
1. Some Preliminary Concerns

**Risk vs. Uncertainty.**

The academic literature distinguishes between risk and uncertainty. *Risk* refers to a situation in which outcomes are random draws from a known probability distribution (e.g., a roulette wheel.) *Uncertainty* refers to a situation where the random outcome has an unknown probability distribution. Although virtually all of the problems individuals face are best described as choices over uncertain outcomes (because people rarely have full knowledge of the probability distribution generating random outcomes), most academic models treat individuals as having full knowledge of the underlying probability distributions. Thus, they model individuals as making choices over risky outcomes, but ignore uncertainty. In this chapter, we use the terms “risk” and “uncertainty” interchangeably, although as a technical matter most models do not really handle uncertainty. In section V of this chapter, we briefly discuss new modeling approaches that can handle the subtle distinction between risk and uncertainty. In academic circles, however, the practical payoff from attempting to model uncertainty (as opposed to risk) is regarded as unclear.

**Modeling Priorities: Risk vs. General Equilibrium Effects.**

A brief review of the beginnings of the academic literature on Social Security suggests that allowing for behavioral and general equilibrium feedbacks had a higher payoff than incorporating risk into non-behavioral or non-equilibrium models. Indeed, the earliest work in this literature was successful in capturing Social Security’s role as a mechanism for intergenerational transfers and its impact on private saving without considering risk. The seminal contribution by Samuelson (1958) used a deterministic 3 period overlapping generations model to demonstrate that Social Security can be viewed as a “social contrivance” that corrects a fundamental market incompleteness -- the impossibility of writing binding contracts with the unborn. In Samuelson’s model, Social Security facilitates a Pareto-improving system of intergenerational transfers that makes everyone better off in the presence of Social Security than in its absence.

For many years, economists continued to ignore uncertainty and focused instead on extending Samuelson’s simple endowment economy with perishable goods (“chocolates”) to more realistic economies in which capital is subject to less than 100% depreciation each period and wages and interest rates are determined endogenously. These models include Diamond’s (1965) analytically solved 2-period overlapping generations (OG) model and Auerbach and Kotlikoff’s (1987) numerically solved 55-period OG model. In contrast to Samuelson’s model with perishable commodities, individuals in these models can save for old age by investing in physical capital at an interest rate equal to the marginal product of capital less the rate of depreciation. The key policy question addressed in these models is whether it is better to finance Social Security on a pay-as-you-go or fully funded basis, or more radically, whether it is better to eliminate Social Security altogether. The main conclusion that emerged from these deterministic models is that if the economy is dynamically efficient (i.e. if the real interest rate exceeds the rate of growth of the real wage base), then in the absence of myopia and borrowing constraints, a pay-as-you-go Social Security system is Pareto inferior to a lassiez faire economy without Social Security. This is because Social Security, by transferring output from the young (who have low
marginal propensities to consume) to the old (who have high marginal propensities to consume), increases aggregate consumption in the short run, thereby reducing aggregate saving, which lowers the aggregate capital stock and aggregate consumption in the long run.\textsuperscript{48}

Thus, deterministic models lead to the conclusion that in the absence of redistributational and paternalistic motivations for establishing a Social Security program, a pay-as-you-go system generally makes individuals worse off than a fully funded or “privatized” system. There is reason to distrust these conclusions, however, because this class of models fails to assign any value to the important risk-sharing features that Social Security provides. Indeed, it has been known at least since Diamond’s (1977) “Framework for Social Security Analysis” that another motivation for a government-run Social Security program is to correct market failure that limits individuals’ ability to insure against key risks using private mechanisms. In particular, it has been known since the seminal work of Rothschild and Stiglitz (1976) that moral hazard and adverse selection problems can lead to the non-existence of competitive equilibrium in markets for insurance of various risks. This appears to be the key explanation why markets for annuities, disability insurance, unemployment insurance, and health insurance either do not exist or are highly undeveloped. Analyses dating back to Eckstein, Eichenbaum, and Peled (1986) and Karni and Zilcha (1986) show how Social Security can “fill in” for these missing markets and make all individuals better off. Thus, in order achieve a better understanding of whether Social Security can be effective in counteracting market failure we need models that account for various forms of uncertainty, and also models that explain why private markets are incomplete or non-competitive.

2. Examples Illustrating the Need to Account for Risk and General Equilibrium Feedbacks

The remainder of this section presents several examples that illustrate the importance of accounting for risk and general equilibrium feedbacks.

\textit{Example 1: Projecting the Financial Impacts of an Aging Population.}

The Board of Trustees provides annual reports on the financial and actuarial status of the OASI and DI Trust Funds. These reports are based on forecasts by the Office of the Actuary at SSA of the revenues and costs of the Social Security program under current law for the next 75 years. The acceleration of increases in payroll tax rates and taxation of a portion of Social Security benefits in the 1983 Social Security amendments, combined with a temporary increase in

\textsuperscript{48}Note that if the economy happens to be at the golden rule allocation (i.e., if the rate of return to capital equals the rate of growth of the real wage base), then changes in Social Security have no effect on equilibrium allocations. If the economy is dynamically inefficient (i.e. if the return on capital is less than the rate of growth in the real wage base), then increasing the size of an unfunded Social Security program is Pareto improving. Also note that these conclusions were derived in overlapping generations models that ignored the possibility of bequests. If successive generations are linked by operative intergenerational transfers, then any change in Social Security can be neutralized by offsetting changes in individual bequests. This is the so-called Ricardian Equivalence proposition that was re-introduced by Barro (1974). Most economists do not believe Ricardian equivalence to be empirically plausible (see, e.g. Elmendorf and Mankiw, 1998). The conventional view is that the unfunded pension liabilities in a pay-as-you-go Social Security system can be viewed as a form of government debt that “crowds out” private capital formation, leading to lower wages and higher interest rates.
the ratio of workers to retirees, have increased OASDI Trust Fund balances in recent years. But in spite of scheduled benefit cuts associated with an increase in the normal retirement age, the steady aging of the U.S. population resulting from a decrease in birth rates and an increase in life expectancy will eventually cause Social Security benefit payments to increase faster than contributions and interest earnings. The Actuary’s 1999 forecasts predict that the OASDI trust fund will peak at about 4.6 trillion dollars in 2021 and will decline rapidly thereafter, reaching 0 by 2036. In order to keep the Trust Fund approximately balanced over the entire 75 year projection horizon, the Actuary calculates that the combined OASDI tax rate would have to be increased by 2.1 percentage points, from 12.4% to 14.5%.

Using the same demographic assumptions as the Actuary’s intermediate case projections, the dynamic stochastic general equilibrium (SGE) model of De Nardi, Imrohoroglu and Sargent (1999) results in very different predictions of Trust Fund balances and tax increases needed to maintain solvency. Their SGE model predicts that if current law benefits are maintained, OASDI contribution rates would have to be raised by 17.2 percentage points to keep the system in long run balance. Their SGE model predicts a much larger increase in tax rates than the Actuary because it accounts for the endogenous responses in individual saving and labor supply to increases in distortionary taxation. Their model also projects substantial increases in Medicare and Medicaid payments. These benefits increases require an additional 12.7 percentage point increase in tax rates to keep the combined OASDI and HI trust funds balanced, amounting to a total increase in tax rates of 29.8 percentage points over the 75 year forecast horizon. This large increase in tax rates reduces labor supply by 21% and the capital stock by 12%, causing total output to be 17% lower than it would have been if it had increased over this period at the exogenous rate of technical progress of 1.2% per year.

Of course, there is considerable uncertainty in any long-term forecast, and there is no easy way to determine whether the Actuary’s forecast or the forecast of the general equilibrium model is closer to the “truth”. The point of this example is simply to suggest that accounting for general equilibrium feedbacks can have a substantial impact on long term forecasts, and may lead to very different conclusions about the costs of delaying the fiscal adjustments to balance the system.

**Example 2: The Value of Social Security in a World of Incomplete Annuity Markets.**

A commonly cited rationale for the existence of a government run Social Security program is that it improves welfare by filling in for an incomplete annuities market. Expected utility calculations by Mitchell, Poterba, Warshawsky and Brown (MPWB, 1999) show that a risk-averse 65 year old retired male life-cycle consumer would be willing to give up one third of his wealth simply to gain access to an actuarially fair market for nominal annuities. Until very recently, however, there was no market for real annuities, and although an increasing fraction of Americans are purchasing nominal annuities, particularly variable annuities (Poterba, 1997), the overall market for individual annuity contracts is still very small. Due to well known adverse selection problems, most existing nominal annuities offer unattractively low rates of return to individuals who do not expect to live unusually long. For example, Walliser (1998) estimates that private annuities are 15 to 25 percent more expensive than actuarially fair annuities based on
average mortality. This had led some analysts, including Friedman and Warshawsky (1988) and MPWB (1999), to suggest market failure as the primary reason for the limited private annuities market in the United States. Because Social Security provides an approximately actuarially fair real annuity, a partial equilibrium analysis would imply that it would increase the welfare of Social Security beneficiaries by substantially reducing longevity risks.

Storesletten, Telmar and Yaron (1999) showed, however, that introducing an actuarially fair annuity market or a Social Security program that provides an equivalent benefit stream would reduce the precautionary savings motive, resulting in lower capital accumulations and income. Thus, the potential welfare gains that come from bolstering an incomplete market for private annuities are overwhelmed by the welfare losses resulting from the general equilibrium effects of this policy change.


Many proposed policy changes involve subtle changes in who bears various risks. An example is the Clinton Administration’s recent proposal to invest a share of the Trust Fund in equities while maintaining the existing “defined benefit” structure of Social Security. The only way such a policy can provide a guaranteed benefit is by shifting the risk of poor equity returns from current beneficiaries to future taxpayers. In contrast, recent proposals by others to increase the Medicare eligibility age shift risk in the opposite direction, from taxpayers to Social Security beneficiaries. The framework one uses to evaluate how different individuals are affected by such shifts in risks and rewards can have a big impact on one’s conclusion about the desirability of the policies in question.

Applying ideas from financial economics, Geanakoplos, Mitchell and Zeldes (1998a) argue that individuals would completely offset the impact of investing the Trust Fund in equities by corresponding changes in their private portfolios, so the policy change will have absolutely no effect. In contrast, Smetters (1999) appeals to another financial theory, option pricing, to value the implicit cost of the benefit guarantee. He estimates that if 40 percent of the OASDI Trust Fund were invested in equities, the cost of this benefit guarantee is equivalent to a 2.1% permanent rise in the OASDI contribution rate.

Simple portfolio models, however, are unable to value the potential benefits of equity investments. To do this, general equilibrium effects must be considered. For example, Bohn (1999) evaluated both the costs and the benefits of the policy change in the context of a simple SGE model. He finds that the net benefit to investing the Trust Fund in equities depends on individual preferences, the risk characteristics of equities, and their correlation with other macroeconomic and demographic shocks. He constructs plausible examples using calibrated parameter values to show that equity investments can be welfare improving for future taxpayers by improving intergenerational risk sharing. Abel (1999) also constructs a calibrated two period overlapping generations model and finds a net welfare gain, albeit a small one, from investing the Trust Fund in equities.
SSA needs models that can help it quantify how much various groups stand to gain or lose from reform proposals that shift risks between beneficiaries and taxpayers. A complicating factor is that many of the intended and unintended effects of policy changes, including the gains and losses experienced by various individuals, can be partially nullified by offsetting private responses that occur from individuals’ attempts to “self-insure” against various risks. These private include changes in saving, portfolio behavior, and other responses. Saving/portfolio responses include private insurance contracts, pensions and changes in holdings of financial securities and real assets. Other possible responses include changes in labor supply, human capital investments, altruistic transfers within families, and private charitable transfers among different individuals and organizations.

**Example 4: Social Security’s Effect on Markets and Institutions.**

Over a longer period, changes in Social Security policy can affect private markets and the structure of economic and social institutions. Much less is known about how to model these effects. The creation of Social Security Old Age benefits in 1935 and its subsequent expansion in the following four decades is thought to be one of the biggest contributing factors to the decline in labor force participation by elderly workers and the breakup of the extended family in the United States. For example McGarry and Schoeni (1999) estimate that increases in Social Security benefits account for over half of the expansion in the fraction of elderly widows living alone, which increased from 18 percent in 1940 to 62 percent in 1990. Expansions in Medicare and Medicaid coverage, particularly for nursing home care, may also lead to some shift of health and long-term care responsibilities for the elderly from the extended family to the government.

There are also allegations that the creation and expansion of Old Age and Disability Insurance programs has partially displaced private markets for annuities and disability insurance, but establishing these claims empirically is difficult. Eckstein, Eichenbaum and Peled (1985) showed theoretically that the creation of a mandatory Social Security program would partially crowd out the market for private annuities, although the overall effect of such an expansion would be Pareto-improving. Walliser (1998) used computational methods to show a similar result. He estimated that the rate of return and presumably the overall size of the annuities market would increase if the scale of Social Security were reduced. Empirical analyses suggest that the potential impacts of Social Security on private markets and institutions are substantially more complex than these theoretical models imply. For example, Poterba (1997) showed that the market for individual annuities in the United States did not start to grow rapidly until the 1930s, the same decade that Social Security was introduced. He notes that “Extended families, common in the 19th century, provided an informal alternative to structured annuity contracts. The falling incidence of multi-generational households in the early 20th century contributed to the growing demand for annuity products.” (p. 8). This suggests that it might be a fairly subtle problem to estimate Social Security’s marginal impact on markets and institutions, because the expansion or contraction of Social Security may coincide with broader changes in economic conditions and household preferences that are pushing private markets in the same direction.

The common theme in all these examples is that our analyses and forecasts of the impact of changes in Social Security policy can depend critically on the way we account for risk,
endogenous changes in behavior, equilibria, and the longer run structure of economic and social institutions. The complications noted above suggest that fairly sophisticated models will be required to capture these effects. There is still no clear consensus as to the best way to develop and empirically validate models of Social Security and the amount of new investment required to develop next generation policy models is likely to be substantial. But the rate of return on these investments could also be very high. An analogy to computerized automobile crash tests suggests that it will be far cheaper to evaluate proposed policy changes by sufficiently realistic computer simulation models than to attempt to discover improved policies by trial and error. It is clearly prohibitively costly and time-consuming to use amendments to Social Security as a form of “social experimentation.” Furthermore, recent analyses (e.g. Butler, 1999, Diamond, 1994, McHale, 1999) suggest that policy changes can impose significant “political risks” on current and future beneficiaries and taxpayers that are very difficult to insure against. This may be one of the reasons why it is so difficult to attain political consensus to make any significant change to the status quo.

3. Examples that Illustrate the Benefits of a Comprehensive Treatment of Risk in Social Security Policy Models

Social Security does much more than help individuals prepare for retirement. It is best viewed as a comprehensive social insurance program that covers individuals against a wide range of risks, including longevity (Old Age Insurance), mortality and widowhood (Survivor’s Insurance and the death benefit), the costs of health care (Medicare), disability and incapacity (Disability Insurance and Supplemental Security Income), and unemployment (Unemployment Insurance). Social Security also plays a significant role as a redistributive anti-poverty program, and few economists would disagree that it deserves much of the credit for significant reductions in the poverty rate among Americans over 65 (from 35% in 1959 to 10.5% in 1997).

It not as easy, however, to quantify Social Security’s success in insuring against the array of other risks that it covers. The huge political popularity of the program does suggest that its wider role in improving the “safety net” may be just as important as its impact on poverty rates among the elderly.

This suggests that we need a model that accounts for multiple risks in the context of a unified model of social insurance for an adequate understanding of the role of Social Security. Of course, it will be significantly more costly to build a unified model of social insurance that provides an integrated treatment of the key risks individuals face than to rely on separate models that treat each component of the Social Security program in isolation. There are, however, significant dangers associated with a piece-meal approach to modeling Social Security. The following examples show how ignoring important interactions between the different components of the Social Security program can lead to misleading predictions and policy conclusions.

Example 1: Interaction Between Old Age and Medicare Insurance.

Rust and Phelan (1997) show that a combination of risk aversion and incompleteness in the market for private health insurance causes changes in the Medicare eligibility age to affect the
timing of retirement greatly. They show that over half of the “spike” in retirements at age 65 is
due to “health insurance constrained” individuals who have employer provided health insurance if
they continue to work, but who would not have access to retiree health insurance if they were to
quit working prior to being eligible for Medicare. Most previous analyses that did not account for
the interaction between Social Security and Medicare were unable to explain the magnitude of the
age 65 spike, attributing it to a “social custom” (e.g. Lumsdaine and Wise (1994) and Gustman
and Steinmeier, (1995)). The Rust-Phelan model suggests that policies that attempt to delay
retirement by increasing either the normal retirement age or the delayed retirement credit that
were legislated in the 1983 Social Security amendments will have a relatively small impact unless
accompanied by a corresponding increase in the Medicare eligibility age. Other studies using a
variety of alternative econometric methods have come to similar conclusions about the importance
of the interaction between labor supply and health insurance.

In their survey of the literature on the relationship between health insurance and labor
supply, Currie and Madrian (1998) concluded:

“A large body of evidence supports the notion that health insurance affects
employment outcomes by giving individuals who rely on their current employer for
health insurance an incentive to remain employed, and by giving individuals with
other sources of health insurance provision less reason to participate in the labor
market. The effects appear to be strong among both older workers and married
women, although there appear to be effects on prime age males as well.” (p. 57).

Example 2: Interactions Between Old Age and Disability Insurance.

Another example of potentially important interaction between components of the Social
Security program is between Disability Insurance and Old Age Insurance. The Rust-Phelan model
and other models suggest that another form of market incompleteness --- liquidity constraints ---
can account for much of the peak in retirements at the early retirement age (currently 62).
Legislation designed to delay retirement by increasing the early retirement age has a powerful
effect in models that treat Old Age insurance in isolation. The total impact on retirements and
Social Security costs might be partially offset, however, by increased incentives to apply for
disability benefits. Even in the absence of changes in the early retirement age, the gradual increase
in the normal retirement age will reduce the fraction of the PIA paid at age 62 and thereby
increase the incentive to apply for DI benefits. The Office of the Actuary currently relies on
judgemental forecasts of these effects in its 75 year projections of the Trust Funds, predicting that
as the normal retirement age is gradually increased from age 65 to 67 over the period 2000 to
2022, the fraction of 60 to 64 year old males who will apply for DI benefits will increase from
17% to 26% and the fraction of 65 and 66 year old males who apply for DI benefits will increases
from 0 to 17%. I am aware of only one econometric study that predicts the effect on employment
levels and DI rolls of the scheduled increase in the normal retirement age (NRA) from 65 to 67
between 2002 and 2022 (see, Wittenburg, Stapleton, et. al.1998). This study finds that 16
percent of Old Age beneficiaries and 10 percent of Medicare beneficiaries would retain their
coverage via DI even if the NRA and Medicare Eligibility Age (MEA) were increased to 67.
Thus, when we account for the interaction between the DI and OA programs, the total budgetary
saving from raising the NRA and MEA to age 67 is far less than would be predicted by models that ignore this interaction.

The failure to treat Social Security programs in an integrated fashion can also lead to misleading conclusions about its distributional effect. For example, a recent analysis of the internal rate of return (IRR) to Social Security by Beach and Davis, (1998a) found very low and sometimes even negative rates of return for certain groups such as young black males. However, even ignoring criticisms about actuarial mistakes in carrying out these calculations (see Beach and Davis, 1998b), their results are suspect because they ignore disability insurance as one of the “payoffs” received by Social Security beneficiaries. Given the higher propensity of black males to apply for and be awarded DI benefits, an analysis that considers both OA and DI benefits would be less likely to find the large racial disparities in rates of return that Beach and Davis found. Indeed, Goss (1995) showed that when DI benefits are included in the IRR calculation, the Social Security system continues to remain progressive, with low earners receiving higher internal rates of return on contributions than higher earners, even after accounting for the higher mortality rates of low earners.

These examples show the potential dangers in making policy forecasts and welfare evaluations in the context of piecemeal models of individual behavior that only consider individual components of the Social Security program in isolation. SSA should give serious consideration to developing more comprehensive integrated models of Social Security.

