Essays on Annuity Valuation, Bequests and Social Security

by

Alain Jousten

Licencié en Science Economique, Université de Liège (1994)

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1998

© Alain Jousten, 1998. All rights reserved

The author hereby grants to MIT permission to reproduce
and to distribute publicly paper and electronic
copies of this thesis document in whole or in part.

Signature of Author .................................................. Department of Economics April 16, 1998

Certified by ...................................................................... Peter A. Diamond
Institute Professor
Thesis Supervisor

Certified by ...................................................................... Jonathan Gruber
Professor of Economics
Thesis Supervisor

Accepted by ...................................................................... Peter Temin
Elisha Gray II Professor of Economics
Chairperson, Departmental Committee on Graduate Students

JUN 09 1998
Essays on Annuity Valuation, Bequests and Social Security

by

Alain Jousten

Submitted to the Department of Economics
on April 16, 1998, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Chapter 2: This paper analyzes the implications of introducing a linear bequest motive into
a standard life-cycle model, both allowing for credit and annuity market imperfections. First,
we characterize the consumption and wealth processes. We find that the consumption profile
is non-increasing in the linear bequest parameter for the simplest certainty case, but that the
same is not true when allowing for life-span uncertainty. Second, we study the issue of annuity
valuation in the presence of annuity market imperfections. We find that for a sufficiently strong
bequest motive, the value of an annuity is equal to its actuarial value. This invalidates a
previous claim that in the presence of imperfect annuity markets, the true value is close to the
simple financial value.

Chapter 3: We generalize the standard joy-of-giving bequest motive by including inter-vivos
gifts. Within a life-cycle framework, we analyze the implications of the choice of different
discount factors for the utility of gifts and bequests. For a linear utility of giving, we characterize
the gift and bequest pattern of a liquidity constrained individual over the life-cycle. We find
that discounting at the interest rate is very interesting as the linear utility parameter can be
interpreted as a summary measure for the strength of the motive of giving, net of all gift and
bequest timing issues over the life-cycle.

Chapter 4: This paper focuses on Social Security benefit claiming behavior, a take-up decision
that has been ignored in the previous literature. Using financial calculations and simulations
based on an expected utility maximization model, we show that delaying benefit claim for a
period of time after retirement is optimal in a wide variety of cases and that gains from delay
may be significant. We discuss the implications of our findings for the large literature on Social
Security.

Thesis Supervisor: Peter A. Diamond
Title: Institute Professor

Thesis Supervisor: Jonathan Gruber
Title: Professor of Economics
Für meine Eltern
Acknowledgments

I am very grateful to my advisors Peter Diamond, Jon Gruber and Jim Poterba, who were always available for discussions and for advice. Their willingness to read through my numerous drafts was almost endless! By hiring me as a research assistant in my second year, Peter and Jon helped me over a period of serious financial problems. The lengthy discussions I had with Jon over the last few years sharpened my thoughts and his comments proved instrumental in improving the writing of my thesis. I am happy that he let me share his excitement for, and expertise in, research and teaching. A special 'thank you' also goes to Jim who 'jumped in' after Jon's departure for Washington, and de facto became my third advisor. He proved to be an invaluable resource both for the thesis and on the job market. I am particularly indebted to Peter, who was extremely supportive over these last few years. I really enjoyed our long and numerous discussions, which truly broadened my horizons. I am confident to say that he is the one individual who shaped the way I see and understand economics. I would also like to thank Pierre Pestieau. He first talked to me about the idea of coming to the US for graduate school, and I am convinced that this was one of the best decisions of my life. Over the last few years, he has proven to be a constant and reliable source of support and advice that I would not want to miss.

Further, I thank all those who supported me financially during these four years. The Belgian American Educational Foundation (BAEF) paved me the way to the US by generously awarding me a Francqui Foundation Fellowship for my first year of graduate studies. The support of the NBER through a Fellowship in the Economics of Aging, the CREPP, MIT, as well as the Deutschsprachige Gemeinschaft Belgiens are also gratefully acknowledged.