III. DIMENSIONS OF RISK AND UNCERTAINTY

1. Sources of Risk

Some analysts within SSA believe that analyses of risk and uncertainty should be structured according to a “two-by-two” classification that distinguishes between 1) labor earnings risk and rate of return risk on savings invested in equities, and 2) idiosyncratic vs. aggregate risk. This is a useful classification, but it neglects other sources of risk that may be equally important to consider in policy models:

- *financial risks* associated with adverse wage changes, unemployment, and uncertainty about future job opportunities;
- *financial risks* associated with uninsured health care costs;
- *risk of lost job earnings* and pain and suffering associated with poor health and disability;
- *personal risks* associated with divorce or death of spouse and or other family members;
- *financial risks* associated with uncertain asset returns, including equities and housing;
- *longevity risks* arising from limited access to actuarially fair real annuities and pensions;
- *political risks* associated with unanticipated changes to Social Security tax rates and benefit formulas;
• insolvency risks associated with privately provided defined benefit pensions, annuities, and health insurance; and
• mobility risks associated with loss of pension coverage due to incomplete vesting and lack of portability of private pension plans.

2. How Individuals Deal With Risk

Individuals can deal with the risks listed above through a number of mechanisms. These include:

• private insurance purchased from firms;
• employer pensions;
• private transfers from within the family and from private charitable organizations;
• “self-insurance” through precautionary savings, human capital investments, and labor supply decisions;
• Social Security benefits, including OASDI, UI, and Medicare; and
• transfers from other state and federal government programs.

For most Americans, personal labor supply, savings, and Social Security are the main strategies for dealing with these risks, with Social Security becoming increasingly important as they age. As noted in section II of this chapter, the market for individual annuities is very small in the United States. Aside from private pensions and personal saving, Social Security provides the main retirement annuity for a majority of Americans.

Similarly, large fractions of the U.S. population are not covered by private disability insurance or health insurance, while Social Security provides nearly universal disability coverage (although with low replacement rates for individuals in high income brackets) and health insurance coverage for individuals over 65 (although with limited coverage for certain catastrophic health care costs such as extended nursing home care).

---

49 Medicare is also available to DI recipients after two years. We refer to DI coverage as “near universal” since workers are disability insured only if they had 20 quarters of coverage (QC) during the last 40 quarters, and are currently insured (i.e. they earned 6 QCs in the preceding 13 quarters, including the quarter of death).
The level of non-housing, non-pension financial savings appears too small to finance a significant share of retirement income for most Americans or to provide an adequate reserve against the possibility of disability (see, Gustman and Steinmeier, 1999). It does, however, appear to provide a significant buffer stock for unexpected short term events such as job loss or unexpected uninsured health care costs. Recent theoretical and empirical work by Hubbard, Skinner and Zeldes (1994, 1995) and Carroll and Samwick, (1995) suggests that a substantial share of private wealth, as much as 50 percent, can be attributed to such precautionary, self-insurance motives. A growing body of recent empirical work also suggests that family transfers is an important mechanism that partially plugs gaps in incomplete private insurance markets (Altonji, Hayashi and Kotlikoff, 1996, and McGarry and Schoeni, 1995).

3. Risks That Social Security Models Should Address

A comprehensive model of Social Security would incorporate the main components of the program (OA, SI, DI, UI and Medicare), the key risks facing individuals listed in Section III.1 above, and the effects of private responses to risk listed in Section 111.2, especially responses of private saving and labor supply. Of course, in the short run it is not feasible to build integrated models that encompass all dimensions of risk. It is therefore necessary to begin with models that incorporate the most important risks that concern individuals.

Evidence of Risks that Matter to Americans.

We are unaware of data or research that would allow us to gauge accurately how Americans rank in importance the various risks they face. Public opinion and voter surveys, however, provide some indirect evidence of Americans’ concerns. These polls indicate that the stability and solvency of Social Security is high on the list of worries of older Americans. For example, the Voter News Service conducted interviews of 1,109 voters aged 60 or over who were leaving polls after the November, 1998 elections. Among these voters, 31 percent listed Social Security as the issue that mattered most in deciding their votes for Congress; 11 percent listed taxes; 10 percent listed the economy and jobs, and 8 percent listed health care.

Individuals’ ranking of their most important issues in choosing political candidates may not necessarily be the same as their ranking of the risks that concern them the most. For example, even though health care was listed as only the sixth most important political issue in the voter survey, introspection suggests that it should rank nearly as high as concern over the solvency of Social Security and risks of job loss and adverse changes in labor earnings. Indeed, the latest statistics indicate that over 44 million Americans, 16 percent of the U.S. population, have no health insurance. Despite the strong economy, over a million families are losing health insurance each year (Marmor and Mashaw, 1997). Of particular concern to many elderly is the cost of uninsured nursing home expenditures. Even though 10 percent of public medical spending is devoted to long-term care, the majority of nursing home costs are still uninsured (Cutler and Sheiner, 1998). Catastrophic health costs that exceed medicare reimbursement limits are also a major concern. Feenberg and Skinner (1994) used simulations of dynamic programming models to show that “catastrophic health care insurance could yield large utility gains.”
These data and research findings are consistent with the apparent large concern over preserving and extending Medicare among senior citizens. Politicians who emphasize health-related issues may be in touch with the “issues of real importance to seniors --- i.e., the patient bill of rights, getting prescription drugs to people on Medicare at a good price, long-term care insurance and tax breaks for long-term care, shoring up and keeping strong Social Security and Medicare.”(quote from New York Times article, “Shift by Older Voters to G.O.P. Is Democrat’s Challenge in 2000”, May 31, 1999).

**Subgroups of the Population at Risk.**

SSA may want to prioritize certain socio-demographic groups for which particular concern is warranted. It might be reasonable to assume that most high income individuals are able to “fend for themselves”, so that attention can be focused on the most vulnerable parts of the population including those in the lowest income brackets, minorities, widows, children, and divorcees with dependent children. According to Burkhauser and Smeeding (1994), “The single greatest risk of falling into poverty in old age now comes after the death of a spouse, as the survivor faces life after marriage.” (p. 1). For example, poverty rates among single women are 4 times higher than the poverty rate for married women. Data from the Health and Retirement Survey (see, e.g. Glaes, Estes and Smeeding, 1999) indicate that 41% of single women aged 51 to 61 fall into the lowest population decile of wealth, with a mean total wealth (including discounted Social Security Old Age benefits) of only $50,500. Social Security comprises over 92% of the wealth of these women, making them especially vulnerable to recently discussed changes in program rules such as increases in the minimum retirement age or decreases in the minimum benefit level.

Another particularly vulnerable group of Americans are those who suffer from disabilities and chronic health problems, including the nearly 5 million Social Security Disability Insurance recipients. A recent analysis by Benitez-Silva et. al. (1999b) using the Health and Retirement Survey (HRS) data documents that the disabled are severely economically disadvantaged relative to the non-disabled population, with annual labor income that is one tenth as large, and household income and net worth that are less than one half as large as the mean values for these quantities in the non-disabled population. In addition to having far worse health and higher incidence of impairments according to virtually all available measures of health and “activities of daily living”, the disabled also have much higher medical expenditures, and significantly lower life expectancy than the non-disabled. For example, the incidence of severe medical conditions such as cancer, diabetes, lung disease, stroke, and congestive heart failure is more than three times as large among the disabled than the non-disabled. Approximately 67 percent of the non-disabled population expects to live to age 75, but fewer than 50 percent of the disabled population expects to live that long.

Compounding their problems, the fraction of the disabled population that have support from other family members such as a spouse is actually lower than in the non-disabled population. For example, while over 80 percent of the non-disabled population is married, less than two thirds of the disabled population is currently married, and the fraction of the disabled who have never been married is twice as large as in the non-disabled population. Divorce rates among the disabled
are twice as high as the non-disabled: 30 percent compared with 15 percent. Thus, the disabled individuals who qualify for DI or SSI are likely to be highly dependent on these benefits, and would probably be greatly affected by reductions in benefit levels or restrictions in their access to medical care that might result from policies designed to reduce Medicare costs such as shifting recipients into managed care. Benitez et al. (1999a) find that DI applicants face a significant risk of rejection, but this risk can be reduced by exercising the option to appeal at the cost of long delays between an appeal and the final decision. Further, Benitez et al. (1999b) provide evidence of substantial classification errors in the DI award process, with many rejected DI applicants appearing to be “observationally equivalent” to DI recipients. Although their analysis also identifies a significant fraction (nearly 20%) of “non-disabled” DI recipients, a major concern is the relatively large fraction of rejected DI applicants (nearly 50% after accounting for eligibility restrictions such as means-testing of other related sources of support such as SSI) who appear to be truly disabled according to SSA’s definition of disability, but may have few other sources of support to fall back on.

Despite all the press and recent attention to the issue of investing the Trust Fund in equities, it is not clear whether the risks of equity investments would be extremely high up on the priority ranking of the risks that many elderly are most concerned about. There are several reasons why this might be the case. First, only a minority of Americans, most with relatively high incomes, have substantial equity holdings at the time of retirement. Second, stocks have performed so well in recent years that many people have become less aware of their downside risks relative to their upside potential. Finally, many proposals for creating individual accounts or investing the Trust Fund in equities involve implicit minimum benefit guarantees that effectively shift much of the downside risk of poor equity returns to future taxpayers. Other risks discussed above, including the risk of job loss, health risks including disability, and uninsured health care costs, particularly for catastrophic care and nursing home care, may justify higher priority than equity risks when considering potential modeling improvements.

The following sections of this chapter do, however, devote special attention to models that are capable of evaluating policies that affect investment risks. The emphasis on modeling investment risks reflects the substantial current interest in proposals to introduce individual accounts or to invest a portion of the Social Security Trust Fund in equities. In the long run, however, models need to address how changes in the Social Security rules affect the broader types of risks that most people face and need to consider how behavior will change in response to changes in these risks.
IV. “SHORT-CUT” APPROACHES TO INCORPORATING RISK AND INSURANCE IN POLICY MODELS

This section reviews short cut or simulation-based approaches to incorporating uncertainty in policy models. This group of methods presupposes that we have access to a reliable behavioral or non-behavioral probability model to generate stochastic simulations of various outcomes of interest for the policies under consideration. The analyst assesses the relative desirability of different policies based on a subjective evaluation of the distributions of simulated outcomes.

Because stochastic simulations generate a large amount of “artificial data” that can be difficult to digest, short-cut approaches use relatively “objective”, low-dimensional, and easily calculated summary measures of individual and social welfare resulting from alternative policies. Examples include expected discounted utility, expected discounted wealth including the expected present value of Social Security benefits net of taxes, internal rates of return on Social Security contributions, and related “money’s worth” measures. In cases where policies involve evaluating financial risks, such as investing the Trust Fund in equities, one could use a risk-adjusted discount rate to compute expected present values or utility levels. While this is one possible way of correcting for risk, the process of choosing a risk-adjustment factor is rather ad hoc. A better approach is the use of market-based methods such as option pricing theory to evaluate the price of risk. Option pricing theory can be used to calculate the “premium” that would be required to insure against the risk in question. It could be used, for example, to estimate the costs imposed on taxpayers by the Administration’s recent proposal to invest Social Security Trust Funds in equities, while maintaining minimum benefit guarantees to Social Security beneficiaries.

The common features of proposals we refer to as “short cut” is that they do not derive simulation results endogenously from the solution to an individual’s dynamic optimization problem and do not account for general equilibrium feedbacks by solving an SGE model explicitly. Instead, major simplifications are obtained by obtaining simulation results from an exogenously specified stochastic process. Endogenous feedback and policy effects are either ignored or obtained from extrapolations or judgemental adjustments to the underlying simulation model. This avoids the need for more complicated numerical dynamic programming methods to derive individuals’ “best replies” and general equilibrium methods to obtain simultaneous solutions in which all markets clear. It also avoids the associated econometric problems involved in selecting empirically justified values of the unknown parameters characterizing individuals’ and firm’s beliefs, preferences, and profit (objective) functions. These problems require a more ambitious approach that is the topic of section V.

1. An Illustration of the Short-cut Approach to Accounting for Risk.

A recent paper by Feldstein and Ranguelova (FR) (1998) provides a good example of how the short cut approach can account for risk at the individual level in analyzing the impact of “privatizing” Social Security by creating individual accounts that can be invested in equities. Risk analysis at the individual level and analysis of policies that directly or indirectly invest money used
to fund Social Security retirement benefits in equity markets are both central topics of this chapter.

FR evaluate the desirability of moving from the current Social Security system to a privatized system of individual retirement accounts which are invested 60% in equities and 40% in bonds. Individuals contribute a fixed percentage of their wages to personal retirement accounts (PRAs) during their working years. It is assumed that they work continuously between age 21 and retirement at age 67. Various contribution (tax) rates are considered, including 4%, 6%, and 9%. Upon retirement, the cumulative value of the PRAs are used to purchase a variable annuity. FR model stock and bond returns as independently and identically distributed random variables. They conduct 10,000 stochastic simulations of bond and stock returns in order to calculate the distribution of the cumulative value in an individual’s account at age 67, and payoffs from a variable annuity between age 67 and the individual’s death. Death occurs sometime between age 67 and the maximum age of 100. They compared the distribution of retirement income streams from this PRA/variable annuity plan to a benchmark equal to the stream of benefits that the same individual would receive under the current Social Security program.

If all the assets in the PRA were invested in risk-free assets yielding a 5.5% real return, individuals would only need to save 3.1% of their incomes to purchase an annuity stream equal to that promised under the current Social Security program, or about one sixth of the tax rate that the Office of the Actuary predicts is necessary to maintain the current pay as you go Social Security system in long run actuarial balance. To model the case where PRA balances are invested in a portfolio containing 60% equities and 40% (risky) bonds, FR assumed that the returns on this portfolio would have a mean of 5.5% and a standard deviation of 12.5%. They simulated the cumulative value of the investment portfolio up until retirement at age 67 and the subsequent variable annuity streams that individuals would receive under the PRA system. They then compared this to the retirement income stream an individual would receive under the current Social Security system. At a 6% contribution rate, the median payment of the combined PRA/variable annuity plan at age 67 would be 2.14 times as large at the payment under the existing Social Security system. There is approximately a 10% chance that the PRA plan would lead to lower payments than the under the existing Social Security system, but only a 1% chance that the payments from the variable annuity would be less than half of the benefits received under the status quo. Based on their evaluation of this simulated distribution of outcomes, FR conclude that

“a pure defined contribution plan, with a saving rate equal to one/third of the long-run projected payroll tax, invested in a 60:40 equity-debt Personal Retirement Account, could provide a retirement annuity that is likely to be substantially more than the benchmark benefit while exposing the retiree to relatively little risk that the annuity will be less than the benchmark. Even this risk can be completely eliminated by a conditional guarantee plan that imposes only a very small risk on future taxpayers.”

There is nothing mysterious in FR’s finding that the PRA plan generates substantially higher benefits from a contribution rate that is only one third as large as under the status quo. This result follows from the simple logic of compound interest. Abstracting from risk, it is easy to
show that whenever there is a sufficiently long contribution horizon, a retirement plan that has a higher rate of return will be far more attractive than one that offers a lower rate of return. The plan with the higher return will generate higher retirement consumption streams for an equivalent contribution (savings or tax) rate, or require a smaller contribution rate to produce the same retirement consumption (benefit) stream, or both. Historically, the rate of return on Social Security contributions has fallen steadily from the extraordinarily high rates paid to the “startup generations” who were alive at the inception of the program in 1935 to less than 2% for the baby boomers and subsequent cohorts. Part of the decline in returns is due to the fact that the startup generations received large net transfers, which will effectively be paid by declining rates of return to subsequent generations. The other reason for the decline is the “demographic transition” resulting from the wave of baby boomers and the downward trend in the total fertility rate to 1.9 children per female.

In steady state, the rate of return to a pay-as-you-go Social Security system is roughly equal to the sum of the growth in population plus the growth in real wage rates. It follows that if the rate of population and real wage growth is high relative to the return on bonds, then a pay-as-you-go system will seem more attractive than a privatized system. However the U.S. economy is now in a situation where the growth rate of the real wage base is significantly lower than the rate of return on government bonds, so a pay-as-you-go system seems unattractive relative to an individual account that earns the government bond rate or higher. This is essentially what FR’s simulations have demonstrated, but with the added wrinkle that rates of return on the PRA plan are both higher than the yield on government bonds and risky. This implies that there is a small chance that realized returns in the PRA plan could be lower than the low asymptotic steady state rates of return earned under the status quo, which is financed mainly on pay-as-you-go basis.

2. **Limitations of the Short Cut Approach to Policy Forecasting**

The FR paper illustrates some of the potential benefits of an individual account approach and as such is a valuable contribution to the literature on Social Security policy analysis. We select this article for critical evaluation because as a prototype “short cut” method it is also useful for illustrating some of the limitations of this approach.

**Absence of Objective Criteria For Evaluating Welfare Gains and Losses.**

The first limitation is the subjective nature of the evaluation of the outcomes. Other analysts may look at the same distribution of outcomes and conclude that the PRA system imposes too much risk on investors. In particular, risk-averse individuals may be willing to pay large sums of money to insure against low probability outcomes. Even though the probability that benefit streams under the PRA system will be substantially smaller than the status quo is fairly small, sufficiently risk-averse individuals might prefer the status quo to the PRA system.

Although FR state that “we believe displaying the probability distributions of possible outcomes … is the best way to indicate the risks and rewards of alternative investment-based options”, they do include a comparison of the expected discounted utilities for an individual at age
The discounted utility of an individual at reference age \( a = 21 \) or \( a = 67 \) is given by:

\[
V_a = E \left\{ \sum_{t=a}^{T} \beta^{(t-a)} u \left( \tilde{c}_t \right) \right\},
\]

where \( T \) is the random age of death, and \( \tilde{C}_t \) is the random realized consumption at age \( t \).

Ignoring the utility of leisure, the endogenous labor supply decision, and other private saving decisions simplifies their analysis considerably. This allows FR to evaluate discounted utility by simply averaging the realized discounted utilities for different outcomes, using the realized incomes, taxes, and benefit levels from their 10,000 simulations. That is, they approximate \( V_a \) by the sample average \( \hat{V}_a \) given by

\[
\hat{V}_a = \frac{1}{N} \sum_{i=1}^{N} \sum_{t=a}^{\tilde{T}_i} \beta^{(t-a)} u(\tilde{c}_{it}),
\]

where \( N=10,000 \) is the number of simulations, \( \tilde{T}_i \) is the date of death in simulation \( i \), and \( \{\tilde{c}_{it}\} \) is the simulated path of consumption in simulation \( i \).

The problem with assuming the arbitrary utility function \( u(c) \) is that different individuals may rank the PRA as more or less desirable than the status quo depending on their degree of risk aversion. FR used constant relative risk aversion utility functions, \( u(c) = \frac{(c^\rho - 1)}{\rho} \) and tried various values of the coefficient of risk aversion \( \rho \) and the discount factor \( \beta \). Not surprisingly, FR found that sufficiently highly risk averse individuals prefer the status quo to the PRA. However, their calculations suggest an implausibly high degree of risk aversion is required in order for an individual to prefer the status quo to the PRA plan. There is no need to make subjective judgements about what level of risk aversion is empirically “implausible” because the appropriate functional form of the utility function and the degree of risk aversion is essentially an empirical issue that can be resolved using econometric methods. However, this approach leads us towards the more involved utility-based methods that we will describe in the next section.

Note that the use of empirically estimated utility functions lessens but does not obviate the need for subjective judgements in policy evaluation exercises because most policies have distributional consequences for which subjective interpersonal utility comparisons are required. Thus, the problem of subjectivity is not unique to the simulation-based approach to policy evaluation, although it is fair to say that the nature of the approach requires the analyst to make a greater number of subjective judgements about the “desirability” of simulated outcomes than most of the approaches described in the next section.

**Uni-dimensional Analysis of Risk.**

A second limitation of the FR study is that it is uni-dimensional. It considers only one risk facing an individual, stock market risk. FR do not consider the many other risks individuals face,
including the risk of health problems and job loss during their working years that interrupts their ability to contribute to their PRA. If they included these risks, the PRA plan would not appear as attractive as it presently appears. The interruptions in contributions would effectively reduce the realized return on the PRA relative to the status quo Social Security benefit, which is relatively insensitive to temporary reductions in earnings because the 5 lowest years of earnings are dropped in computing the AIME. These risk-sharing benefits of the current Social Security system can be substantial and are captured in the stochastic general equilibrium models discussed below in section V of this chapter.

It is unclear how to extend the FR study to incorporate these extra risks, absent a reliable probability model of labor force participation and earnings over the life cycle. Stochastic models of labor force earnings can be estimated via a variety of parametric, non-parametric classical and Bayesian reduced-form approaches, see, e.g. Geweke and Keane (1997), Hirano (1997), and Storesletten, Telmer, and Yaron (1999). Indeed, there a variety of new modeling efforts such as the MINT model and other microsimulation models (e.g. the LIFEPATHS mode of Statistics Canada, see Wolfson, 1995) that allow one to use reduced-form statistical methods to simulate individual-level stochastic processes governing fertility, marriage, divorce, labor force participation and earnings, retirement, and death to obtain distributions of multiple outcomes of interest under different policy scenarios. So, in principle, the stochastic simulation approach appears to be quite flexible in allowing evaluation of multidimensional sets of outcomes.

However, there are a number of important drawbacks to the approach of simulating multiple outcomes and risks simultaneously. The first drawback is that as the number of outcomes of interest increases (e.g. incomes, labor supply levels, marital states, and so forth in various years) the policy analyst is forced to evaluate potentially complicated multidimensional distributions of outcomes in order to make subjective comparisons between alternative policies. Although many people might agree with FR’s claim of the self-evident superiority of the distribution of outcomes from the PRA plan relative to the status quo, one can imagine other more complicated cases where it is very difficult to achieve any consensus in subjective rankings of different policies based on a mass of simulation output. Even if individuals can come up with their own rankings, there might be considerable disagreement resulting from differences in the ways various analysts subjectively evaluate this simulation output. A second drawback of the multidimensional simulation approach is that many policies change the stochastic processes governing the simulated outcomes due to endogenous behavioral responses, a result of a problem known in the econometrics literature as the “Marschak-Lucas critique” that will be discussed in the next section. This problem would occur even if the FR model were expanded to include more realistic stochastic processes. It is easy to illustrate how the Marschak-Lucas problem could arise in the FR model: if high equity returns lead to substantial accumulations in the PRA accounts in the FR analysis, individuals may be tempted to retire well before the age 67, reducing the ultimate size of the realized retirement accumulations, unless this reduction is offset by increased savings earlier in an individual’s working career as suggested by Feldstein (1974). It is clear that allowing for endogenous determination of retirement age could substantially affect the results of their analysis, but there is no obvious way to determine a “hazard rate” for early retirement in this framework without solving the individual’s dynamic programming problem to obtain optimal labor supply and saving. Thus, for many interesting policies, there may not be any
reliable way to evaluate the policy without introducing a utility function and solving the individual’s DP problem. Section 5.2 discusses DP models of the joint labor/leisure and consumption/savings decision.

**Treatment of Unfunded Liability of Social Security System.**

A misleading aspect of the FR analysis is that it does not account for the transition costs involved in moving from the status quo to the PRA plan. As most analysts recognize, there is an implicit debt (unfunded pension liability) to current beneficiaries and workers who paid Social Security taxes and expect to receive benefits under the status quo policy. If we reject the possibility that Social Security would renege on its obligation to current and future retirees, then a proper accounting of the risks and returns of moving to a privatized system like the PRA must consider how these transition costs will be borne by current and future taxpayers. If these costs were subtracted from the higher returns earned by the PRA accounts, the FR analysis would result in a less optimistic conclusion about the gains to privatization. Subsequent papers by Geanakoplos, Mitchell and Zeldes (GMZ) and responses by Feldstein and co-authors address this key issue. Once transition costs are recognized, the differences then come down to what one assumes about how a privatized system would affect saving, what is the appropriate discount rate for risky assets, and what portfolio adjustments would individuals make to offset increased equity holdings in PRAs.

Geanakoplos, Mitchell and Zeldes (GMZ) (1998a) conclude that there is no gain to privatizing Social Security if individual accounts are invested in riskless securities such as government bonds. Even if the rate of return on government bonds is higher than the real internal rate of return in a pay-as-you-go Social Security system, the interest on “recognition bonds” (the expected present value of unfunded liabilities to current beneficiaries and contributors), completely offset the higher returns earned on the individual accounts. The present value of unfunded liabilities is substantial, in excess of $10 trillion dollars, or about twice the size of the explicitly recognized part of the national debt.\(^\text{30}\)

This estimate by GMZ was taken from a previous study by Leimer (1994) that showed that the internal rate of return of the current Social Security system was initially very high (in excess of 20%) for cohorts born prior to 1896 who would have been eligible for benefits shortly after the start up for Social Security in 1935. These individuals received these high returns as a “startup windfall” at the initiation of the Social Security system. The first beneficiaries received a stream of Social Security benefits even though they paid little or no Social Security taxes. Subsequent cohorts received similar windfalls because both the tax and benefit levels of the program continued to increase until the 1970s, so that benefits for these cohorts were high relative to the taxes they paid while working. Leimer’s results show that all birth cohorts up to 1936 received net transfers in the sense that the present value of the cohort’s Social Security benefits exceeded the present value of its Social Security contributions.

---

\(^{30}\) The $10 trillion figure is derived from Leimer’s study. Other studies have arrived at lower estimates of the unfunded liability. For example, Duggan et. al. 1993 estimated that the net transfer to cohorts born from 1895 to 1992 was equal to $3.5 trillion in 1988 dollars.
Even though benefits and taxes may diverge on a year to year basis, Social Security must 
obeys a long run budget constraint that requires the sum of the present values of all net transfers to 
al past and future generations to equal zero. This long run budget constraint implies that the $10 trillion in net transfers to the “startup generations” born prior to 1936 must be by paid for by net 
taxes on generations born after 1936.

“Privatizing the system would allow households to earn market returns on 
their Social Security contributions. But if those contributions were taxed just 

enough to keep the unfunded liability a constant fraction of GDP, then in the 
absence of uncertainty, extra market returns would be entirely dissipated. Future 
workers would not perceive any higher net investment returns than they would 
under the current Social Security system, because for every dollar of Social 
Security contributions, they would have to pay 29 cents in tax to meet the unfunded 
liability inherited from transfers from previous generations.” (p. 45)

GMZ’s calculations, however, fail to value the additional output resulting from the higher 
long run capital stock that would result from moving from a pay-as-you-go to a privatized system. 
As we will see in section V.3, the welfare gains resulting from higher savings and a higher capital 
stock in a fully-funded or privatized system can be substantial, although there may be a temporary 
reduction in welfare for the “transition generations” who are saddled with the dual burden of 
paying off the unfunded liabilities of the transition generations and saving for their own retirement 
at the same time. GMZ’s conclusions also conflict with Feldstein (1998), who used similar 
calculations to derive the gains to privatizing Social Security that explicitly account for transition 
costs. Feldstein’s formula for the present value of the gains to privatizing, PVG, is given by 

\[
\text{PVG} = U \left[ \frac{r - \gamma}{\delta - \gamma} - \frac{r}{\delta} \right],
\]

where \(U\) is the expected discounted value of unfunded Social Security liabilities, \(r\) is the riskless 
rate of interest (the rate of return on government bonds), \(\gamma\) is the growth rate of the economy (the 
sum of the population growth rate plus the rate of growth of real wages), and \(\delta\) is the “social 
discount rate”. Notice that \(\text{PVG} > 0\) only if three conditions hold, 1) \(r > \gamma\), 2) \(r > \delta\) and 3) \(\gamma > 0\).

If we assume that the appropriate social discount rate is the risk free rate of interest, then \(r = \delta\), 
\(\text{PVG} = 0\), and Feldstein’s results are consistent with GMZ’s conclusion that there are no 
gains to privatizing Social Security. Feldstein, however, uses a variety of arguments to suggest 
that the appropriate social discount rate is less than the risk-free rate of interest. “If 
considerations of risk are ignored, this implies that \(\delta = (1 - \tau) r_N\) where \(\tau\) is the marginal individual 
tax rate and \(r_N\) is the return after individual taxes but before individual taxes. With a relatively 
conservative estimate of \(\tau = .2\) and with \(r_N = 0.055\), this approach implies \(\delta = 0.044\). If the real net 
return on government bonds is regarded as a more appropriate risk adjusted measure, \(\delta = r_{GN} = 0.010\).” (p. 25-26). Feldstein states that “experience in the U.S. since 1960 implies that \(\gamma = 0.026\) 
and \(r = 0.093\)”, but he also suggests that “the certainty equivalent rate of return that replaces the 
return to portfolio investors with the yield on government bonds is 6.4 percent.” (p. 25).
Feldstein’s real economic growth rates and rates of return are much higher than those considered in GMZ, which are closer to values used by the Social Security Actuary. GMZ assume that $\gamma = 0.012$ and $r = 0.023$. GMZ specifically criticize the approach of using high expected rates of return for projecting benefits of a privatization plan, but discounting the expected benefits at a low risk-free rate: “calculations using an appropriate risky discount rate deviate considerably from those found in the literature, since all benefit payouts that depend on stock market returns must be discounted by a higher rate than the Treasury risk-free rate. In general, a money’s worth measure that assumes high returns on stock but discounts by a risk-free rate will overstate the benefit of Social Security diversification, and the overstatement will be greater the more stock assumed held in the Social Security accounts.” (p. 59).

Regardless of whether a risk-free or risk-adjusted discount rate is used, Feldstein finds that $r > \delta$ and $r > \gamma > 0$, which implies that $\text{PVG} > 0$. The higher $r$ is relative to $\delta$, the higher the gains to privatization. Thus, the simple logic of compound interest suggests that the gains to privatization will be even higher if one can invest in equities and take advantage of the equity premium, i.e. the excess in the expected return to equities relative to stocks. However, as usual there is a tradeoff between risk and return that needs to be evaluated. This has been done by Feldstein, Rangelova, and Samwick (FRS) (1999), by extending the analysis in FR (1998) to account for transition costs in addition to uncertainty in returns. The analysis is very similar to FR (1998), but also borrows from the approach of Feldstein and Samwick (FS) (1998) which accounted for transition costs but abstracted from uncertainty. FRS find that accounting for transition costs does not radically affect the conclusions reached by either FR or FS. In particular, they find that “transitions to either a completely investment-based system or a mixed system that maintains current law benefits can be done with little additional saving in the early years (a maximum of three percent) and substantially lower combinations of taxes and savings deposits in the later years. The extra risk to retirees and/or taxpayers is relatively small, making investment-based plans preferable to a pure pay-as-you-go system for reasonable degrees of risk aversion.”

GMZ also extend their analysis to allow for risky stock market investments that yield higher returns than the risk free rate. However, they do not find the advantages to privatizing Social Security and investing in equities to be as clear-cut as FRS. The reason is that “if all households have access to stock investments of their own, then permitting equities in the individual accounts will have no effect on anyone’s well-being. In effect, the stocks that workers buy for their individual Social Security accounts will be purchased from their own private portfolios. Their overall portfolios will end up absolutely unchanged.” (p. 58). GMZ recognize, however, that by helping to “prefund” Social Security, privatization is likely to increase the long run capital stock and economic well-being. Further, the hypothesis that all households have access to stocks is empirically false: the most recent available data indicate that only 50% of all U.S. households have any equity holdings whatsoever, even indirectly via defined contribution pension holdings.\footnote{This situation is changing quickly, however. The huge stock market returns in recent years coupled with sharp decrease in costs in stock trading via the WWW and “electronic communication networks” has lead to a rapid rise in popularity of “day trading” by small investors that is rapidly making participation in the stock market more of a...} If there are a significant fraction of “constrained” households, then privatization could have a beneficial impact since
“the fundamental rationale for Social Security investment in the stock market rests on the existence of people who are currently constrained from holding equities. It is interesting to note that those who would benefit the most from Social Security investments in equity are probably the poor, since this group is least likely to hold stocks now. Quantifying the money’s worth to a constrained household of a large movement into the stock market is more difficult. It is clear, however, that a discount rate higher than the risk free rate should be used in computing the money’s worth of the stock payoffs. The reason is that as households gain more exposure to stocks, they would perceive their old age income as more and more at risk.” (p. 32).

3. Using Options to Calculate the Value of Benefit Guarantees

A corollary to the problem of failing to account fully for gains and losses in policy simulation models is what Bazelon and Smetters (1999) have termed the problem of “super implicit debt”. An example is the implicit cost of minimum benefit guarantees to Social Security beneficiaries. For example, in considering moving to a privatized system with individual accounts that can be invested in equities, such as the PRA plan considered in FR, it is easy to see that if the government offers a guarantee that benefits will not fall below a specific minimum benefit level, then future taxpayers necessarily bear the risk of paying for this guarantee in the event of poor stock market performance. To be fair, FR did consider the impact of the implicit minimum benefit guarantees in their simulations by computing the value of a social welfare function, a weighted sum of the expected discounted utilities for 80 birth cohorts who are between ages 21 to 100 in a steady state after the PRA plan is fully phased in. A limitation of this calculation is that it does not consider the endogenous long run changes in saving and labor supply that result from such a change, but accounting for these feedbacks moves us toward a more complicated utility-based, general equilibrium analysis that we will consider in the next section. This section considers another shortcut approach to dealing with the problem of “super implicit debt” inherent in minimum benefit guarantees—the option-pricing approach.

The option pricing approach is discussed in Smetters (1999). The discussion below also borrows from an example presented in Abel and Lucas (1999). The basic idea is to use a well developed theory of option pricing in finance to calculate the value of minimum benefit guarantees implicit in many Social Security reform proposals, and then use these values to calculate an adjusted cost of the program. A call option gives the holder the right to buy a stock (or a stock index) at a predefined exercise price. A put option gives the holder the right to sell a stock at a predefined exercise price. If the options can only be exercised at a fixed exercise date, then at this date the value of the call $C$ and the put $P$ can be written in terms of the exercise price $E$ and the stock price $S$ as:

$$C = \max[S - E, 0]$$
$$P = \max[E - S, 0].$$

universal phenomenon in the U.S.
Finance theory, particularly the famous Black-Scholes option pricing formula, enables us to calculate the value of these options at any time prior to the exercise date. This formula was derived from a no-arbitrage condition that the price of the option should be equal to the value of a hedge portfolio consisting of stocks and bonds that can exactly duplicate the payoff of the option at the exercise date. Alternatively, Harrison and Kreps (1983) showed that these option prices are simply the conditional expectations of the discounted values of the values of the options at the exercise date, where the expectation is taken with respect to an equivalent risk-neutral martingale measure. This representation is useful since it enables the use of Monte Carlo simulation methods to estimate the option values, see, e.g., Boyle, Broadie and Glasserman, (1999).

Options can be used to insure against stock price fluctuations and guarantee minimum rates of return. For example, to guarantee a certain end of period payoff of \( E \), an investor could form a portfolio consisting of a long position in the stock and one put together with a short position in one call. This portfolio would then have a value at the exercise date of

\[
E = S + \max(E - S, 0) - \max(S - E, 0)
\]

with probability 1. The cost, or “insurance premium”, for this guarantee equals the value of the put less the value of the call (which was sold short). Suppose, for example, that you wanted to invest $100 in the stock market and obtain a certain return of 12% after one year. You would buy the stock for $100, and then buy a put option with an exercise price of \( E = 112 \) and sell a call option with an exercise price of $112, both with an exercise date one year in the future. In this case, the premium or cost of the guarantee is very easy to determine since it is equivalent to a risk-free bond. Assuming the risk-free rate of return is 6%, the present value of the payoff from this portfolio is \( \frac{112}{1.06} = 105.66 \). Thus, the implicit cost of the “insurance” that guarantees a certain return of 12% is $5.66.

Many policy discussions center on providing minimum benefit guarantees, such as a system of individual accounts where a beneficiary gets the higher of the value of the individual account which is invested in stocks, call this \( S \), or the benefit level under current law, call this \( E \). Then the benefit payable upon retirement is given by

\[
\max(E, S) = S + \max(E - S, 0)
\]

Thus, this benefit guarantee can be viewed as a portfolio consisting of a stock \( S \) and a put option with exercise price \( E \). The value of the guarantee then equals the market valuation of the put option.

In some cases, market prices of put options can be directly used to value this minimum benefit guarantee. Otherwise the option pricing formulas can be used to compute the value of the put option, basically by computing the expected discounted value of the put option at exercise date using the appropriate “risk neutral” or equivalent martingale measure for the security \( S \). Smetters (1999) discusses these details and provides other examples of how option pricing methods can be used to value the implicit insurance provided by a number of other benefit guarantees. The basic idea in all cases is simply to charge taxpayers the estimated “market price”
of these guarantees, even though in practice the risk is not explicitly insured with options, but is almost always “self-insured” via uncertain changes in future tax rates.

Abel and Lucas (1999) used this approach to evaluate the implicit cost to taxpayers of the Clinton Administration’s recent proposal to invest 20% of the Social Security Trust Fund in equities, while maintaining a guarantee that current Social Security benefits would remain unchanged. Assuming an expected return of 9% on equities, they calculated that the aggregate value of the benefit guarantee was $254 billion. Smetters (1999) estimates that the Clinton Administration proposal is actuarially equivalent to increasing the payroll tax in 30 years by 0.8 percentage points in perpetuity, from the current 12.4% to 13.2%. Investing 40% of the Trust Fund in stocks is equivalent to a 2.1 percentage point rise in future tax rates.

The option pricing approach is an attractive way to estimate the value of the “super implicit debt” created by various benefit guarantees, but it has a number of important limitations. Since the government does not explicitly purchase this insurance by hedging in the stock market, but rather “self insures” this risk via stochastic adjustments to future tax rates, the calculated cost of the guarantee could overestimate or underestimate an individual’s certainty equivalent valuation of this risk. Stated differently, many risk averse taxpayers would find the “insurance premium” necessary to insure a given percentage return to be too high, and will decide to be long in stock without incurring the cost to hedge this risk via purchases and sales of puts and calls. For these investors the insurance premium is higher than their certainty equivalent valuation of the risk. Other investors would find the insurance premium to be relatively cheap since their certainty equivalent valuation of the risk is higher than the implicit insurance premium. Therefore ranking Social Security reform proposals via money’s worth measures that adjusted the expected present values of taxes for the estimated market cost of implicit benefit guarantees may not necessarily provide a correct social welfare ranking of the alternative proposals. Also, many proposals are rather vague about whether there are any minimum benefit guarantees. In practice, most plans will lead to uncertain changes to future benefits and taxes and there is no clear way to value these risky changes in the Social Security rules via the standard option pricing formulas used to value simple put and call options. In these more complicated and realistic situations, one may have no choice but to attempt to evaluate the risk explicitly via direct calculations of expected utility that account for the subjective probability distribution of uncertain future changes to the Social Security system. These calculations are not easy to do, but examples will be provided in section V below.

4. Pitfalls of Using Money’s Worth Measures and Other Summary Statistics

This section discusses other ways of summarizing the costs and benefits of different policies via relatively low dimensional “sufficient statistics” that are designed to capture the main welfare and behavioral impacts of various policies. A standard way of comparing alternative random sequences of cash flows that occur over a sequence of years is the internal rate of return. This is the discount rate that equates the expected present value of positive cash flows (benefits)

52 Their calculations presume that the Trust Fund is continually rebalanced to ensure that equities never constitute more than 20% of the Trust Fund.
to the expected present discounted value of negative cash flows (costs, taxes). In the case of Social Security, we can evaluate the internal rate of return of the existing Social Security system by finding the discount rate that equates the expected present value of Social Security tax contributions to the expected present value of benefit payments. A large number of studies have carried out such calculations in recent years, including the recent study by Beach and Davis (1998a) that was discussed in section III. In addition to the criticisms already noted, calculated internal rates of return do a poor job of accounting for risk because any policy that produces a mean-preserving spread on Social Security benefits and taxes will not affect the internal rate of return. However, it should be clear that mean-preserving spreads of taxes and benefits make risk averse individuals strictly worse off.

This shortcoming can be addressed by computing the expected net present value (NPV) of Social Security wealth, which equals the expected discounted value of benefits less contributions using a risk-adjusted discount rate. However, this approach has its own problems because it is often difficult and somewhat arbitrary to determine what the risk adjusted discount rate should be. Do we use the capital asset pricing model to compute the risk-adjusted rate, or try to define different risk adjusted discount rates for different individuals that reflect their varying levels of risk aversion? How do we adjust for idiosyncratic risks such as mortality, health problems, and job loss, etc.?