Many thanks also to all my friends and fellow students at MIT, who made my life in Boston a more enjoyable experience. In particular, I want to mention Christian Frois, Markus Möbius, David Myatt, Jaime Ortega, Valéry Panier, Koen Vermeylen and Ksenia Yudaeva. Very special thanks go to my dear friend Maciej Dudek, who has been an excellent dinner companion over the last two years. I will never forget our numerous and sometimes endless discussions about 'life' which proved to be so much fun.
But there is little doubt in my mind that I owe by far the most to my family. My parents, my brother René and my sister Nicole have always let me enjoy their unwavering support in whatever I decided to do. I am convinced that without their help and encouragements, I would never have been able to obtain this Ph.D. My parents made every possible effort to give me a good education, as well as a chance to fulfill my dreams, even though this meant huge financial sacrifices for them. I am happy to say that one of my dreams just became true. It is therefore to my parents that I dedicate this thesis.
Contents

1 Introduction ............................................... 11

2 Life-cycle modelling of bequests and their impact on annuity valuation 13
  2.1 Introduction ............................................ 13
  2.2 Certainty model ......................................... 15
    2.2.1 Kuhn-Tucker problem ............................. 16
    2.2.2 Summary and discussion ........................... 19
  2.3 The life-span uncertainty model ......................... 20
    2.3.1 No liquidity constraints ......................... 22
    2.3.2 Liquidity constraints ............................... 28
    2.3.3 Comments on the uncertainty case ............... 37
  2.4 Annuity valuation ...................................... 39
    2.4.1 Case 1: Interior solution, \( \frac{\rho - r}{\alpha - 1} < r, g < r \), budget constraint not binding . 40
    2.4.2 Case 2: Liquidity constraints binding at time \( E, \frac{\rho - r}{\alpha - 1} < r, g < r \), budget constraint slack ................................................. 41
    2.4.3 Case 5: Budget constraint binding at infinity, \( \frac{\rho - r}{\alpha - 1} > g > \frac{\rho + p - r}{\alpha - 1} \) ............... 42
    2.4.4 Case 7: Budget constraint binding at infinity, \( \frac{\rho + p - r}{\alpha - 1} > g \) ......................... 43
    2.4.5 Summary of the annuity results .................... 43
  2.5 Conclusion .............................................. 44

3 Timing of Gifts and Bequests .......................... 46
  3.1 Introduction ............................................ 46
  3.2 The framework ......................................... 47
List of Figures

2-1 Different scenarios ................................................. 17
2-2 $C_t$ Profile, Budget Constraint binding .................. 25
2-3 $W_t$ Profile, Budget Constraint binding ................. 25
2-4 $C_t$ Profile, Interior solution ................................. 31
2-5 $W_t$ Profile, Interior solution ................................. 31
2-6 $C_t$ Profile, Constraint binding at $E$ ...................... 32
2-7 $W_t$ Profile, Constraint binding at $E$ ...................... 32
2-8 Effect of $b$ on consumption profile ......................... 34
2-9 $C_t$ Profile, Constraint binding at $t > E$ ................ 36
2-10 $W_t$ Profile, Constraint binding at $t > E$ ............ 36
2-11 $C_t$ Profile, $\frac{e^t + p - r}{a - 1} > g$ .................... 37
2-12 $W_t$ Profile, $\frac{e^t + p - r}{a - 1} > g$ .................... 38

3-1 Different scenarios for $\xi > r$ ................................. 54
3-2 Different scenarios for $\xi \leq r$ ............................... 57

4-1 EPDV by Months of Delay, One-Earner Couple ............. 72
4-2 EPDV by Months of Delay, Single Worker ................. 73
4-3 EPDV by Months of Delay, Two-Earner Couple ............. 76

A-1 $C_t$ Profile, Constraint binding at $t > E$ .................. 84
A-2 $W_t$ Profile, Constraint binding at $t > E$ .................. 85
A-3 $C_t$ Profile, Constraint binding at $E$ ...................... 86
A-4 $W_t$ Profile, Constraint binding at $E$ ...................... 86
A-5 $C_t$ Profile, $\frac{\rho t \rho - r}{\alpha - 1} > g$, Constraint binding at $E$ ............................. 87
A-6 $W_t$ Profile, $\frac{\rho t \rho - r}{\alpha - 1} > g$, Constraint binding at $E$ ............................. 87
List of Tables

2.1 Effect of $\beta$ on the consumption and wealth levels .................................. 19
2.2 Time pattern of $\hat{c}$ .................................................................................. 26
2.3 Summary of the different possible scenarios ..................................................... 29

4.1 Financial Calculations: One-Earner Couple and Single Worker ..................... 70
4.2 Financial Calculations: Two-Earner Couple ................................................... 75
4.3 Mortality Risk and Discount Rate Required for Optimal Claiming at 62 ........... 77
4.4 Expected Utility Maximization: Single Worker .............................................. 78
Chapter 1

Introduction

Retirement income programs have long been recognized as important determinants of individual behavior. As a result, a large empirical and theoretical literature has developed, analyzing the impact of these programs on decisions such as the date of retirement, or alternatively the life-cycle consumption and savings pattern. In some of the empirical work, authors have explicitly tried to take bequest motives into account when specifying their models. But unfortunately, a theoretical analysis of gift and bequest motives and their implications for empirical work has largely been missing from the literature. There has only been very little work directed at studying the impact of a motive of giving on the consumption, gift and bequest patterns over the life-cycle, as well as on the value an individual attaches to an annuity payout stream, such as social security payments. In a first part of the thesis, we present two models that try to fill this gap in the literature.