Despite these problems, IRRs, NPVs and other related “money’s worth” measures have played a prominent role in the U.S. Social Security reform debate. For example, the Social Security Advisory Council (1997) scored three reform plans using money’s worth measures and concluded that all three plans had higher money’s worth than the status quo, with the implicit suggestion that most individuals would find these plans preferable to the status quo. Geanakoplos, Mitchell, and Zeldes (GMZ) 1998 provide a thorough analysis of some of the pitfalls that can arise from naive analyses using money’s worth measures, with particular attention to the issue of whether it is possible to increase the returns to Social Security by increasing the Trust Fund or moving to a privatized system with individual accounts. They show that ranking alternative reform proposals based on a simple money’s worth measure such as the IRR or NPV will generally not yield a correct welfare ranking: “In the most general situation, then, accurate money’s worth numbers require a detailed knowledge of each household’s preferences including its attitude towards risk. Since these preferences are generally not known, it might seem hopeless to derive a useful set of summary statistics that can be used for comparing different Social Security systems.” (p. 23). However, they show that there are conditions under which rankings based on money’s worth do yield valid welfare rankings: i.e., conditions under which all individuals would agree that a Social Security plan with a higher money’s worth is preferable to one with a lower money’s worth.

Unfortunately the conditions for this ranking to be valid are too restrictive to be relevant in practice. Indeed, their key “spanning” assumption (households can duplicate all Social Security income streams via securities traded in the marketplace) implies that there is no useful role for Social Security beyond providing a system of taxes and transfers that redistributes lifetime income. When spanning fails there are incomplete markets that Social Security can help to “complete”. That is, when private markets don’t offer individuals the opportunity to design their
own preferred retirement, investment and insurance plans, the analysis of Eckstein, Eichenbaum
and Peled (1986) suggests that Social Security can lead to welfare improvements by helping to fill
in these missing markets. Actually, a safer statement is that Social Security can cause changes in
welfare by helping to fill in missing markets, because an example discussed in section V.3 shows
that introducing an actuarially fair annuities market can actually reduce welfare in a plausibly
specified model. Examples of incomplete markets where Social Security is likely to have a strong
effect include markets for annuities, disability, and health insurance. Unfortunately, money’s
worth measures do not provide reliable measures for the practical case where markets are
incomplete. GMZ conclude that “money’s worth measures are well-suited to comparing benefits
and costs for different income groups, and even different cohorts, under the same Social Security
system. But the typical approaches do not fare well at making comparisons across different reform
plans, primarily because they do not properly account for differences in risk and/or transition
costs.” (p. 46).

In section V.3, we provide an example where a comparison of internal rates of return to
Social Security contributions provides a misleading policy conclusion even when transition costs
are accounted for. This example provides a stark illustration of the fact that at the aggregate level,
a policy change (in this case, the creation of an actuarially fair annuities market) that leads to a
higher IRR on Social Security contributions may actually result in a lower level of welfare for
individuals in the economy.

5. Concluding Remarks about the Short Cut Approach

The main advantage of the simulation-based approach is that once SSA has a reliable
probability simulation model, it is relatively easy to conduct stochastic simulations to study the
implied distributions over a variety of different outcomes of interest. Stochastic simulation per se
does not impose huge computational burdens. The approach is quite straightforward and could
be implemented by SSA relatively quickly. The simulations can distinguish between aggregate
shocks, which are common to all individuals (e.g. stock prices, stochastic changes in tax rates,
unemployment rates, etc.), and idiosyncratic shocks. A single simulation would provide one
realization of the aggregate shocks and \(N\) realizations of the idiosyncratic shocks, where \(N\) is the
number of individuals included in the simulation. The individual-level simulations are typically
more time consuming to perform since there are \(N\) of them where \(N\) is typically a large number,
and because each individual’s simulation requires sequential updating of many associated
variables such as their earnings histories, benefit and asset levels, and so forth. However, these
updates generally involve only trivial calculations, so that even the individual level simulations can
be carried out extremely rapidly. A single simulation of a population consisting of \(N=1,000,000\)
for 80 periods that requires updating of 20 individual level state variables would take well under a
minute to run on modern desktop workstations that are capable of speeds of performing 50
million floating point operations per second or higher. Given a sufficiently large number of
simulated individuals, the law of large numbers guarantees that in the aggregate the idiosyncratic
shocks average out. Thus, current technology makes it possible to produce simulated data sets
rapidly, from which population moments and distributions of individual outcomes can be easily
tabulated. Even with a million simulated individuals, the population moments may be fairly noisy
estimates in cases where analysts are interested in extremely small segments of the overall
population or certain rare events (e.g. the number of young individuals entering the DI program due to strokes). Although this problem can always be solved by increasing the size of the simulated population, in many cases, the problem can also be solved by conducting simulations that condition on particular population subgroups of interest, and then reweighting the estimated simulated moments for each of these subpopulations by their share in the overall population.

Note, however, that the approach described above appeals to the law of large numbers to average out idiosyncratic variation conditional on a particular realization of the macro shocks. If we also want to tabulate the distribution of possible aggregate outcomes we need to make multiple simulations of the entire population for independently drawn realizations of the macro shocks. For these types of analyses, the cpu time necessary to conduct simulations can start to become significant, since even though a single simulation of a population of a million individuals may take less than a minute, many thousands of such simulations would be required in order to estimate accurately the overall joint distribution with both micro and macro shocks. One way to deal with this is to reduce the size of the simulated population, tolerating some additional imprecision in the estimates of the distribution of outcomes resulting from idiosyncratic shocks in order to get more precise estimates of the distribution of outcomes resulting from the macro shocks. Note that simulations are highly parallelizable, so that networked clusters of workstations could perform many simulations of large populations with different processors being assigned to simulate a population with a different realized sets of macro shocks. Overall, the cpu time needed to conduct simulations will probably not be a significant obstacle for most of the analyses that are done under the short cut approach.

Aside from the practical problems involved in subjectively assessing the desirability of different probability distributions for various outcomes implied by different policies, the key prerequisite for this approach is a probability model that accurately simulates individual behavior under the status quo and a range of hypothetical policies of interest. This prerequisite is the “Achilles heel” of the short cut approach. As discussed in the next section of this chapter, it is relatively easy to specify reduced-form probability models that can accurately simulate (observed) individual behavior under some historical policy regime for which there is sufficient data to reliably estimate observed behavioral relationships. However, the Marschak-Lucas critique leads to the presumption that reduced-form models will not lead to accurate predictions of policy changes, especially those for which haven’t occurred in the past. If we want a fully behaviorally justified probability model that is more likely to predict policy changes accurately, we have few options except to work within a “closed” model that endogenously derives how policies affect the probability distributions affecting individual behavior. These include the structural DP or SGE approaches that will be described in the next section. The only other possibility is to assume that policy-induced changes to the probability distributions governing individual behavior are determined by some undefined, ad hoc process reflecting the analyst’s “judgement” of how individuals will respond to various policies. In certain cases, such as Feldstein and Rangelova’s (1998) analysis of personal retirement accounts, policy changes can be modeled as changes in certain exogenously specified probability distributions, with endogenous behavioral and general equilibrium responses ignored. However, the discussion in this chapter suggests that the assumption that there are no endogenous responses to policy changes can be very difficult to justify.
V. THE DP/SGE APPROACH TO INCORPORATING RISK AND INSURANCE IN POLICY MODELS

The main difference between the short cut approaches discussed in section 4 and the dynamic programming (DP) and stochastic general equilibrium (SGE) models reviewed in this section is that the latter models endogenously derive the behavioral and general equilibrium responses to changes in Social Security policies. To do this, we need to specify the objectives of individuals, firms and the government, the risks and constraints they face, and the actions they can take to maximize their preferences subject to these risks and constraints. The actions available to individuals include labor supply and saving, the purchase of private insurance contracts, and applications for various types of Social Security benefits. There is a large, rapidly growing literature on these models. Collectively, the models in this literature incorporate nearly all of the individual-level risks that section III of this chapter listed as being relevant for policy analysis, as well as most of the various private, governmental, and non-governmental institutions and mechanisms individuals have at their disposal to deal with these risks. Although some of these academic models could be used almost immediately with little or no modification for use in policy forecasting within SSA, most of the models considered in this section should be viewed as prototypes for models that SSA might consider as part of a longer term strategy for policy modeling. We divide this review of the literature on DP/SGE models into three parts: 1) dynamic programming models, used to model individual decision making under risk and uncertainty, 2) a brief review of the literature on mechanism design and the design of optimal social insurance systems, and 3) stochastic overlapping generations models used to model general equilibrium impact of social insurance on the overall economy.

1. Dynamic Programming Models

The method of dynamic programming allows us to model optimal sequential decisions under uncertainty, and constitutes the foundation on which virtually the entire literature reviewed in this section is built. Unfortunately, only relatively specialized, highly simplified dynamic programming problems admit analytic closed-form solutions. But the use of computers has allowed economists to solve substantially more realistic models. It has also spawned a literature on econometric estimation and testing, where the unknown parameters of the DP models are chosen so that the predictions of the DP model best fit observed behavior. This literature allows us to test the validity of strong rationality hypotheses underlying DP models.

This section briefly reviews the method of dynamic programming, and discusses the literature on numerical solution and econometric estimation of DP models, with particular attention to the pros and cons of the structural DP approach to econometric estimation of behavioral relationships, compared with simpler reduced-form estimation methods that do not rely on any particular theory of individual decision-making. It provides examples that show that even relatively simplified numerical DP models succeed in providing very good empirical models of individual behavior, and provide more accurate predictions of individuals’ endogenous responses to policy changes than alternative reduced-form econometric models. It discusses the Marschak-Lucas critique that provides a theoretical presupposition for why this should be the case. The
The term “dynamic programming” (DP), introduced by Richard Bellman (1957), refers to a recursive approach for solving for optimal decision rules that constitutes the standard approach to solving these problems by both analytic and numerical methods. However, it has proved most useful as a basis for numerical methods that solve DP problems on digital computers (see, e.g. Rust, 1996). Computers allow us to formulate more detailed and realistic models of individual behavior because we are freed from the restriction of having to specify simple functional forms that permit closed-form solutions to the individual’s optimization problem. The advent of sufficiently powerful hardware and software has enabled us to solve trial versions of the DP problem repeatedly in a process of searching for the unknown parameters of individuals’ beliefs and preferences that make the predicted behavior from the DP model “best fit” observed behavior, using various measures of goodness of fit. Furthermore, DP models can be used to derive optimal decision rules associated with alternative Social Security policies. These optimal rules constitute a precise quantitative prediction of an individual’s “best response” to changes in policy. It is also easy to use DP models to compute “certainty equivalent” valuations of a policy change, i.e. the fraction of the individual’s wealth or income flow that they would need to pay (or be paid) to make them indifferent between the new policy and the status quo. This enables us to study the welfare and distributional consequences of policies in a relatively objective manner, quantifying the individual-level impacts in dollars rather than in “utils”.

For concreteness, this paper describes the DP method in the context of discrete time Markovian decision theory. The key variables entering this model are a vector of state variables $s_t$ and a vector of control variables $c_t$. Examples of components of $s_t$ include income and hours of work in the previous year, assets, age, health status, family status, and indicators of pension status (eligibility for various types of Social Security and pension benefits). Examples of components of $c_t$ include labor supply, consumption, and indicators of the decision to apply for Social Security or pension benefits. There is no significant loss of generality in modeling the evolution of various uncertain outcomes as a Markov process. Higher orders of correlation in

---

53 Note that, as in Rust and Phelan (1997), having separate state variables that distinguish between labor force status and pension acceptance of individuals obviates the need for arbitrary, subjective classifications of states such as “retirement”. It is entirely possible for someone to apply for Social Security Old Age benefits and continue working, or to quit working and not apply for Social Security, etc. Different people will have different definitions of whether or not they are “retired” in these situations, but by focusing on objectively measurable states and decisions we avoid much of the arbitrariness inherent in the definition of retirement. Furthermore we avoid any presumption that retirement is an “absorbing state”, because it is always a logical possibility to return to work.
various state variables can be accounted by a variety of methods that retain the Markovian formulation at the expense of a higher dimensional state vector. Although it may be more realistic to think of this Markov process as evolving in continuous time, for data and computational reasons we will assume that it is a discrete time process. In practice, annual time intervals are typically assumed, although for some policies (e.g. unemployment policy, or modeling disability application and appeal processes), monthly or even weekly time intervals might be necessary. However, it is hard to imagine that SSA would require a model with finer time intervals for most of its policy forecasting needs.

The DP model is completed by specifying the individual’s beliefs and preferences. Preferences are given by the utility function \( u(s_t,c_t,\theta_1,\alpha) \) where \( \theta_1 \) represents time-invariant parameters of the individual’s utility function (relevant for the discussion of econometric estimation of DP models below) and \( \alpha \) represents a vector of policy parameters that affect preferences. By not indexing \( \alpha \) with a subscript \( t \), we adopt the Stackelberg assumption that the government commits to a fixed policy regime. This can be relaxed at the expense of allowing components of \( s_t \) to represent policy variables representing the state of the government’s policy at time \( t \). Examples of policy parameters include objective quantities such as tax rates and more subjective quantities such as parameters that affect the “stigma” of being in various states such as disability. Beliefs are represented by the transition probability density \( p_t(s_{t+1}|s_t,c_t,\theta_2,\alpha) \) that also depends on a vector of parameters \( \theta_2 \) (also included to motivate discussions of econometric estimation) and policy parameters \( \alpha \). It is easy to provide examples of how policy parameters can affect beliefs. For instance, parameters characterizing the levels and the bend points in the formula relating AIME to PIA affects an individual’s beliefs about their future Social Security benefits. Another example would be parameters describing the conditional probability of being awarded DI benefits, or parameters affecting the distributions of the delays between application and award of benefits.

The method of dynamic programming amounts to calculating the value functions \( \{V_t\} \) via the recursive formula known as Bellman’s equation given by:

\[
V_t(s) = \max_{\{\delta_t\}} \left\{ \bar{T} \sum_{j=0}^{\bar{T}} \beta \int_{s_{t+1}}^{s_t} u_t(s,c,\theta_1,\alpha) \left[ S_t = s, \theta, \alpha \right] \right\} = \max_{c} \left[ u_t(s,c,\theta_1,\alpha) + \beta \phi_t(s,c,\theta_2) \int V_{t+1}(s') p_t(ds'|s,c,\theta_2,\alpha) \right], \tag{5.1} \]

where \( \bar{T} \) is the uncertain date of death of the individual and \( \phi_t(s,d,\theta_2) \) is the conditional probability of survival from age \( t \) to \( t+1 \) given \( (s,d) \).8 The solution consists of the sequence of

\[\text{\footnotesize 54 We omit the policy parameters } \alpha \text{ from } \phi \text{ on the assumption that the government policy does not directly affect mortality. However it is clear that government funded medical research has had a huge indirect effect on improving mortality, and it is not a stretch to think of other policy variables that SSA has control over (e.g. rules over who is eligible for Medicare and disability benefits) that has a much more direct impact on mortality. However, I am not aware of any studies that directly model the impact of policy on mortality.}\]

210
optimal decision rules \( \{ \delta_t \} \) which are the values of the control variables that maximize Bellman’s equation:

\[
\delta_t(s) = \operatorname{argmax}_c \left[ u_t(s,c,\theta_1,\alpha) + \beta \phi_t(s,c,\theta_2) \int V_{t+1}(s') p_t(ds'|s,c,\theta_2,\alpha) \right] \tag{5.2}
\]

The decision rule \( \delta_t \) represents the behavior implied by the DP model and the value function \( V_t \) represents the individual’s welfare. Thus, \( c_t = \delta_t(s_t) \) represents the optimal decision that an individual should take in state \( s_t \) at time \( t \) and \( V(s_t) \) represents the expected discounted utility over their remaining lifetime.

It is possible to turn generic DP models into behavioral econometric models using algorithms that estimate “structural parameters” \( \theta \) characterizing an individual’s beliefs and preferences (see, e.g. Rust, 1994). Examples of structural parameters include an individual’s subjective discount factor, parameters affecting their willingness to trade-off consumption for leisure, parameters characterizing their mortality, etc. The estimation algorithm repeatedly resolves the individual’s DP problem until values of the structural parameters are found that enable the optimal behavior implied by the DP model to “best fit” observed behavior. There are a variety of different metrics of goodness of fit and associated estimation methods, including maximum likelihood. A likelihood function can be derived from the solution to the DP model by assuming that the state variable \( s_t \) is partitioned as \( s_t = (x_t, \varepsilon_t) \) where only \( x_t \) is observed by the econometrician and \( \varepsilon_t \) is treated as an unobservable from the standpoint of the econometrician.

Under certain conditional independence assumptions, one can derive a conditional choice probability \( P_t(c_t|x_t,\theta) \) from the solution to the DP problem by “integrating out” the unobserved state \( \varepsilon_t \):

\[
P_t(c_t|x_t,\theta) = \int I\left\{ c_t = \delta_t(x_t,\varepsilon,\theta) \right\} q(\varepsilon|x_t) d\varepsilon \tag{5.3}
\]

and this conditional choice probability serves as the basis for maximum likelihood estimation of the unknown preference and belief parameters \( \theta \) using the likelihood function \( L(\theta) \) given by

\[
L(\theta) = \prod_{i=1}^{N} \prod_{t=1}^{T_i} P_t(c_{it}|x_{it},\theta) p_t\left(c_{it} \mid x_{i,t-1}, c_{i,t-1}, \theta_2 \right).
\]  

where \( \{ x_{it}, c_{it} \}, t = 0, \ldots, T_i, i = 1, \ldots, N \) represent the observed states and decisions of a sample of \( N \) individuals where individual \( i \) is followed for \( T_i \) periods.

Policy analysis is conceptually straightforward within the DP framework. First we estimate the unknown structural parameters \( \theta \) using observed data \( \{ x_{it}, c_{it} \} \) on individuals under some historical policy regime represented by the policy parameter vector \( \alpha \). Then to predict how behavior and welfare change under an alternative policy \( \alpha' \) we simply re-solve Bellman’s equation with \( \alpha' \) substituted for \( \alpha \) and using the estimates \( \theta' \) for the unknown belief and preference.
parameters. The resulting decision rules \( \{ \delta_i(\alpha') \} \) represent the DP model’s prediction of how behavior will change under the new policy regime \( \alpha' \). The value functions \( \{ V(\cdot|\alpha') \} \) represent the DP model’s predictions of how the individual’s welfare will change under the new policy regime \( \alpha' \). If one of the state variables is wealth or income, then certainty equivalent valuations can be computed by solving for the change in wealth or income necessary to give the individual the same expected discounted utility after the policy change as they expected to receive prior to the policy change. Note that these certainty equivalent valuations will typically depend on the values of other state variables, and will reflect population heterogeneity: different people (as represented by different state vectors) \( s_t \) will be affected in different ways by any given policy change. The validity of these policy forecasts depends on the maintained hypothesis that the structural parameters are invariant to changes in the policy parameters, and on the assumption that an individual’s behavior is a “best response” to any given policy regime. This means that the individual behaves “as if” he or she had solved a dynamic programming problem with full knowledge of the values of all relevant policy parameters.

Clearly, the accuracy and reliability of quantitative predictions of numerical DP models depends on the validity of the key hypothesis underlying these and nearly all analytic models in economics, namely that individuals are rational. There are two aspects to this hypothesis. First, we generally assume that individuals have rational expectations, i.e. they have full knowledge of the transition probability \( p(s_t|s_{t-1}, c_{t-1}, \theta_2, \alpha) \) including all relevant policy parameters. In this context this implies that they are fully informed about the rules of the Social Security program and that their beliefs about uncertain events such as death coincide with objective probability distributions that can be estimated from observed outcomes. Although it is possible to weaken the rational expectations assumption and allow individuals’ subjective beliefs to differ from the objective probability measure governing observable outcomes, identification of such models is extremely difficult. In simple terms, there are generally many different combinations of beliefs and preferences that can rationalize the same observed outcomes, so to identify the model we generally invoke the strong identifying hypothesis of rational expectations. The second aspect of the rationality hypothesis is that individuals behave optimally, i.e. they behave “as if” they were maximizing the expected value of a discounted time separable utility function. Due to the identification problems noted above (see Rust, 1994 for details), the rationality hypothesis can be empirically tested only if one is willing to make some additional identifying assumptions about the functional forms of individuals’ preferences and beliefs.

In summary, if the rationality hypothesis is correct, and the a priori specification of the parametric functional forms for preferences and beliefs are correct, then it is possible to forecast the impact of arbitrary policy changes correctly, even if they have no historical precedent. The impacts that can be analyzed via these “structural” econometric methods include changes in observable behavioral characteristics such as saving, labor supply, and the dates of application for Social Security benefits, as well as changes in non-observable features such as an individual’s welfare. DP models constitute a subclass of structural models that are rich enough to model almost any problem involving individual sequential decision making under uncertainty. However, they are not the only class of models that could be used for this purpose. I briefly review, and discard, two extensions to the DP approach below.
One class of models allows us to distinguish between choices involving risky outcomes from choices of uncertain outcomes, i.e. where the decision-maker is not certain of all the parameters $\theta$ characterizing their preferences $u_t(c,s,\theta_1,\alpha)$ or beliefs $p_t(s'|s,c,\theta_2,\alpha)$. There is a well-developed literature on Bayesian decision theory where individuals have priors over the unknown parameters in the probability distributions governing uncertain outcomes. Individuals rationally update their beliefs via Bayes’ rule, but otherwise this theory is formally isomorphic to standard Markovian decision theory, where agents are assumed to have complete knowledge of all probability distributions. Essentially the only change involved is to use the individual’s posterior distribution in place of a known distribution over random outcomes. Other than substantially complicating the analysis, there is no clear evidence that suggests that use of Bayesian decision theory would significantly affect any major conclusions about Social Security policy that have been obtained in the existing literature. The predictions of DP models may be sensitive to specifications of individuals’ beliefs, but this can usually be determined without resorting to Bayesian decision theory.

For similar reasons, we also ignore the recent literature on non-Bayesian models of choice under uncertainty and choice under risk with non-expected utility and non-time separable specifications of preferences. While there is substantial experimental evidence that suggests that individuals do not behave in accordance with the expected utility model, there is no definitive alternative to the expected utility model that is analytically tractable and that succeeds in resolving a majority of the various experimental anomalies. Expected utility maximization of time separable preferences may appear highly restrictive, but if one has the freedom to choose preferences and beliefs arbitrarily, then it can be shown that virtually any type of behavior can be rationalized within this framework (see, e.g. Rust, 1994). Nearly all existing models of Social Security have adopted the expected utility framework, although a few recent exceptions have solved SGE models where non-expected utility preferences actually lead to a tractable analysis. Gertler (1999) analyzes the effect of Social Security using the following non-time separable, non-expected utility specification of preferences

$$V_t(s) = \left( [u_t(c,s)]^\rho + \beta (EV_{t+1}(s,c))^\rho \right)^{1/\rho}.$$

Note that when $\rho = 1$ these preferences reduce to ordinary time separable, expected utility specification. Huang, Imrohoroglu and Sargent (1999) also analyzed the impact of Social Security using the following time-separable, but non-expected utility function

$$V_t(s) = \left[ u_t(c,s) + \frac{\beta}{\sigma} \log \left( \mathbb{E} \exp \{ \sigma V_{t+1}(s,c) \} \right) \right].$$

Note that when $\sigma = 0$ these preferences reduce to the standard time separable, expected utility specification. If $\sigma > 0$, the individual prefers to delay the resolution of uncertainty. If $\sigma < 0$, the individual prefers early resolution of uncertainty and exhibits a precautionary motive for saving that is distinct from the precautionary savings motive that arises due to nonexistence of
complete insurance markets. This specification of preferences can also generate an increase in the inequality in the distribution of consumption (due to stochastic shocks to earnings) within a cohort of individuals as the cohort ages. There is no clear consensus as to whether the “payoff” in terms of improved ability to account for observed behavior from using the more complicated non-time-separable non-expected utility preferences is higher than the payoff from developing richer models that account for other aspects of Social Security and other types of risks and decisions within the expected utility framework. As we will see, it is already quite challenging to construct realistic models of Social Security that account for the many risks individuals face and the variety of ways for coping with them within the standard expected utility model.

The main drawback of the numerical dynamic programming approach to policy forecasting is the computational burden involved with solving increasingly realistic formulations of DP models. This problem, which Bellman referred to as the curse of dimensionality, will be discussed in more detail below. It is useful to consider the pros and cons of a much less computationally intensive alternative approach to policy forecasting, namely to extrapolate historical trends using a variety of ad hoc models estimated from a combination of individual-level panel data, cross-sectional data, and aggregate time series data. This is the approach that would generally be used to construct the probability simulation models underlying “short cut” approaches to policy evaluation, as discussed in section IV of this chapter. Ad hoc models have the advantage of being relatively simple to use and can be flexibly parameterized to approximate a wide variety of stochastic processes. Somewhat more computationally intensive non-parametric approaches can also be used to estimate these models. As a result, ad hoc models do not require as many strong maintained behavioral assumptions as numerical dynamic programming models.

The validity of forecasts from ad hoc models does, however, depend on a number of very restrictive assumptions, including stationarity. A sufficiently large change in the policy regime could significantly alter behavioral relationships, which could invalidate the predictions of ad hoc models based on historical estimates of behavioral relationships estimated under the previous policy regime. This is the essence of the well-known “Marschak-Lucas critique” (Marschak, 1953, Lucas, 1973). This critique amounts to the simple observation that the stochastic process generating the data prior to the policy change is generally different than the stochastic process generating the data after the policy change. This is the reason why forecasts of ad hoc or non-utility based statistical methods generally cannot be relied upon to generate accurate forecasts of the effects of policy changes. In some cases, policy changes are sufficiently frequent that they can be treated as “regressors” within an ad hoc framework, from which reliable predictions can be generated. However, in the case of Social Security, most policy changes are infrequent or have no previous analogs, so the historical evidence is generally insufficient to enable one to predict how individual behavior and the resulting stochastic process generating the data will change.

Of course, in most cases we view models as abstractions that are at best crude approximations to reality. The Marschak-Lucas critique ignores the possibility that a structural model could be misspecified. As discussed in section IV, it is not obvious whether a misspecified structural model will necessarily forecast a policy change better than an ad hoc statistical model. Thus, the problem reduces to a practical question of whether one obtains superior policy forecasts from 1) flexible statistical models that place fewer restrictions on equation specification in order
to obtain better fits with observed data (ad hoc models) or 2) highly simplified structural models that impose considerable a priori assumptions about rationality and optimization.

In some cases, including the problem of predicting how retirement behavior is affected by changes in retirement pension incentives, relatively simple parametric structural models have provided substantially more accurate forecasts than ad hoc reduced-form models. For examples see Rust, (1994) and Lumsdaine, Stock and Wise (1993).

2. Information Economics, Mechanism Design, and Models of Optimal Social Insurance

One of the most important insights in the economic literature on risk and insurance in the past three decades is the realization that private information about individual characteristics and actions could lead to market failure due to problems of adverse selection and moral hazard (see, e.g. Rothschild and Stiglitz, 1976 and Wilson, 1977). The high cost of private insurance in the annuity and medical care markets is well documented empirically (see, e.g. Walliser, 1997, 1998, Friedman and Warshawsky, 1988, Brown, J. O. Mitchell and J. Poterba 1999, and Finkelstein and Poterba 1999) and is one of the main theoretical rationales (along with paternalism and redistribution) for the existence of Social Security (Diamond, 1977). Eckstein, Eichenbaum and Peled (1985) built a simple overlapping generations model that formalized how a mandatory Social Security system, by forcing a pooling of risks at a common actuarily fair rate for the population as a whole, could lead to a pure Pareto improvement relative to private markets for annuities.

However, these same information problems can limit the effectiveness of government run social insurance programs. This is particularly true for the Medicare, Disability Insurance, and Unemployment Insurance components of the Social Security program, where moral hazard has long been recognized to be a significant problem that increases the costs of the program. A literature on optimal taxation (see Diamond and Mirrlees, 1978 for an early contribution) showed the government could achieve second-best outcomes, i.e., social welfare maximizing outcomes subject to the constraints imposed by incomplete information by relatively simple policies that provide incomplete insurance to enable self-selection. These insights have been formalized in the subsequent literature on mechanism design, which provides formal tools for characterizing the form of efficient mechanisms that account for individuals’ private information (see Krishna and Perry, 1999 for an overview of this more recent literature).

The government is modeled as a Stackelberg leader in this literature. It chooses a social insurance mechanism that maximizes a social welfare criterion, typically a weighted average of the utility of the individuals, subject to the constraints that individuals in the system possess information that the government does not observe and will form strategies that constitute “best replies” to any mechanism chosen by the government. The solution to such a problem is referred to as a “second best” solution because it is natural to compare it to the “first best” solution under the hypothesis that the government was able to observe the private information of each individual. In the case of disability insurance, the first-best solution under complete information would be to
provide a worker full insurance, i.e. the worker would receive the same consumption level regardless of whether she was working or disabled. However, if the government cannot observe disability status and individuals experience disutility from working, then full insurance will cause too many people to apply for disability benefits. The Diamond and Mirrlees solution is to set the consumption level received by disability beneficiaries sufficiently lower than the net consumption level received by workers. The higher consumption level received by workers provides just enough incentive to induce all able-bodied individuals to work. Disabled individuals do receive a benefit, but it is substantially lower than what they would receive under the complete information solution --- they bear the brunt of the “welfare costs” of incomplete information.

Most of the early mechanism design literature abstracted from the time dimension, and focused on problems that were inherently static decision problems. The resulting efficient mechanisms were also static, even though potentially dynamic mechanisms (such as English auctions in the design of efficient government procurement procedures) were allowed to be considered in the search for an efficient mechanism. This result was justified by the revelation principle, i.e. that the search for efficient mechanisms (considering all possible static and dynamic games between the government and individuals) could be restricted to a subclass of static direct revelation games where the individual is given the right incentives to report their private information to the government truthfully. However, a subsequent literature on repeated moral hazard (Spear and Srivastava, 1987) extended the revelation principle to inherently dynamic incentive problems for which it is necessary to account explicitly for the dynamics in order to determine the form of an efficient mechanism. Although both the static and dynamic mechanism design problems can be formulated as linear programs, the size of these linear programs quickly becomes infeasibly large. The method of dynamic programming using “promised utility” as a state variable is the preferred method for solving dynamic incentive problems (see, e.g. Phelan and Townsend, 1991, Prescott, 1997 and Fernandes and Phelan, 1999). The remainder of this section provides examples of how ideas from the theory of mechanism design ideas have been applied to characterize optimal structures for various components of the Social Security program.

**Optimal Disability Insurance.**

The early work by Diamond and Mirrlees (1978) characterized the form of optimal disability insurance under the restriction that the government could only use variation in benefit levels to induce self-selection. Parsons (1996) recognized that the government also has access to a monitoring technology i.e. it operates a disability award bureaucracy that provides noisy signals about whether an applicant is truly disabled. Parsons modeled the optimal structure of the disability program when there is unavoidable noise in the disability award process (e.g. the Disability Determination Services and Administrative Law Judges, etc.) that causes some truly disabled applicants to be denied benefits and some non-disabled applicants to be awarded benefits. Parsons characterized an optimal “dual negative income tax” (NIT) benefit scheme that maximizes the ex ante expected utility of an individual in the population (prior to knowing disability status) subject to a resource constraint (equal to the production of the non-disabled fraction of the population). The optimal benefit structure involves, counter-intuitively, a substantial premium to individuals who are awarded disability benefits and decide to return to work. There is also a “welfare benefit” for individuals who don’t work. This provides partial
protection for individuals who apply for DI and are rejected, but are truly disabled. This return to work premium is similar in principle to the policy of allowing a DI recipient to keep their DI benefits during a trial work period and is required to provide the proper incentives for healthy individuals to work rather than to rely on disability.

In Parsons’ model the DI benefit is significantly larger than the welfare benefit, but the welfare benefit is smaller than the disability benefit in the Diamond and Mirrlees model, a consequence of the fact that the monitoring technology tends to help the government screen truly disabled applicants. The existence of a monitoring technology causes Parsons' dual-NIT solution to Pareto-dominate the Diamond-Mirrlees solution that ignored it. An unrealistic feature of Parsons' dual NIT scheme is that 100% of the population has an incentive to apply for DI benefits (which was assumed to be costless in his initial model). An even better solution results from charging an application fee to apply for DI benefits. The application fee can be set in such a way that only truly disabled individuals have an incentive to apply for DI benefits, resulting in lower DI taxes for non-disabled workers and higher benefits for disabled awardees than under the dual-NIT system that didn’t have an application fee. The main limitation of Parson’s analysis is that it ignores the fact that the actual DI award process is a multi-stage “game” that allows applicants the option of appealing or reapplying. The application and appeal process involves substantial delays, which might be interpreted as a type of “in-kind” application fee. Also, Parson’s ignored the substantial cost of running the huge DI bureaucracy, with over 15,000 employees at a cost of approximately 2% of total DI benefits (payroll tax receipts). Analysis of the impact of the proposed disability process reform (see, e.g. SSA, 1999) requires a more detailed model of the disability screening process, with multidimensional signals about health status and a dynamic model of individuals’ incentives to apply and appeal at various stages in the process. Empirical models are currently being developed (see, e.g. Benitez-Silva, et. al 1999a,b) to analyze these aspects of the disability award process.

**Optimal Unemployment Insurance.**

Unemployment insurance is also subject to significant moral hazard problems. The U.S. unemployment insurance (UI) program has grown rapidly, with the fraction of covered workers increasing from 60% of the working population in 1950 to 90% in the 1980s, and UI benefits as a fraction of GDP increased 0.4 percent by 1990. UI programs have been criticized, especially in European countries, for reducing incentives for an unemployed individual to return to work. On the other hand UI may provide a valuable source of insurance to risk averse workers who experience job losses, and if optimally structured, may help provide sufficient temporary liquidity to enable an unemployed individual to find a good match, reducing the costs of subsequent turnover. Determining the optimal structure of UI benefits involves a complicated balancing of the insurance benefits against moral hazard costs. Hopenhayn and Nicolini (HN,1997) provide a formal derivation of the optimal structure of UI benefits and taxes, treated as a repeated principal-agent problem involving the unemployed worker (the risk-averse agent) and the government (the risk neutral principal) which cannot monitor the agent’s search effort.

HN explicitly account for the dynamic nature of the optimal benefit structure, characterizing the optimal contract using recursive DP-like methods developed in the literature on
dynamic mechanism design and repeated moral hazard discussed above. They consider unemployment tax/transfers of the form \( \tau_t : h_t \rightarrow \{a_t, z_t\} \), where \( h_t \) is the individual’s employment history at time \( t \), \( a_t \) is the search effort of the agent, and \( z_t \) is the net transfer from the government to the individuals (a tax if negative). If the worker obtains expected discounted utility \( V \) under Social Security’s existing UI policy (which involves a constant benefit profile for the first 26 weeks and zero thereafter), then the optimal UI policy minimizes \( C(\tau) \) subject to \( V(\tau) = V \) where \( C(\tau) \) is the expected present value to the government of UI policy \( \tau \) and \( V(\tau) \) is the worker’s expected discounted utility under \( \tau \). HT show that the optimal benefit should decline monotonically with the duration of the unemployment spell in order to provide sufficient incentives for unemployed individuals to search for a new job, plus a tax rate once reemployed that increases with the length of the previous spell of unemployment. In a numerical illustration of a calibrated example, HN estimate a typical worker’s expected utility under the status quo UI policy and then estimate the cost savings of adopting an optimal UI policy. They find that the savings from adopting an optimal UI benefit policy can be as large as 30%, depending on the level of other precautionary wealth of the agent, his degree of risk aversion, and the cost of job search.

**Optimal Health Insurance.**

Of course, the market for medical insurance is also subject to severe informational and market failure problems. I noted in section 3 that over 40 million Americans have no form of health insurance, and the number is growing rapidly as firms discontinue employee health plans in growing numbers. A review of the literature on health insurance would require a survey in itself. We simply note that, despite the importance of the problem, this literature is far less developed due to the substantially more complicated nature of health insurance. Health insurance is complicated to model because it involves non-standardized services whose quality is hard to verify and has at least three different agents -- the insurer, the individual, and the physician. Given these difficulties most of the work in the area of health care is empirical, with many exploratory data analyses. See Cutler and Zeckhauser (1999) for an accessible introduction to the theoretical literature on health insurance.

There has been some pioneering initial work on the structure of optimal health insurance contracts by Ma and Riordan (1997), albeit in a highly simplified framework that ignores the role of a physician as an independent agent by assuming that the physician’s and patient’s interests are perfectly aligned. Their model deals with the moral hazard problem of overutilization of health care services that occurs under standard insurance contracts because consumers do not confront the full cost of the treatment. This analysis ignores moral hazard problems associated with imperfect monitoring of excessive and in some cases fraudulent claims submitted by physicians, hospitals, and other health care providers. Given the magnitude of Medicare spending, it is clear that there are substantial payoffs from research into managed care and other means of reducing moral hazard problems (including fraud).

3. **Stochastic General Equilibrium Models**

This section reviews the literature on stochastic overlapping generations models that has evolved from the early work by Samuelson and Diamond and others on deterministic two and
three period overlapping generations models. The work of Auerbach and Kotlikoff (1987), who studied 55 period overlapping generation models, initiated a literature on computational SGE models of Social Security. Computing solutions to SGE models is significantly more demanding than solving individual-level DP models. Indeed, most algorithms for solving SGE models require repeated solution of individual DP problems as a subroutine of an “outer” equilibrium solution algorithm that searches over trial values of prices and quantities until values are determined that equate supply and demand in capital, labor, and financial markets, and government budget constraints and aggregate resource constraints are satisfied. If we want to choose values of the unknown parameters of the SGE model so that the simulated outcomes of the model “best fit” the corresponding observed outcomes in the economy, a further “do-loop” is required. For each trial value of these parameters, the SGE model must be re-solved so that simulated values of the model’s outcomes can be measured against observed outcomes according to some relevant measure of goodness of fit. This final loop presses the limits of current technology, so that computationally simpler calibration methods are typically used to specify unknown parameters of SGE models.

Calibration is a somewhat controversial procedure because it does not provide an explicit metric of goodness of fit that could be used for testing the validity of the models, and there is no statistical theory that could enable users to assess the sampling variability in calibrated parameter values and the associated statistical uncertainty in the predictions of the model. The values of unknown parameters characterizing production technologies and preferences are taken from a variety of different sources. Some are taken from previous empirical studies and others are chosen in order to enable certain predicted quantities from the model (such as the capital output ratio) to match observed values. For a more detailed description of the process of calibration and its problems, see Hansen and Heckman (1998), and Srinivasan and Whalley (1998).

At the present state of our knowledge and technical capabilities, calibration is arguably a justifiable procedure, especially since many of the models in this literature are still too simplified to be taken seriously for statistical estimation and testing. We know that these highly stylized models will be rejected at the outset. The purpose of calibration is mainly to provide a way of choosing values for the unknown parameters of these models that have some degree of plausibility. However, due to the inferential problems with calibration, it is not yet possible to test the validity of SGE models and their implicit rationality or equilibrium assumptions rigorously, and therefore one cannot assess their predictive accuracy. Similar to DP models, the main way to assess the credibility of SGE models for use in policy forecasting is by informal “out-of-sample” predictive tests. The smaller stylized SGE models are best evaluated in terms of their qualitative insights rather than their quantitative predictions. Nonetheless, SGE models can be a very important tool for SSA, even if some of them are best employed for qualitative predictions, and even if it isn’t yet possible to quantify the uncertainty we have about the quantitative predictions of the larger and more realistic SGE models.

The STY Model.

There are several different large scale models that could be used to illustrate the SGE approach. We use the model by Storesletten, Telmer and Yaron in their paper, “The Risk-Sharing
Implications of Alternative Social Security Arrangements” (STY, 1999), as an illustration because it provides a particularly detailed analysis of specific policy reforms advocated by the 1996 Social Security Advisory Council. As noted in the introduction, one of the important goals of this chapter is to review models that have the capability to analyze these specific types of policy questions. To provide some context for STY’s analysis, recall that our review of the literature on deterministic overlapping generations models in section II of this chapter suggested that most deterministic models have difficulty finding a useful role for Social Security, other than serving as a redistributive tax/transfer program. The general conclusion of the deterministic literature is that either Social Security is irrelevant (if Ricardian equivalence holds), or if it does have an effect (due to binding liquidity constraints or the existence of corner solutions with zero intended bequests) rational individuals would prefer to abolish Social Security and replace it with a non-distortionary system of lump-sum transfers to redistribute lifetime income. If the existence of Social Security is tied to paternalistic reasons (i.e. public short-sightedness in preparing for retirement), then the general conclusion is that it is better to run it as a fully funded system than as a pay-as-you-go system because the former leads to higher long run savings and a larger capital stock, which increases consumption and overall welfare. Social Security also reduces welfare via distortionary taxation, by taking away income from young individuals during their working years when they would like to borrow against future income to increase current consumption. Because of liquidity constraints, it is believed that many young workers are at a corner solution involving zero saving except for the implicit saving via their mandatory Social Security tax contributions. In this situation, increases in the Social Security tax rate generally lowers lifetime utility (see, e.g. Hubbard and Judd, 1987).