The first essay, chapter 2, analyzes the impact of a bequest motive on the behavior of the giver. We model the motive of giving by inserting a linear utility of bequests function of the joy-of-giving type to a standard life-cycle model. We allow for both credit and annuity market imperfections. First, we characterize the consumption and wealth profiles towards the end of the life-cycle. We find that the consumption profile is non-increasing in the linear bequest parameter for the simplest certainty case, but that the same is not true when allowing for life-span uncertainty. That is, consumption at some ages can be higher for a person with a stronger bequest motive. This finding contradicts a common belief, that consumption levels are monotonic in the strength of the bequest motive. Second, within the same framework, we study
the issue of annuity valuation by individuals in the presence of annuity market imperfections. We find that for a sufficiently strong bequest motive, the value of an annuity is equal to its actuarial value. This invalidates a previous claim that in the presence of imperfect annuity markets, the true value is close to the simple financial value.

Chapter 3 generalizes the standard joy-of-giving bequest motive by including inter-vivos gifts. Within a life-cycle framework, we analyze the implications of the choice of different discount factors for the utility of gifts and bequests. For a linear utility of giving, we characterize the gift and bequest pattern of a liquidity constrained individual over the life-cycle. We find that discounting at the interest rate, such as used in chapter 2, is very interesting as the linear utility parameter can be interpreted as a summary measure for the strength of the motive of giving, net of all gift and bequest timing issues over the life-cycle.

The second part of the thesis, chapter 4, is an extract of Coile, Diamond, Gruber and Jousten(1997). This essay focuses on Social Security benefit claiming behavior, a take-up decision that has been ignored in the previous literature. Using financial calculations and simulations based on the expected utility maximization model presented in chapter 2, we show that delaying benefit claiming for a period of time after retirement is optimal in a wide variety of cases and that gains from delay may be significant. For the case of the utility maximization model with liquidity constraints, we further show that optimal delays follow an inverse u-shape pattern as wealth rises. For individuals with low wealth, liquidity constraints limit their ability to delay the onset of the Social Security benefit stream. For households with high wealth levels, bequest motives play an important role and reduce the incentives to annuitize the household income. We conclude this chapter by discussing the implications of our findings for the large literature on Social Security.
Chapter 2

Life-cycle modelling of bequests and their impact on annuity valuation

2.1 Introduction

Bequest motives have long been recognized as potentially important determinants of saving patterns. Surprisingly, there has been very little work directed at describing how they affect optimal consumption patterns. We add a utility of bequests term of the joy-of-giving type to a standard life-cycle model.\footnote{This modelling strategy follows Yaari(1965), Fischer(1973), Friedman-Warshawsky(1988) and Hurd(1989).} We use a quasi-linear utility function in consumption and bequests and allow for both life-span uncertainty and liquidity constraints. We study the consumption and wealth profiles towards the end of the life-cycle. We focus on the behavior of people who have retired from active work and who are or will be eligible for some form of annuitized government benefits.

Within our framework we analyze two questions: First, we characterize the consumption and wealth profile of the elderly in the presence of a bequest motive. Even though joy-of-giving models have already been used before, a characterization of their impact on the consumption and wealth profile, both in a setup with and without liquidity constraints, is missing. Second, we analyze the implications for the valuation of annuities under credit constraints.

The liquidity constraints analyzed mimic the U.S. law, which prohibits the use of social
security benefits as collateral. The importance of liquidity constraints is indicated inter alia by the results of Hausman and Paquette (1987). The authors found strong empirical evidence of jumps in consumption levels at the time involuntary early retirees become eligible for retirement benefits. The finding of consumption jumps at the time of first eligibility implies that the legal limitations actually matter, and that the market does not completely offset them through private arrangements.