The main drawback of deterministic models is that they are unable to evaluate properly the many potentially valuable risk-sharing features of Social Security, such as its ability to “fill in” for poorly functioning annuities markets. The absence of annuities markets forces individuals to accumulate extra wealth to self-insure against the possibility they will live longer than expected. To the extent Social Security can be viewed as an actuarially fair annuity, it enables individuals to guarantee a higher stream of consumption in old age without the “overaccumulation” of assets that is required in a world without Social Security and actuarially fair private annuities markets. This leads to the possibility that individuals would be better off under a mandatory actuarially fair pay-as-you-go Social Security system than they would be under a privatized system, even though aggregate capital accumulation under the latter system might be greater. As noted in section II of this chapter, this result was suggested by the analytical model of Eckstein, Eichenbaum and Peled (1986). However, their model ignored the general equilibrium feedbacks on capital accumulation which subsequent models, including STY’s model, have addressed.

There is also a question as to the extent to which individuals might value a progressive redistributive Social Security tax and transfer system as a type of “insurance” against lifetime earnings risk. It is clear that 
ex post 
a high income/high ability person would prefer to “opt out” of Social Security, whereas a low income/low ability person may prefer the status quo Social Security system to a privatized system. But from the 

ex ante 
perspective of an individual who is unsure whether he or she will be born into a rich or poor family or have a low or high “ability endowment”, a pay-as-you-go Social Security system might provide greater insurance than a
privatized system in which individual accounts are more directly proportional to lifetime earnings histories.

The SGE model developed by STY leads to important new insights into these issues. It also provides a detailed quantitative comparison of how the U.S. economy would be affected by various proposals described in the Advisory Council on Social Security (1996), including the “Personal Security Account” (PSA) plan. This is a hybrid that combines aspects of the current largely pay-as-you-go Social Security system with the Personal Retirement Account (PRA) plan analyzed in Feldstein and Rangelova, (1998). Under the PSA plan the long run tax rate would remain the same as under the status quo OASI system, 12.4%. However 5 percentage points of this tax would be invested in personal retirement accounts, and the remaining 7.4% tax rate would fund a scaled-down pay-as-you-go Social Security system that pays $410 per month regardless of contributions. Further, 100% of Social Security benefits would be treated as taxable income. In order to cover the transition costs of paying the unfunded liabilities to contributors under the current system, the PSA plan also includes a supplemental payroll tax of 1.5% until 2070.

STY model these issues using a stochastic overlapping generations model where individuals are “born” at age 22, work until retirement at age 65, and live at most to age 100. Individuals can die prior to age 100 according to mortality hazards observed for U.S. females in 1991. All individuals have identical constant relative risk averse utility functions, \( u(c) = (c^\rho -1)/\rho \) with \( \rho = -1 \), and maximize expected discounted utility with subjective discount factor \( \beta = 1.011 \), which corresponds to a negative subjective discount rate. STY model the risk of labor earnings as follows: an individual \( i \) of age \( a \) works an exogenously specified amount of hours per period given by

\[
\log(n_{ia}) = \kappa_a + \alpha_i + z_{ia} + g_{ia}, \tag{5.5}
\]

where \( \kappa_a \) is an age-specific constant included to enable the model to capture the age-earnings profile, \( \alpha_i \) is an individual-specific “random effect” realized at birth, and \( \{z_{ia}\} \) and \( g_{ia} \) are time varying shocks designed to provide a “reduced-form” account of factors such as job loss, health problems and so forth that may temporarily interrupt an individual’s labor supply trajectory. They assume that \( \{g_{ia}\} \) represents transitory shocks which are independently and identically distributed (IID) whereas \( \{z_{ia}\} \) follows a first order Markov process given by

\[
z_{ia} = \omega z_{i,a-1} + \eta_{ia} \tag{5.6}
\]

where \( \{\eta_{ia}\} \) are IID shocks. Thus, the STY model incorporates two basic types of risk: mortality risk and earnings risk. There are three different types of earnings risk: 1) persistent individual differences in ability (or labor “endowment”) accounted for the \( \alpha_i \) random effects, 2) serially correlated shocks to labor supply accounted for by the \( \{z_{ia}\} \), and 3) transitory shocks to labor supply represented by the \( \{\eta_{ia}\} \). STY used econometric methods (exactly identified generalized methods of moments estimation) to choose the parameters of this parameterization of the stochastic process of labor earnings \( y_{ia} = wn_{ia} \) that matched observed moments (means, covariances, etc.) of labor earnings in the Panel Study on Income Dynamics (PSID).
The final individual-specific state variables, are non Social Security wealth \( w_a \) and “Social Security wealth” \( s_a \). The latter variable is a scalar summary of the cumulative Social Security contributions of an individual of age \( a \) under various alternative policies \( \alpha \), including \( \alpha = \text{psa} \) (the PSA plan) and \( \alpha = \text{pg} \) (a pay-as-you-go Social Security system). \( s_a \) is updated according to the following scheme,

\[
\begin{align*}
\alpha s_a &= \begin{cases} 
\alpha s_a - 1 r_a + \min(\tau_a n_a w, \bar{s}_a) & a \leq 65 \\
\alpha s_a & a > 65 
\end{cases}
\end{align*}
\] (5.7)

where \( \bar{s}_a \) denotes a government-mandated floor on annual Social Security contributions—a form of insurance against years in which labor earnings are very low. \( r_a \) denotes the rate of return to Social Security contributions under policy \( \alpha \). This is equal to the real rate of growth in the wage base when \( \alpha = \text{pg} \) and equals the real rate of return on capital when \( \alpha = \text{psa} \). Social Security benefits are a nonlinear function of contributions given by

\[
(\tau_a n_a w, \bar{s}_a)
\] (5.8)

where \( \bar{s}_a \) denotes a minimum benefit level, and \( d_a \) is a concave function relating Social Security contributions to benefit payments above the minimum benefit level. In the case \( \alpha = \text{pg} \), the \( d_a \) function can be thought of as an approximation to the relationship between AIME and PIA. Wealth evolves according to the standard budget equation

\[
w_{a+1} = w_a (1 + r) - \tau_k w_a r + n_a w(1 - \tau_l - \tau_a) + b^a(s_a)(1 - \tau_l) - c_a,
\] (5.9)

where \( \tau_k \) is the tax rate on interest earnings, \( \tau_l \) is the tax rate on wages, and \( \tau_a \) is the Social Security tax rate (which may be different under a pay-as-you-go system than under the PSA plan). The individual’s “state” at age \( a \) can be summarized by the 3-dimensional vector \( (z_a, w_a, s_a) \) where \( z_a \) is the serially persistent shock to labor supply at age \( a \), \( w_a \) denotes accumulated (non-Social Security) wealth at age \( a \), and \( s_a \) denotes the accumulated Social Security contributions discussed above. Because labor supply is given by the exogenous process in equation (5.5), and everyone is assumed to retire at age 65, the individual’s only choice variable is how much to consume, \( c_a \). The Bellman equation for this problem is given by:

\[
V_a(z_a, w_a, s_a) = \max_{c_a} \left[ u(c_a) + \beta \phi_a \left[ V_{a+1}(z_{a+1}, w_{a+1}, s_{a+1}) p(z_{a+1}, s_{a+1} | z_a, s_a) \right] \right]
\] (5.10)

where \( \phi_a \) is the survival probability for a person of age \( a \) and where increases in \( c_a \) result in one for one reductions in \( w_{a+1} \) due to the budget constraint in equation (5.9). The Bellman equation is solved via backward induction from the maximal lifespan, age 100. As we will see, the DP
problem must be repeatedly solved in an inner loop of an outer algorithm that searches for
equilibrium values of endogenously determined quantities such as aggregate capital and labor and
wages and interest rates. Thus, the feasibility of this approach depends on the feasibility of rapidly
solving individuals’ DP problems as noted in section V.2.

Aggregate output \( Y_t \) is generated from a Cobb-Douglas production function,
\[ Y_t = Z_t K_t^\theta L_t^{1-\theta}, \]
where \( K_t \) is the aggregate capital stock, \( L_t \) is the aggregate labor supply, and \( Z_t \) is
a term that can be thought of as capturing technological progress or “knowledge capital”, and is
assumed to grow as a constant rate of \((1 + \lambda)\) per period. The population in this economy is
assumed to grow at a constant rate of \((1+n)\) per period. Thus, the overall rate of growth of the
economy is \((1+\gamma) = (1+n)(1+\lambda)\). There is also a government that spends an aggregate amount \( G \)
on items other than Social Security, and issues an aggregate amount of bonds \( B \) that are held
constant as a fraction of GNP, and thus grow at rate \( \gamma \).

**Definition:** An equilibrium for the stochastic overlapping generations economy consists
of market clearing wages and interest rates, and a set of value functions \( \{V_a\}, a = 21,...,100 \)
satisfying:

1. Wages and interest rates are equal to the marginal product of labor and capital,
   respectively:
   \[
   \begin{align*}
   w &= \partial Y / \partial L \\
   r &= \partial Y / \partial K - \delta
   \end{align*}
   \]
   where \( \delta \) is the depreciation rate of capital.

2. Individuals behave optimally, i.e. \( \{V_a\}, a = 21,...,100 \) solves Bellman’s equation (5.10).

3. Markets clear and aggregate quantities are the sum (integral) over individuals in
   the economy with respect to the population distribution \( \mu \) the joint (stationary)
distribution of the ages, labor supply shocks (both persistent and transitory), and
wealth levels and Social Security contribution levels.

\[
\begin{align*}
K + B &= \int (w_a + s_a^{PSA})d\mu \\
L &= \int n_ad\mu \\
Y &= ZK^\theta L^{1-\theta} = \int c_a d\mu + (K' - K) + \delta K
\end{align*}
\]

4. The government budget constraint is satisfied:
\[
G + (r - \gamma) B = \int (\tau_k rw_a + \tau_l w_a) d\mu + E
\]
where the second term on the left hand side of (5.13) is net interest cost on the
government debt, the integral on the right hand side of (5.13) is total revenue from
taxes on  capital and labor income, and $E$ is government estate tax revenue, equal
to 100% of the wealth of any individual who dies.

5. The pay-as-you-go component of the Social Security program is balanced period-
by-period

$$
\int \left( \tau_{pg} wn_a + \tau_{s} b_{pg}^s \right) d\mu = \int b_{pg}^s (s_{a}^{pg}) d\mu \quad (5.14)
$$

An equilibrium to the SGE model is computed as follows. Initial values for $G$ and $B$ are
specified, along with guesses for tax rates that satisfy the government budget constraint (5.13)
and values of $w$ and $r$ that satisfy equilibrium in the market for labor and capital given in (5.11).
Then the individual’s DP problem is computed using the Bellman equation (5.10), and from this
solution the stationary distribution $\mu$ of individuals’ ages, idiosyncratic labor shocks, and wealth
levels can be computed. Using $\mu$, aggregate capital, labor supply and output can be calculated
from equation (5.12). The initial guesses for tax rates, $w$ and $r$ are then updated via a quasi-
Newton algorithm designed to solve the system of nonlinear equations (5.11), (5.12), (5.13) and
(5.14). Thus, the SGE model requires repeated solutions of a DP problem that is nested within an
outer equilibrium solution algorithm, which in this case is simply a nonlinear equation solver.
Because it may take hundreds or thousands of trial evaluations of the DP problem until
equilibrium wages, interest rates, and tax rates are found, it is clear that the DP problem must be
solved very efficiently (in a matter of minutes) for the SGE approach to be computationally
feasible.

With the ability to compute equilibria for different policy regimes, the only remaining
detail is to specify a way of comparing welfare losses/gains across different regimes. STY define
a certainty equivalent measure of the welfare impact of a policy change by solving for an age-
invariant proportional increase or decrease in per period consumption that equates the expected
discounted utility under some alternative policy to the expected discounted utility that an
individual would receive under the status quo. This certainty equivalent will depend on the initial
state of the individual when this calculation is made (e.g. age, wealth levels, and the realized
values for the various shocks to labor supply), so STY solve for the certainty equivalents for all
possible combinations of this initial state and then compute an average certainty equivalent value
by integrating the state-dependent certainty equivalents with respect to the stationary distribution
$\mu$.

STY computed equilibria for several different policy regimes, measuring welfare changes
relative to the status quo as noted above. Their idealization of the status quo (SQ) policy regime
is intended to approximate the current U.S. Social Security system, but without the Trust Fund.
Thus, STY assumed that the Trust Fund balance is zero and that Social Security is run as a pure
pay-as-you-go system. They parameterized the benefit function $b_{pg}(s_{a}^{pg})$ in equation (5.8) to
approximate the benefits payable under the current Social Security rules. They assumed that 25%
of Social Security benefits are taxable at the ordinary income tax rate, $\tau_{t}$, and set values for the
minimum benefit level $b^{pg}$ and the maximum Social Security $\tau^{PSA}$ rate to approximate the actual values in 1995 of $572$ per month and $61,750$ per year, respectively. When they solved for the Social Security tax rate $\tau_{pg}$ that satisfies the budget balance condition in equation (5.14), they obtained a rate of $\tau_{pg} = .1092$ for the SQ policy regime, which is reasonably close to the actual OASI tax rate of 12.4%. They chose the parameters of the PSA plan to approximate the plan detailed in the 1996 Advisory Council Report. They did this by setting $\tau_{psa}$, the tax rate that finances the individual accounts in the PSA plan, to be $\tau_{psa} = .05$, and assumed that these individual accounts would earn interest at the equilibrium interest rate $r$ calculated in equation (5.11). The accumulated value in these individual accounts are assumed to be paid out as a lump-sum distribution at age 65. The PSA plan also contains a scaled-down version of a pay-as-you-go Social Security system that pays a flat benefit of $410$ per month to all retirees regardless of income or wealth. They solved for the payroll tax rate that balances the budget of the pay-as-you-go component of Social Security and obtained $\tau_{pg} = .0513$, so that the total tax rate under the PSA policy regime amounts to 10.13%. Finally STY consider the radical policy option of abandoning the pay-as-you-go component of Social Security altogether and letting individuals rely completely on their individual accounts, a policy regime that they describe as “privately provided pensions” (PP). In the PP policy regime the Social Security payroll tax is set to zero, and individuals have complete discretion about how much of their income to save for their retirement each period. It is important to note that STY assume there are no private annuity markets, so individuals are constrained to consuming a fraction of their remaining wealth during each year of retirement. However STY do evaluate a version of the PP scenario where there is an actuarily fair annuity market. This leads to an interesting estimate of the potential welfare gains to “filling in” a “missing market”.

Initially, STY ignored Social Security’s unfunded pension liabilities, and calculated equilibria and welfare under new long-run steady states for the PP and PSA regimes under the implicit assumption that SSA reneges on its obligations to existing retirees under the SQ policy regime. They find welfare gains equal to 0.9% and 7.3% of consumption, respectively, for the PP and PSA regimes. The PP economy is dynamically inefficient with an equilibrium interest rate of $r = 0.0096$ that is less than the sum of the rate of growth of the economy, $\gamma = 0.025$ (equal to the sum of the 1% population growth rate and a 1.5% rate of growth in real wages). The dynamic inefficiency is partially a result of the assumption that $\beta > 1$ (which implies that individuals value future utility more than current utility), and partially a result of the substantially larger precautionary savings that individuals undertake in the PP economy due to the lack of a redistributive Social Security system and the absence of private annuity markets. Altogether, these factors cause the aggregate capital stock to be nearly twice as high under the PP policy regime than under the SQ regime. This is the reason why the welfare gain under the PSA plan is so much larger than the PP plan.

In all subsequent calculations STY carefully treat the issue of how to account for Social Security’s obligations to current beneficiaries and contributors in the aftermath of a policy change to a new policy regime. They assume that Social Security issues an amount of debt $D$ equal to the expected discounted value of the unfunded liabilities under the SQ policy regime. Once this extra “Social Security debt” is accounted for, the welfare calculations are rather different: individuals would be willing to pay 3.7% of per period consumption to adopt the PP regime, and 4.0% of per
period consumption to adopt the PSA regime. The fact that the welfare gain is larger once we account for the Social Security debt seems non-sensical. However, it has been known since Diamond’s (1965) analysis that government debt “crowds out” private capital accumulations, lowering the aggregate capital stock. When there is dynamic inefficiency, as there is in the case of the PP economy prior to issuance of Social Security debt, then issuing additional debt actually results in a welfare gain. Indeed, the issuance of Social Security debt causes capital stock to fall by 30%, increasing the before tax rate of return \( r \) to 3%, which exceeds the 2.5% rate of growth of the economy. Note that in the case of the PSA plan, there is implicit “Social Security debt” due to the fact that the PSA plan contains a pay-as-you-go component, which is also predicted to crowd out private capital accumulation in a manner similar to explicit government debt. This crowding out implies that the equilibrium for the PSA regime is dynamically efficient even in the absence of explicit Social Security debt, so it follows that issuing additional Social Security debt has the expected effect of reducing the welfare gains of the PSA plan from 7.3% to 4.0% of per period consumption.

An interesting feature of STY’s analysis is their ability to decompose the change in welfare into four components: 1) general equilibrium effects (i.e. the overall effect of the policy on aggregate capital accumulation and market clearing prices), 2) taxation effects (i.e. reductions in the incentive to work and accumulate capital due to distortionary taxation of capital and labor), 3) incomplete market effects (i.e. the effect of the policy on individuals’ access to actuarially fair annuities, and 4) risk sharing effects (i.e. the effect of redistributive Social Security taxes and transfers on the welfare of different individuals in the economy). Under the PP plan, the total welfare gain of 3.7% of per period consumption equals the sum of a 6.8% welfare gain due to general equilibrium effects and a 3.1% welfare loss due to the other effects, including a 1.6% welfare loss due to reduced income risk sharing under a privatized system, a 0.8% welfare loss due to removing Social Security as substitute for missing annuity markets, and a 0.7% welfare loss due to distortions caused by increased capital tax revenues on interest earnings on individuals’ retirement savings.

“The main message of this decomposition is simple. The lion’s share of the welfare gain associated with Social Security reform derives from the general equilibrium effects. As we demonstrate explicitly below, this is a manifestation of the fact that, as a whole, society saves more under the PP or PSA arrangements, leading to lower interest rates, a higher capital stock, and a higher level of aggregate output and consumption. Income risk sharing effects are also important as evidenced by the 1.6% loss associated with the PP alternative, whereas effects directly attributable to the provision of annuities and capital income tax distortions play a relatively minor role.” (STY, 1999, p. 24).

“The primary force driving these welfare gains is a kind of externality associated with retirement savings. Social Security provides a participant with an imperfect annuity. When that annuity is removed --- either completely or partially --- individuals save more during their working lives in order to insure against the possibility of outliving their resources during retirement. The collective effect of this increase in saving, something which is external to each individual’s choice problem, is an increase in aggregate capital, output and consumption. This increase in aggregate resources lies at the heart of the welfare gains we uncover.” (STY, 1999, p. 31).
A final, interesting policy experiment that STY study is the impact of perfect annuity markets on welfare under the SQ, PSA and PP policy regimes. *Ceteris paribus* the addition of a perfect annuities market should make all individuals better off, because it provides an additional option for hedging against risk. Paradoxically STY find that the addition of perfect annuities markets reduces welfare under each of the three policy regimes. The reason for the welfare reduction is the “savings externality” noted above:

“We argue that this finding --- that the provision of annuities can reduce welfare --- is not unlike a classic set of results from the literature on general equilibrium with incomplete markets: the endogenous nature of the set of investment opportunities generates an externality which can make changes in market structure welfare decreasing. Our example of this is stark in that it abstracts from privately provided (imperfect) annuities which are, to some extent, available in actual financial markets. Nevertheless, it makes a point which is often overlooked in the debate on Social Security reform: that the savings response to a change in the system can very much depend on how the availability of annuities is altered. Our results suggest that the quantitative magnitude of this response is substantial.” (STY, 1999, p. 31).

It is also interesting to note that a “money’s worth” measure, the internal rate of return on Social Security contributions, provides a completely misleading indicator of the welfare change resulting from the addition of annuity markets. For example, the internal rate of return on total contributions under the PSA plan is 3.7% without complete annuity markets versus 4.2% with complete annuity markets. The higher IRR suggests that individuals should be better off with complete annuity markets, but the discussion above shows that individuals are actually worse off, and the higher IRR is due to the fact that the marginal product of capital necessarily increases when there is less capital relative to the exogenously determined aggregate labor supply in this economy.

**Other SGE Models.**

Imrohoroglu, Imrohoroglu and Joines (1998) provide a survey of other recent computational SGE models of Social Security. There is no need to repeat that survey here, although, I do briefly discuss papers by Huang, Imrohoroglu and Sargent (HIS,1997) and De Nardi, Imrohoroglu and Sargent (DIS,1998) because they deal with the important issue of transition dynamics in moving from the status quo to an alternative policy regime, an issue that STY and most previous work in this literature have abstracted from.\(^{55}\) HIS study transition dynamics in an economy where birth rates and mortality rates are time-invariant, and the main

---

\(^{55}\) Although original computational SGE models in Auerbach and Kotlikoff (1987) did explicitly calculate transition dynamics, these calculations were simplified by the fact that the models were deterministic so that it was only necessary to compute paths for the key aggregate quantities. When uncertainty is incorporated in the model, we have a form of heterogeneity in terms of different realized outcomes for uncertain variables in the model. It is then necessary to describe the evolution of the *distributions* of these uncertain quantities, which is a much more difficult computational task. The initial stochastic SGE models of Social Security such as Hubbard and Judd (1987) and Imrohoroglu, Imrohoroglu and Joines (1995) showed that it was possible to incorporate various types of uncertainty, but at the cost of abstracting from transition dynamics. To our knowledge the HIS and DIS papers are the first SGE models to explicitly account for transition dynamics.
non-stationarity results from the anticipated consequences of an initially unanticipated policy change and the subsequent fully anticipated transition to a new steady state. They consider two cases, a “small open economy case” where the interest and wage rates $r$ and $w$ are exogenously determined, and a “closed economy case” where $r$ and $w$ are determined endogenously from the marginal product of capital and labor, respectively, similar to the procedure described for STY. They compute two different policy experiments describing how the economy responds to a change from a pay-as-you-go to a fully funded Social Security system. In the first experiment, the government suddenly terminates Social Security and issues “recognition bonds” equal to the present value of the unfunded liabilities under the pay-as-you-go regime. This debt is fully retired over a 40 year transition period to the new steady state. In the second experiment, Social Security benefits are untouched, but the system is funded by building up a Trust Fund over a 40 year transition period at a sufficient rate that all benefits can be paid from interest earnings on the Trust Fund by the end of the transition period. These experiments are performed in a model where individuals live for 65 years, with the interpretation that an individual is “born” at age 20, retires at age 65, and lives at most until age 85. The main modeling innovation is the use of non-expected utility linear-quadratic preferences that deliver linear decision rules. This simplifies computation and aggregation considerably, while allowing for a form of risk-sensitivity that does not satisfy the standard version of certainty equivalence of standard linear-quadratic expected utility models. This allows the model to capture a precautionary savings motive, and an increase in the within-cohort variance in consumption spending as the cohort ages. These preferences are described recursively via a generalized form of the Bellman equation (5.10):

$$V_a = -\frac{1}{2}(\pi c_a - \xi_a)^2 + \beta \phi_a \frac{2}{\sigma} \log \left[ E\left\{ \exp \left( \frac{\sigma V_{a+1}}{2} \right) I_a \right\} \right]. \quad (5.15)$$

where $I_a$ denotes information available at age $a$, $\phi_a$ is the survival probability at age $a$, and $\xi_a$ can be regarded as a “target consumption level” at age $a$.

Under experiment 1, consumption decreases for the first 20 years of the transition, reflecting the initial transition costs involved in saving for individual retirement and paying the interest and principal on the recognition bonds that cover Social Security’s unfunded obligations. After 40 years, aggregate consumption levels are the same as they were pre-transition, and by 100 years after the policy change, the capital stock has risen by 40%, the labor income tax rate decreases from 34% to 14%, and consumption levels asymptote to a level 9.8% higher than prior to the transition. However, during the 40 year transition period, the income tax must be raised to 38% to pay off the entitlement debt, which is initially equal to 2.7 times real GDP. Under experiment 2, aggregate consumption immediately drops by 2.5%, representing the extra taxes that are collected to build up the Trust Fund, but thereafter consumption levels steadily rise, asymptoting to a level 8.4% higher than prior to the transition. All cohorts born 25 years or more after the transition gain from the change to a fully funded Social Security, but the first 25 cohorts born after the date the policy is announced have a lower lifetime utility.

HIS quantified the welfare gains associated with the two reform plans as follows. First they computed the present value of additional wealth required to make the individuals along a transition path indifferent between the reform or remaining under the status quo pay-as-you-go
system. Then they computed the annuity equivalent value of this certainty equivalent wealth measure and expressed it as a fraction of GDP at the initial stationary equilibrium under the status quo Social Security policy. The privatization plan yields welfare gains of 1.3% of GDP for the small open economy scenario and 2.0% of GDP under the closed economy scenario. They found larger gains to preserving the defined benefit plan and simply building up the Trust Fund, which is equal to 2.1% of GDP under the small open economy scenario and 2.8% under the closed economy scenario. The reason the policy of building up the Trust Fund delivers larger welfare gains than the policy of privatizing Social Security is due to “the scheme’s public provision of insurance both against life-span risk and labor income volatility.” (p. 7). They conclude that while the welfare of transition generations falls on average during the 40 year transition period, these policies “redistribute enough of the permanent gains from future to current generations to induce a majority of each cohort to assent to the transition.” (p. 8).

It is interesting to contrast HIS’s conclusion about the welfare gains to privatization with the conclusion obtained from the “short-cut” approach to the analysis of transition costs by Geanakoplos, Mitchell and Zeldes (GMZ) (1998a,b) discussed in section IV.2 above. Recall that their analysis lead GMZ to conclude that there is no gain to privatizing Social Security because the cost of paying off the entitlement debt completely offsets the gains from the higher returns earned in privatized accounts or by investing Trust Fund balances in equities. In HIS’s analysis, equilibrium interest rates actually fall in the new steady states (a simple consequence of the increased steady state capital stocks). If we were to use these lower rates of return in GMZ’s calculations, then it would appear that their analysis leads to the conclusion that privatization leads to a clear welfare loss. Note also that GMZ do not assume that the “recognition bonds” are ever retired, but instead are continually refloated as a constant fraction of GDP. In HIS’s analysis, the recognition bonds are paid off in a relatively short span of 40 years. If we account for the cost of paying off a principal balance of 2.7 times GDP, this would be an additional factor that would lead the GMZ type of short-cut analysis to conclude that privatization involves a clear welfare loss. How is it that HIS are able to find a net welfare gain to policies of privatization or fully funding Social Security net of all transition costs? The reason is fundamentally due to the general equilibrium feedbacks that the short-cut analyses ignore. The long run build up in capital stock leads to higher consumption and welfare levels that are correctly accounted for within the SGE model, but which are ignored by the short-cut approaches.

The final large scale SGE model reviewed here is a paper entitled, “Projected U.S. Demographics and Social Security” by De Nardi, Imrohoroglu and Sargent (DIS) (1998). This paper illustrates how to account for nonstationarities in survival probabilities and demographic patterns, including the “demographic shift” to higher dependency ratios in the next century as the baby boom generation approaches retirement. The paper is noteworthy because it introduces several methodological innovations, including modeling consumption and labor decisions jointly within a linear-quadratic specification of preferences and allowing for a bequest motive. Demographic shifts constitute an important and hard to diversify “macro risk” confronting SSA and individuals. An unfavorable demographic shift (i.e., one that increases the dependency ratio by lowering birth rates or increasing life expectancies) makes the current pay-as-you-go system even more out of long run balance at current tax rates and benefit levels than it is now, increasing the “political risk” that major changes will be made to the system, possibly including “radical reform”
such as privatization. The Social Security Actuary projects a doubling in the dependency ratio (the ratio of Social Security beneficiaries to taxpayers) between 1997 and 2050. The less favorable demographic patterns effectively increase the transition costs involved in various Social Security reform proposals, which may invalidate conclusions obtained from previous SGE models on the desirability of privatizing or building up the Social Security Trust Fund. But the paper also shows that a number of unpleasant adjustments must be made even under the status quo pay-as-you-go system.

DIS consider 8 experiments involving relatively minor changes to the structure of a status quo similar to those that have been considered in recent policy discussions. These changes are designed to restore long run fiscal balance following the demographic transition, but without altering the fundamentally pay-as-you-go character of the Social Security system. The fiscal adjustments include an increase in the payroll tax rate on labor earnings, the introduction of a consumption tax, various combinations of benefit reductions and tax increases, and a final experiment designed to increase the linkage of benefits to cumulative earnings, while also adjusting the tax rate on labor income. The outcomes of the 8 experiments are summarized below:

1. In experiment 1, SSA uses a rise in the payroll tax rate to balance the system, increasing it from 29.7% to 59.5% between 2000 and 2060. Over this period the amount of labor supplied falls by 20.8%, capital stock falls by 11.7% and output (GDP) falls by 17.4% relative to what it would have been if the economy had grown at the exogenous productivity growth rate of 1.6% per year.

2. In experiment 2 the government uses a consumption tax to balance the system. The outcomes are better than in experiment 1. The consumption tax rate must rise from 5.5% to 36.9% to balance the system, aggregate labor input falls by 12.4%, capital stock rises by 11.3% and GDP falls by 4.6%, substantially less than the 17.4% decrease when a labor income tax is used to balance the system. This shows the impact of using a more distortionary labor income tax instead of a consumption tax.

3-4. Experiments 3 and 4 postpone the retirement age to 69 and then increase the labor tax rate (experiment 3) or the consumption tax rate (experiment 4) to finance the remaining burden. Postponing the retirement age necessitates smaller tax increases. When labor income taxes are increased to balance the system, they must only rise to 52.9% instead of 59.5% (if the retirement age was not postponed). If the consumption tax is used to balance the system, it must rise to 31.2% instead of 36.9% (if the retirement age was kept at 65). The fall in labor supply and output is therefore smaller. Labor supply falls by 14.8% in experiment 3 and 4.4% in experiment 4. Output falls by 10.7% in experiment 3 and 2.3% in experiment 4. Capital stock falls by 3.4% in experiment 3 and rises by 13.5% in experiment 4.

5. In experiment 5, Social Security benefits are taxed and the labor income tax is raised to finance the residual burden. The results are similar to the use of a
consumption tax, as in experiment 2. Labor supply falls by 10.6%, capital stock rises by 11.3% and GDP falls by 3.4%.

6. In experiment 6, there is a stronger linkage between retirement benefits and past earnings in order to reduce the perception that Social Security contributions are a pure tax and reduce the resulting distortion in labor/leisure choices. The results from experiment 6 are somewhat better than experiment 1. The payroll tax rate rises to 51.3%, labor input falls by only 6.6%, the capital stock rises by 5%, and output falls by only 2.4%.

7. Experiment 7 is similar to experiment 6 except that the consumption tax instead of the labor tax is gradually increased to pick up the residual burden of balancing the Social Security system. Thus, the labor tax rate remains at 29.7%, but the consumption tax rate is raised from 5.5% to 30.5%. The consumption tax produces smaller distortionary effects on labor supply than the payroll tax, so GDP increases by 3.7% in the new steady state, labor supply falls by only 4.4% (compared with 6.6% in experiment 6), and the capital stock increases by 20.3%.

8. Experiment 8 is an uncompensated phase out of the current Social Security system in which benefits are gradually reduced to zero between 2000 and 2050. By 2060 individuals must rely on their own savings (either investments in physical capital, stocks, or government bonds) to provide for their retirement. This yields a substantial increase of 38% in mean asset holdings and 42.5% in the capital stock. This causes the interest rate to decrease from 5.9% to 3%. The labor income tax rate falls slightly from 29.7% to 26%, wages increase by 15.2% (the largest increase in any of the experiments), but labor supply still falls by 5.6%. GDP increases by 8.7% (the largest increase in any of the experiments) and consumption falls by 0.2% (the smallest decrease in any of the experiments).

DIS conduct a welfare comparison of these eight experiments by measuring the fraction of initial wealth that makes a person indifferent between experiment \( j \) and experiment 1. They conclude that

“Essentially all future generations are better off under Experiments 2-8 than Experiment 1. In fact, when we compute an overall welfare measure by properly taking into account the welfare gains and losses of all generations, weighing them by their (time-varying) population shares and discounting the future gains and losses by the after-tax real interest rate, all of the experiments deliver a welfare gain. Experiment 2 produces a welfare improvement of 54% of GDP (at the initial steady state) relative to experiment 1. Experiments 3, 4 and 5 yield overall welfare gains of 49%, 84% and 56% of GDP, respectively. Experiments 6, 7, and 8 produce overall welfare gains of 197%, 189%, and 10.9% of GDP, respectively, relative to Experiment 1.

When welfare gains/losses are plotted by birth cohort (measured as a fraction of assets that would need to be given to a person born in year \( t \) living under experiment 1 to make him or her indifferent between experiment \( j \) and experiment 1), DIS find that most of the experiments
produce welfare losses for the transition generations, especially those in the baby boom generations. Subsequent generations, however, (i.e., those born after 2000) are better off under experiments 2 to 8 than under experiment 1. The experiment that yields the highest welfare gains to future generations is experiment 8, the gradual phase out of Social Security, but it produces the highest welfare losses to the transition generations. This is why experiment 8 yields only a 10.9% net discounted welfare gain when the gains and losses of all current and future generations are aggregated. Because the future welfare gains of generations born after 2000 are discounted, the losses of the transition generations weigh more heavily on this measure than the welfare gains of future generations, even though the ultimate steady state performance of the economy under the privatization alternative looks the best.

The only policy that yielded a net welfare gain to all generations (relative to experiment 1) was experiment 6, which amounts to a switch from the status quo Social Security system to a defined contribution system. The welfare gains to future generations is not as large for experiment 6 as for the privatization experiment 8. For example, individuals born in 2150 would pay 1.5% of their assets to move from experiment 1 (higher labor tax rates) to experiment 6 (where individuals contribute to their own govt.run retirement account), whereas these same individuals would be willing to pay 2% of their assets to move to a completely privatized system where they would save for their own retirement. But experiment 6, unlike experiment 8, does not “renege” on the obligations to the transition generations. As a result it yields the highest overall discounted welfare gain (197% of GDP).

The DIS model yields dramatically different predictions of the rise in tax rates necessary to keep the Social Security system in long run actuarial balance than the Social Security Office of the Actuary’s predictions. The 1999 Trustees Report predicts that only a 2 percentage point increase in the 12.4% Social Security OASDI combined employer/employee tax rate will be necessary to keep the current system in long run actuarial balance. The DIS model predicts that much higher tax increases will be necessary because it accounts for the distortionary, disincentive effects on labor supply and saving that increased payroll taxes will have.

It is not clear whether the predictions of the DIS model are empirically realistic, even though the model was calibrated to match the demographic projections of the Office of the Actuary. The main uncertainty is whether the degree of elasticity of labor supply assumed in their model is empirically plausible because this appears to be one of the key factors behind their predictions of large losses in output and reductions in labor supply under the status quo. The 21% reduction in labor supply necessitates correspondingly larger increases in Social Security tax rates to balance the system.

Overall, all of these models suggest the importance of accounting for general equilibrium feedback effects in the analysis of alternative Social Security policies. Future work should concentrate on empirical validation of their estimated utility function parameters and labor supply and savings elasticities to determine whether the rather dire predictions from some of these models are reasonable. If so, they suggest that the current Social Security program will be in much worse shape in the future unless actions to change the system are taken in the not too distant future.
Models With Endogenous Policies.

Most of the literature has modeled the government as a Stackelberg leader that has the ability to commit to any given policy. A brief review of the history of Social Security amendments, however, shows that the government makes relatively frequent changes to the program. A number of studies have noted the importance of political risk, i.e. the uncertain future policy changes that are hard for individuals to predict or diversify against. Bütler (1999) models changes in policy as an exogenous stochastic process and shows that policy changes that are announced sufficiently far in advance have much less severe welfare consequences than sudden and unexpected policy changes because in the former case individuals have lead time to undertake offsetting actions. These actions include building up precautionary savings balances and changing their planned retirement dates. Diamond (1994) and McHale (1999) also discuss the importance of political risk as a consideration affecting the design of alternative social insurance programs.

Another line of papers has endogenized the government’s policy formation process by modeling majority voting over alternative policies (see, e.g. Cooley and Soraes, 1999, Conesa and Krueger, 1998, Rangel and Zeckhauser, 1999, Galasso, 1998, Boldrin and Alonso 1998, and Boldrin and Rustichini, 1999). This literature is important, because it defines the conditions under which a Social Security program can be politically sustainable. As noted above, one of the paradoxes of the current literature is that despite attempts to model various types of risk and market failure that could lead to a beneficial role for a government-run social insurance program, most models have difficulty in demonstrating significant welfare gains from the existence of a pay-as-you-go Social Security program. Indeed, most models find that the system leads to significant welfare losses; the models by Eckstein, Eichenbaum and Peled (1985) and Merton (1983) are among the few exceptions that find a welfare gain. Attempts to explain a positive role for Social Security as a paternalistic system designed to prevent myopic undersaving on the part of individuals also fail to find a significant role for Social Security (Feldstein 1985, Imrohoroglu, Imrohoroglu and Joines, 1999d). Explaining the immense political popularity of the Social Security program requires showing that the number of people who perceive a net gain from Social Security must outweigh those who see it as a loss. Cooley and Soares (1999a) show how such political support could arise from majority voting even in a world where all generations know that a pay-as-you-go Social Security program produces a welfare loss. An initial population votes to initiate such a program because they are beneficiaries of net transfers from future generations. The system can be sustained in subsequent years because there are just enough old people who will vote to continue it out of fear that their own contributions would be lost if the system were terminated.

However, this equilibrium is a tenuous one and can be upset by demographic shocks such as a wave of baby boomers. The model developed by Cooley and Soares (1996) predicts that the rise in tax rates or reductions in benefit levels needed keep a pay-as-you-go system in long run balance will lead to its abandonment. Boldrin and Rustichini (1999) develop a similar model, and show that a pay-as-you-go Social Security system will eventually collapse with a probability of one due to demographic shocks such as the baby boom. Cooley and Soares (1999b) showed that there exist transition policies that make it politically feasible to provide a smooth transition from a
pay-as-you-go Social Security to a privatized system. This requires the right package of “recognition bonds to transfer enough of the welfare gains that future generations receive from privatizing or fully funding the system to the transition generations to convince them that the change is worthwhile. While this literature is rather new and the models are still rather stylized, the issues they address are at the heart of current concerns about the reform of the Social Security system.

**Concluding Remarks on SGE Models.**

This review of several recent SGE models suggests that some of these models are now sufficiently realistic that they can be useful inputs to the evaluation of a range of policy issues of interest to SSA. These models are especially useful for evaluating “radical” policy reforms, such as privatization of Social Security or making the transition to a fully funded system in which general equilibrium feedbacks are likely to be substantial.

But there are limitations of SGE models that could lead to reservations about using existing versions of these models “off the shelf” for quantitative forecasting. Moreover, for some policy issues, general equilibrium feedbacks may be less important than for others, and on these issues simpler partial equilibrium models could provide sufficiently accurate policy forecasts.

The biggest limitation of SGE models from the standpoint of their use in policy forecasting at SSA relates to their failure to incorporate various risks, aspects of reality, and features of the Social Security program. These limitations are largely related to the “curse of dimensionality” associated with solving the DP problem for more realistic formulations of individuals’ decision problem, as discussed in section V.2 of this chapter.