We find that the presence of a bequest motive has strong implications for the valuation of annuity contracts. Previously, Bernheim (1987) claimed that in the presence of liquidity constraints, the value of a marginal dollar of annuity payouts is close to the simple discounted value of future payouts. With sufficiently strong bequest motives, this does not apply: the true value of a marginal annuity payout stream is closer to the actuarially correct value, which takes into account both the interest rate and the survival probabilities. Annuity valuation is of considerable importance for some of today's most acute policy questions, particularly the evaluation of reforms in the area of old-age income provision. For example, to evaluate and understand the implications of a move from a public annuity based retirement income systems towards a private pension savings system, a correct measure of the value of annuity holdings is crucial.

Our paper is divided into two main parts. First, we characterize the implications of a bequest motive on consumption and wealth. In section 2.2, we start by presenting a certainty model with potentially binding liquidity constraints. We find that consumption levels are non-increasing in the strength of the bequest motive. Section 2.3 extends the analysis to the case of a life-span uncertainty model, both with and without liquidity constraints. We find that the previous result does not generalize to this more realistic setup. That is, consumption at some ages can be higher for a person with a stronger bequest motive. Further, we establish that under life-span uncertainty, the rate of growth of consumption is non-decreasing in the strength of the bequest parameter. Section 2.4 constitutes the second part of the paper, where we analyze the question of annuity valuation in the presence of a bequest motive. Section 2.5 contains some concluding remarks.
2.2 Certainty model

Consider an individual who lives for two periods \((t = \{0, 1\})\). Suppose that we can represent his utility function by

\[
U(C_0, C_1, B_2) = u(C_0) + \frac{1}{1 + \rho} u(C_1) + \beta \frac{B_2}{(1 + r)^2}
\]  

(2.1)

where the first two terms correspond to the standard additively separable utility of consumption \((C_0\) and \(C_1\) terms. We suppose that the per period utility of consumption function is strictly concave, and that \(\lim_{C \to 0} u'(C) = \infty\). \(^2\) \(r\) denotes the real interest rate the individual faces on the capital markets and \(\rho\) is the discount rate for utility generated by consumption.

The third term in expression 2.1 captures the utility generated by bequests that we assume to be realized at the time of death and to be linear in the present discounted value of bequests \((B_2)\) with the linear parameter \(\beta\). There are several reasons for choosing this linear form. \(^3\) The first reason is tractability, as it would be more difficult to derive clear predictions in a general setup. Second, the quasi-linearity of the utility function of consumption and bequests captures the reasonable assertion that people are less risk averse with respect to bequests than with respect to their own personal consumption.

Further notice that we use different discount factors for consumption \((\rho)\) and bequests \((r)\). In chapter 3 we show that the choice of \(r\) as the discount rate for bequests is very convenient: First, a linear bequest motive with discount rate \(r\) is equivalent to a more general linear motive allowing for both a gift and a bequest motive. Second, this particular setup allows us to abstract away from timing issues with respect to these gifts and bequests.

The use of the interest rate \(r\) for discounting utility of bequests further allows us to reinterpret the utility-of-bequests term in the utility function. Instead of viewing it as a discounted period two utility, we can see the bequest terms as a linear function of the present discounted value of bequests, which is our preferred interpretation in this paper.

Suppose that the individual has an initial wealth \(W_0\) and is entitled to an income \(Y_1\) in period 1. Further suppose that the individual faces credit constraints.

\(^2\) The strict concavity has been chosen for pure reasons of simplicity.

\(^3\) For a detailed discussion of the quasi-linear utility function in consumption and bequests, see Joulz (1997).
We can write the consumer's problem as the following constrained optimization problem.

$$\max_{C_0, C_1, B_2} \left( u(C_0) + \frac{1}{1+\rho} u(C_1) + \frac{\beta}{(1+r)^2} B_2 \right)$$

s.t.

$$S \equiv W_0 - C_0 \geq 0 \quad (2.2)$$

$$B_2 \geq 0 \quad (2.3)$$

$$W_0 + \frac{Y_1}{1+r} = C_0 + \frac{C_1}{1+r} + \frac{B_2}{(1+r)^2} \quad (2.4)$$

where the first constraint represents the credit constraint at the end of the first period and the second the global budget constraint. We further impose an explicit non-negativity constraints on bequests (expression 2.3). The assumption of having marginal utility tend to infinity as consumption levels tend to zero takes care of the non-negativity of the two consumption variables $C_0$ and $C_1$.