One key limitation is the failure to account for stochastic macro shocks, such as the effect of an oil shock or a stock market crash. The production technologies and rates of return on assets in current generation SGE models are deterministic. None of these models allows for the possibility of stock market “bubbles”, which is a focus of concerns about privatization and investing Trust Fund assets in equities. The failure to incorporate stochastic macro shocks limits the usefulness of these models for studying the role of Social Security as a device for intergenerational sharing of aggregate risks. This type of risk-sharing is tangentially addressed in some of the studies reviewed above (e.g. the De Nardi et al. 1998 paper shows how different cohorts share the burden of a perfectly anticipated demographic shock), but at present there is no SGE model that can address the question of whether Social Security can provide a welfare enhancing transfer mechanism to help out a cohort that suffered the consequences of a sustained stock market crash or depression, which Blinder (1988) ascribed as one of the reasons motivating the introduction of Social Security in 1935.

A second limitation is the lack of detail on firm behavior. Most models abstract from modeling firms. The “firm” is synonymous with a single aggregate Cobb Douglas production function that earns zero profits and pays out all revenue to workers and capital owners at competitive rates equal to the marginal product of labor and capital, respectively. With the exception of a few studies such as Walliser (1997), there have been few attempts to model the
markets for insurance or annuities, or to model why firms offer fringe benefits such as health insurance or private pensions. Information asymmetries that lead to moral hazard and adverse selection problems in insurance markets have not been modeled at the same level of detail as the analytic approaches surveyed in section V.1. Incomplete markets are modeled only to the extent that certain markets are exogenously specified to exist or not to exist, but the pattern of available markets and contracts, such as pensions, are not allowed to emerge endogenously as a competitive outcome. Thus, the current generation of SGE models is unable to address the important question of the extent to which Social Security "crowds out" pensions, annuities and private insurance contracts.

A third limitation is the failure to model labor supply and consumption decisions jointly in the individual’s DP problem (the De Nardi, et al 1998 model is an exception). Instead labor supply is typically treated as an exogenous process, with a fixed retirement date as in the Storesletten, et.al paper. The latter paper allows for exogenous random variation in pre-retirement labor supply, in which case the dichotomous treatment of labor and consumption is slightly less objectionable, but the importance of jointly modeling labor supply, retirement and savings decisions has been recognized since Feldstein (1974). In particular, labor supply is an important mechanism for dealing with risk, at least for healthy, productive workers. If savings turn out to be lower than expected, the individual can delay retirement. The additional complexities and unexpected results from modeling consumption, labor supply and retirement as jointly endogenous decisions has already been discussed in section V.1.

Finally, most SGE models still abstract from “political risk” of changes in the Social Security system. Government policy is typically treated as exogenously specified, although recent specifications, as noted above, have endogenized policy formation using simple median voter models of the political economy of Social Security policymaking. Even remaining within the standard framework for policy analysis where the government is modeled as a “Stackelberg leader” that can commit to any given policy, we are still far from the point of being able to apply mechanism design theory in a systematic search for welfare maximizing or cost minimizing social insurance systems, subject to information and incentive constraints. In view of all this, it should come as no surprise that we are quite far from developing a “grand unified model” of social insurance.

A useful direction for work in the near term is to determine policy issues where it might be reasonable to abstract from general equilibrium feedbacks and thereby avoid the complexities inherent in solving SGE models. This survey suggests that general equilibrium feedbacks cannot be safely ignored for fairly radical changes to the Social Security system, such as privatization or changing from a pay-as-you-go system to a fully funded system. However, it is likely that there are questions for which general equilibrium feedbacks are second order. One example might be the issue of investing the Trust Fund in equities. If we believe that most individuals can undertake offsetting changes in their private portfolios and that the U.S. is increasingly a small player in a global equity market, then it seems that the impact on the equity premium of investing the Trust Fund in equities would be second order and might be safely ignored in policy analyses similar to what was done in the Feldstein, Rangeluova and Samwick (1999) study discussed in section IV.3. It is hard to determine in advance of actually solving an SGE model whether general equilibrium
feedbacks are likely to be important, but common sense and intuition are likely to be good guides. This is standard practice in academic models where certain variables are assumed to follow exogenous stochastic processes even though it is possible to argue that on a grand scale “everything is endogenous.” It is eminently reasonable to try to reduce the computational burdens underlying SGE models by identifying certain processes such as equity returns, trends in fertility, mortality, marriage and divorce that can be reasonably treated as exogenous stochastic processes, at least to a first approximation.

VI. CONCLUSIONS

This essay has surveyed a very rapidly growing academic literature that models how individuals use Social Security and other public and private mechanisms to cope with various types of risk. One indication of the speed at which this literature is growing is that over two thirds of the more than 150 papers in the references were written or published since 1997. Although it is hard to summarize this vast literature in a few sentences, a general conclusion that emerges from reading this literature is that Social Security has big effects on individual behavior and welfare, and on the structure and performance of the entire economy.

This chapter has contrasted the use of “short cut” methods for modeling risk and evaluating the impact of changes in Social Security policies with more formal dynamic programming (DP) and stochastic overlapping generations (SGE) models that endogenize behavioral and general equilibrium responses to policy changes. The latter models are descendants of the simple deterministic two and three period overlapping generations models due to Samuelson (1958). Samuelson considered a simple endowment economy with a perishable good and a fundamental market incompleteness, the inability to trade with unborn generations, and showed that the introduction of a pay-as-you-go Social Security made all generations better off. A subsequent model by Diamond (1965) introduced capital and labor, and showed that introducing a pay-as-you-go Social Security system “crowds out” private saving by transferring income from the young (who have a low marginal propensity to consume) to the old (who have a high marginal propensity to consume). This results in a lower capital stock and consumption and welfare levels compared to either a “privatized” system (i.e. no Social Security) or a fully funded system. However, Diamond’s model abstracted from risk and uncertainty and thus was not able to capture Social Security’s other key roles as a provider of multiple forms of social insurance. Because so many key aspects of Social Security were missing, the policy implications of the early deterministic models could not necessarily be trusted. Starting in the late 1970s, economists used probability theory, statistical decision theory, and the method of dynamic programming to model many of the risks and uncertainties that individuals face. They developed models of private insurance and identified informational problems that can lead to inefficiencies and failures in these insurance markets that motivated and rationalized the need for social insurance. There has been an especially large amount of research on pay-as-you-go social insurance systems similar to our Social Security program, which has been modeled as a mechanism for counteracting the informational problems in private insurance markets. Through this mechanism, Social Security offers potential individual welfare gains from improved risk-sharing, in addition to any gains in social welfare that result from its role in redistributing lifetime income.
Despite numerous attempts, however, economists have found it difficult to demonstrate the existence of clear cut welfare gains from the introduction of a pay-as-you-go social insurance program. Instead, providing a “social safety net” Social Security reduces individuals’ incentive to save, both for precautionary and life cycle reasons. Additional reduction in savings results from the fact that Social Security redistributes income from rich to poor, and from young to old, and both of the latter groups tend to have higher marginal propensities to consume. Most of the recent stochastic multiperiod overlapping generations models predict that the welfare losses due to reduced capital accumulation and consumption outweigh the welfare gains due to improved risk-sharing. Thus, the welfare losses due to the general equilibrium “crowding out” effect of Social Security that was identified in Diamond (1965) appears robust to a variety of ways of introducing uncertainty into SGE models. To paraphrase the title to a (1998) article by Alan Blinder, economists have not yet succeeded in providing a compelling answer to the question, “Why is the government in the social insurance business?”

One understudied aspect of Social Security that affects welfare is Medicare, which is a particularly popular component of Social Security. Partial equilibrium analyses such as Rust and Phelan (1997) are able to quantify the large improvements in welfare that this particular form of risk sharing provides. There are obvious problems in private markets for health insurance and health care that Medicare corrects, although at significant cost to taxpayers. It remains to be seen whether the Medicare program provides net increases in social welfare in a more complete general equilibrium model that accounts for the costs of running the program and its potential distortions on saving decisions and the production of health care.

The fact that the retirement part of Social Security is so politically popular suggests that these economic models are leaving out one or more crucial aspects of the program. The recent literature on the political economy of Social Security suggests that political support for a pay-as-you-go system similar to ours could arise from majority voting, even in a world where all generations know that the distortions it creates leads to a net reduction in social welfare. An initial population votes to initiate such a program because they receive a net windfall benefit from future generations; the present value of their benefits exceeds the present value of Social Security taxes. The system can be sustained in subsequent years even though it imposes a net cost on subsequent generations because at any given time there are just enough old people who will vote to continue it out of fear that their own contributions would be lost if the system were terminated. However, this equilibrium is a tenuous one and can be upset by demographic shocks such as a wave of baby boomers. But, as also discussed in this chapter, recent models show that it is politically possible to move from a pay-as-you-go to a privatized or fully funded system provided the right package of “recognition bonds” is introduced that effectively transfers enough of the welfare gains that future generations receive from privatizing or fully funding the system to the transition generations to convince them that the change is worthwhile.

The academic literature has created the relevant tools and conceptual framework for analyzing risk, uncertainty, and insurance. It has provided us with key insights into the role of social insurance and its political sustainability. Most of these results have been obtained using the DP/SGE approach to modeling described in section V of this chapter.
Some of the questions can be addressed with “short cut” approaches, but the discussion in this chapter points to several serious limitations with these approaches, particularly with respect to their inability to assess the effect of alternative policies on individual welfare and their failure to account for potentially large behavioral and general equilibrium feedback effects. DP and SGE models explicitly address both of these problems. The discussion in this chapter suggests that some DP and SGE models are becoming sufficiently realistic to be taken seriously for use in forecasting and policy analysis. However, the DP/SGE approach still suffers from important limitations that reduce the usefulness of these models for practical forecasting and policy analysis. These include:

1. the absence of a “grand unified model” of social insurance that provides a single consistent, integrated framework for policy evaluation,

2. the computational complexity of solving detailed dynamic programming and stochastic general equilibrium models,

3. the large data requirements necessary to identify and estimate the unknown parameters of these models,

4. the absence of good models of labor demand, including models of employers’ demand for older workers and their decision to offer various types of pensions, health insurance, and other types of fringe benefit packages; and

5. the lack of models that capture a rich array of macroeconomic shocks and the lack of work on endogenous determination of equity returns, particularly in a sufficiently realistic framework that could explain the large equity premium and high volatility of equities.

The last limitation is particularly important in relation to the recent interest in evaluating proposals to privatize Social Security. Although we are beginning to see some simple models of endogenously derived stock price volatility in overlapping generations models (see, e.g., Spiegel, 1998), few of the current generation academic models provide a satisfactory analysis of the role of Social Security as an intergenerational risk sharing mechanism in the face of persistent macroeconomic shocks. As a result, academic models cannot yet provide predictions of how privatization would affect the equity premium and the overall volatility of the stock market, or whether it would contribute to an “asset meltdown” if a large number of retiring baby boomers simultaneously tried to liquidate their retirement savings portfolios. Our understanding and modeling of firm behavior is also very weak, even though firms, particularly as employers, have a critical impact on the risks facing individuals. Firms provide job opportunities, pensions, health insurance, and a variety of types of private insurance contracts that can complement the goals of Social Security in providing a stronger safety net for all American citizens. In the absence of realistic models of firm behavior, we cannot determine whether public policies designed to insure against more types of risk may inadvertently “crowd out” complementary insurance provided by firms.
Rapid improvements in computer hardware and improvements in algorithms and numerical methods have enabled economists to overcome many of the shortcomings of existing models. However, it is important to note that despite the empirical success of a number of the DP and SGE models, the behavioral assumptions underlying most of these models are still a long way from being empirically justified. In particular, there are still many unresolved empirical puzzles connected with saving and wealth accumulation. We are years away from being able to solve and empirically estimate the full life-cycle model from birth until death with a realistic treatment of labor/leisure and consumption/saving decisions and to incorporate various additional choices such as educational and career choices and marital and childbearing decisions. Until then, we do not know whether this extended life cycle model could explain these puzzles or whether they will lead us to reject it in favor of an alternative representation of behavior.

Even focusing on the end of the life-cycle, we are still years away from being able to solve a reasonably complete DP model that incorporates the key components of Social Security, private pensions, and private insurance contracts with an adequate level of detail to be useful to evaluate many policy questions of interest to SSA. None of the existing DP or SGE models have a track record in policy forecasting that would enable us to determine whether they really provide more accurate forecasts than the models it currently uses, and in the short term computation and data constraints still make it prohibitively difficult and expensive to formulate and solve a wide range of more limited academic models that SSA might be interested in using for policy evaluation. In particular we are decades away from developing a comprehensive grand unified stochastic overlapping general equilibrium model of Social Security that could simultaneously handle all of the key risks and private insurance mechanisms listed in section III. If we are interested in the general equilibrium effects of certain policies, we must be willing to work with a patch work of special purpose models that employ relatively crude approximations to Social Security policy. Thus, while there has been a great deal of progress, it is clear that much that remains to be done in academic modeling of Social Security.
APPENDIX 5-A

RECOMMENDATIONS FOR AN SSA MODELING STRATEGY
by John Rust

Given this state of affairs, what is the best long term strategy for SSA to follow in order to develop "practical" policy models as quickly as possible? Clearly, SSA faces a much more pressing time schedule than does the academic research community: the “long term” within SSA is probably measured in years rather than in decades. In order to motivate my own conclusions and recommendations about the directions SSA should pursue in this area, it is useful to recall the recommendation made by the Panel on Retirement Income Modeling (PRIM) at the National Academy of Sciences

“Relevant agencies should consider the development of employer models and a new integrated individual-level microsimulation model for retirement income/income-related policy analysis as important long-term goals. Construction of such models would be premature until better data, research knowledge, and computational methods are available. To respond to immediate policy needs, agencies should use limited, special-purpose models with the best available data and research findings to answer specific policy questions.” (Citro and Hanushek, 1997, p. 9).

Since the time of this report there have been several significant changes that warrant modification of this recommendation. The government is no longer in a severe budget deficit and many government agencies including SSA have benefited from increased funding and autonomy and the possibilities this affords for taking a longer range view of how best to fulfill its mandate. Indeed, in recent years SSA has embarked on a very ambitious program of investing a significant amount of resources on both short and long term model development and on improved liaisons with the academic community. It has also invested significant resources in gathering new data, both for internal data sets such as the Disability Evaluation Survey, and for continuing public use panel data sets such as the Health and Retirement Study (which was previously almost exclusively supported by the National Institute on Aging). I think these are very positive developments and hope that this will continue, although in my opinion the level of SSA’s spending on data gathering, research, and model development is still too low given the potentially large payoffs to these investments that I noted in sections II and III of Chapter 5.

My specific recommendations for SSA’s long term modeling efforts are as follows:

1. It should substantially increase its overall support for policy modeling, both within SSA and via grants and contracts to academic researchers working on various aspects of Social Security,

2. It should invest in developing an in-house capability to work with a portfolio of specialized academic DP and SGE models similar to those described in section V.2 of chapter 5 of this report,

3. It should continue its pioneering work on the MINT and Projected Cohorts models, modifying and extending them to provide a complete “reduced form” overlapping generations model of the U.S. population that allows SSA to analyze the effect of uncertainty and the distributional consequences of alternative policies via stochastic simulations,

4. It should identify the key behavioral components of the models developed under point 3 that are most strongly affected by changes in Social Security policy and gradually replace them with structurally estimated models that provide an explicit expected utility based framework for handling uncertainty and provide endogenous behavioral responses to policy changes.

Recommendations 2 and 3 are consistent with the recommendations of the Panel on Retirement Income Modeling. Recommendation 2 encourages SSA to acquire and assimilate some of the more sophisticated academic analytical and modeling capabilities to complement the less behaviorally justified but more practically oriented microsimulation models that it is already familiar with. Recommendation 3 recognizes SSA’s pressing short term need to develop a capability to analyze policies that involve risky outcomes such as various plans for individual accounts or investing the Trust Fund in equities.

Recommendations 1 and 4 go farther than the Panel on Retirement Income Modeling was willing to go. That Panel was concerned about spending huge amounts of resources to develop a single large, monolithic microsimulation model that would ultimately be treated as a “black box” with underlying components that had little or no behavioral justification. I agree that it would be a mistake to embark on a crash program that attempts to build a grand unified model of social insurance, either now or at any time in the foreseeable future. However I do feel that we are now in an environment where there are sufficient resources to permit SSA to embark on a more ambitious, longer term evolutionary strategy of gradually transforming a non-behavioral microsimulation model such as MINT into a something that approximates a stochastic overlapping generations model that captures the key endogenous behavioral and general equilibrium feedbacks of Social Security policies.

There are a number of difficult conceptual issues associated with the implementation of recommendation 4 that I haven’t considered here, and also some significant risks. However I feel that the payoffs to continuing to extend and refine the MINT and/or Projected Cohorts Models are high. The immediate extensions that I would recommend would be to develop a simulation environment and a data structure that is sufficiently powerful to keep track of large populations of simulated individuals (and ultimately firms as well) from birth until death, including linkages between different simulated individuals that reflect relationships and events such as marriage and divorce, cohabitation, birth of children, changes in employment, and so forth. I think that SSA should attempt to build a model that follows individuals over their entire life cycle --- from birth
until death --- and explicitly accounts for key economic decisions and events such as education, occupation, and career choice; key demographic decisions such as marriage, divorce, and childbearing, as well as tracking decisions that are directly relevant to Social Security such as decisions to apply for OA, UI, and DI benefits, etc. The transition and choice probabilities underlying such a simulation model would initially be estimated via reduced-form statistical methods using existing panel data sets such as the SIPP and the HRS, just as has been done for the MINT model. The main difference is that the MINT model only focuses on a limited number of cohorts and does not attempt to follow individuals over their full life cycle or keep track of the relationships between the individuals that it simulates. Starting with an initial population distribution with overlapping generations of birth cohorts, the model would be run forward following each individual over their remaining life cycle, and tracking new additions to the population via births and immigration and reductions via deaths and emigration. The goal would be to produce a simulated population of agents that is as close to a statistical replica of the U.S. population as possible. Such a model will have both idiosyncratic shocks that affect individuals or families as well as macroeconomic shocks that affect all individuals in the population. A stochastic simulation model of this type would be able to incorporate the key micro and macro risks that SSA is concerned about modeling.

The big open question is whether an evolutionary approach of gradually replacing various reduced form transition and choice probabilities in a microsimulation model such as MINT will ultimately result in something that approximates an SGE model that would be credible for use in policy forecasting exercises. The logic of solving general equilibrium and life-cycle optimization problems suggests that all parts of these are interconnected, and therefore one cannot generate reliable predictions if we are only willing to model parts of these problems and rely on reduced-form behavioral relationships for other parts of the model. However many general equilibrium and life cycle models do have a recursive structure that can be exploited, where no major harm is done in treating certain variables in the model as state variables (as opposed to control variables) that evolve according to exogenously specified stochastic processes. For example even though marriage and child bearing decisions may be endogenous within the overall lifecycle problem, toward the end of the life cycle it becomes more reasonable to treat family status as a state rather than as a decision, especially if empirical work confirms that Social Security policies do not have a big impact on these decisions. Then family status can be modeled as an exogenous stochastic process (estimated via reduced form methods), allowing the structural modeler to focus on endogenously deriving the decision rules for those variables that are more directly affected by Social Security policy (e.g. hours of work and the timing of retirement). So the hope is that there is sufficient recursive structure in the world that an evolutionary approach of gradually replacing various modules of an initial short cut simulation model with behaviorally justified modules derived from structural econometric estimates will result in a more accurate and credible model.
I close with two final observations. First, although it is always desirable to develop more realistic models, they will always be abstractions and it will never be possible or desirable to remove the role of human judgement in forecasting and policy analysis. Second we should keep in mind that perhaps the most important economic insights about the effects of Social Security have come from relatively small, seemingly unrealistic models such as Samuelson’s (1958) seminal overlapping generations model and its subsequent extensions. While the challenge of trying to develop a single grand unified model of social insurance is tempting, it doesn’t seem to be a realistic goal at the present time. There is a real danger that an attempt to build such a model would result in a very cumbersome monolithic “black box” whose predictions could be hard to understand and explain. It is hard to imagine that the predictions of such a model would have much credibility. The problems that an ambitious pioneer such as my Yale colleague Guy Orcutt experienced in the 1960s in his attempt to build a macro model of the U.S. economy from the “bottom up” via aggregation of micro models of individual agents serve as a useful reminder of the risks inherent in attempts to do “big science.” While I think it is worthwhile to develop increasingly comprehensive and realistic policy models (and this is clearly the direction in which the economics profession is moving), it is equally important to support a “cottage industry” of numerous smaller, less ambitious modeling efforts that historically have been the source of some of the most creative and important ideas.
REFERENCES TO CHAPTER 5


*Economic Inquiry* 13-3, 373-388.


251