### 2.2.1 Kuhn-Tucker problem

Using the Kuhn-Tucker method, we can find the following optimality conditions

$$\begin{cases}
    u'(C_0) = \lambda_1 + \frac{1+r}{1+\rho} u'(C_1) \\
    \beta = \frac{1+r}{1+\rho} u'(C_1) - \lambda_3 \\
    W_0 - C_0 \geq 0 , \quad \lambda_1 \geq 0 , \quad \lambda_1 (W_0 - C_0) = 0 \\
    W_0 + \frac{Y_1}{1+r} = C_0 + \frac{C_1}{1+r} + \frac{B_2}{(1+r)^2} \\
    B_2 \geq 0 , \quad \lambda_3 \geq 0 , \quad \lambda_3 B_2 = 0
\end{cases}$$

where $\lambda_1$ and $\lambda_3$ are the Kuhn-Tucker multipliers associated with the constraints 2.2 and 2.3 respectively.

We can regroup the possible optima into four different scenarios depending on the binding or the slackness of the two inequality constraints. We split the analysis of the optimality conditions in two parts. First we derive equations for the borders delimiting these four scenarios depending on which constraints are binding. Then, in a second step, we characterize the four possible scenarios.
The borders

We derive the equations of the expressions delimiting the different scenarios in terms of the bequest parameter $\beta$ and the period 1 income level $Y_1$ keeping total lifetime income $R \equiv W_0 + \frac{Y_1}{1+r}$ constant. Noticing that when a constraint switches from slackness to binding, both the constraint and the associated Kuhn-Tucker multiplier equal 0, we can derive equations for the expressions that delimit the border between the four scenarios. As we will see, the constraints will have the shape shown in figure 2-1.

Figure 2-1: Different scenarios

The expression separating the binding and slackness of the credit constraint, can be written as

$$
\begin{cases}
Y_1 = Y_1^* & \text{for } \beta < \beta^* \\
\beta = u'(R - \frac{Y_1}{1+r}) & \text{for } \beta \geq \beta^*
\end{cases}
$$

(2.5)

where $Y_1^*$ is the implicit solution to $u'(R - \frac{Y_1}{1+r}) = \frac{1+r}{1+r} u'(Y_1)$ and $\beta^*$ is the solution to $\beta = \frac{1+r}{1+r^*} u'(Y_1^*)$. Notice that because of our assumption of unbounded marginal utility at a zero consumption level, $Y_1^*$ is strictly positive.

Characterizing this function 2.5, it is easy to see that $Y_1$ is non-decreasing in $\beta$ because of the sign of $u''(.)$. Further, depending on the sign of the third derivative of the utility function, the expression is either concave or non-concave. For example, for the case of a CRRA utility function, the function is concave.
Similarly, the expression for the bequest constraint is

$$\begin{cases} 
\beta = \beta^* & \text{for } Y_1 \leq Y_1^* \\
\beta = \frac{1+r}{1+\rho} u'(Y_1) & \text{for } Y_1 > Y_1^* 
\end{cases} \tag{2.6}$$

This functional form implies that $Y_1$ is decreasing in $\beta$ and that the concavity depending on the sign of $u''(.)$. For the case of a CRRA utility function, the function 2.6 is non-convex.

Figure 2-1 illustrates the borders and the four scenarios they delimit. Equation 2.6 is represented by the dotted line, and equation 2.5 by the continuous line. The dashed vertical line at $R(1+r)$ represents total income expressed in period 1 units.

The four scenarios

**Scenario 1: No constraint binding**

We start by checking the benchmark case of having both the bequest and liquidity constraint not binding, i.e., $\lambda_1 = 0$ and $\lambda_3 = 0$. In this case, we have that $u'(C_0) = \frac{1+r}{1+\rho} u'(C_1) = \beta$. Hence, $\frac{dC_0}{d\beta} < 0$ and $\frac{dC_1}{d\beta} < 0$. Further it is easy to see that $\frac{dS}{d\beta} > 0$ and that $\frac{dB_2}{d\beta} > 0$.

**Scenario 2: Credit constraint binding**

The optimum is characterized by the following conditions:

$$\frac{1+r}{1+\rho} u'(C_1) = \beta \quad \text{and} \quad W_0 = C_0.$$ 

These two conditions imply that $\frac{dC_0}{d\beta} = 0$ and $\frac{dC_1}{d\beta} < 0$, as well as $\frac{dS}{d\beta} = 0$ and $\frac{dB_2}{d\beta} > 0$. A third relation that holds at the optimum shows us that the liquidity constraint also becomes less binding. Indeed, the above results combined with $u'(C_0) = \lambda_1 + \frac{1+r}{1+\rho} u'(C_1)$ give us $\frac{dA_1}{d\beta} < 0$, which means that the binding constraint becomes less costly in terms of utility. Notice further that this effect on $\lambda_1$ alleviates the distortion in the inter-period consumption allocation that was caused by the liquidity constraints.

**Scenario 3: Bequest constraint binding**

A third possible scenario is when constraint 2.3 is the only binding constraint. We hence have $\lambda_3 > 0$. In this case, the allocation of resources between consumption in the first and second period follow the standard equation $u'(C_0) = \frac{1+r}{1+\rho} u'(C_1)$. Further, the budget constraint 2.4 simplifies to $W_0 + \frac{Y_1}{1+r} = C_0 + \frac{C_1}{1+r}$. It is thus trivial to find $\frac{dC_0}{d\beta} = \frac{dC_1}{d\beta} = \frac{dS}{d\beta} = \frac{dB_2}{d\beta} = 0$ and that $\frac{dA_3}{d\beta} < 0$ meaning that the constraint becomes less binding.

**Scenario 4: Credit and bequest constraints binding**
<table>
<thead>
<tr>
<th>Scenario</th>
<th>Description</th>
<th>$\frac{dC_a}{d\beta}$</th>
<th>$\frac{dC_b}{d\beta}$</th>
<th>$\frac{dS}{d\beta}$</th>
<th>$\frac{dB}{d\beta}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>no constraint binding</td>
<td>$&lt;0$</td>
<td>$&lt;0$</td>
<td>$&gt;0$</td>
<td>$&gt;0$</td>
</tr>
<tr>
<td>2</td>
<td>credit constraint binding</td>
<td>0</td>
<td>$&lt;0$</td>
<td>0</td>
<td>$&gt;0$</td>
</tr>
<tr>
<td>3</td>
<td>bequest constraint binding</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>both constraints binding</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2.1: Effect of $\beta$ on the consumption and wealth levels

A fourth and last possibility is that both the constraint 2.2 and the constraint 2.3 bind at an optimum. It is trivial to see that $\frac{dC_a}{d\beta} = 0$ and $\frac{dC_b}{d\beta} = 0$. The effect on the wealth levels can also easily be determined to be $0$. Further note that we have $\frac{dA_\beta}{d\beta} < 0$ and $\frac{dA_1}{d\beta} = 0$, which means that the constraint that is affecting the allocation between consumption in the second period and bequests becomes less binding. Hence, the optimal choice becomes less distorted.

### 2.2.2 Summary and discussion

Table 2.1 summarizes the above findings concerning the impact of the bequest parameter $\beta$ on the consumption, savings and bequest levels. Notice that, within each scenario, consumption is monotone non-increasing in $\beta$, whereas wealth levels are monotone non-decreasing in $\beta$. A corollary of these findings is that a marginal increase in the parameter $\beta$ makes the binding of the liquidity constraints less costly. The intuitive reason is that, as the marginal utility out of bequests becomes larger, consumption becomes relatively less attractive, hence pushing its level down. Figure 2-1 also indicates the comparative statics across scenarios as we increase the bequest parameter $\beta$. For sufficiently strong increases in $\beta$, we switch from the two scenarios, where either the credit or the bequest constraint binds, into the unconstrained scenario. Supposing the initial optimum lies in the region where both constraints bind, as we increase $\beta$, we first move into the region where only the credit constraint binds and ultimately into the unconstrained region. These comparative statics findings, both within a given scenario and between scenarios, allow us to say that if we increase $\beta$ sufficiently, we will have an optimal consumption profile such that both constraints are slack.

Finally, notice that the case of $\beta = 0$, i.e., of no bequest motive, fits nicely into the current analysis. Looking at figure 2-1 and table 2.1, it is easy to see that the case of $\beta = 0$ can be seen as being nested in our scenarios 3 and 4. Similarly, we would like to emphasize that the no-
liquidity-constraints case is perfectly integrated in the preceding analysis. Taking the extreme example of $Y_1 = 0$, it is trivial to see that the credit constraint in our general formulation of the problem will be slack as $Y_1^* > 0$.

2.3 The life-span uncertainty model.

We now allow for life-span uncertainty. A tractable way of modelling life-span uncertainty is to use an infinite horizon continuous time model with a constant survival probability.

The individual maximizes his expected utility of the consumption ($C_t \equiv C(t)$) and bequest processes ($B_t \equiv B(t)$) over his entire and potentially infinite lifetime. He faces uncertainty about his life-span under the form of a constant instantaneous survival probability $(1-p)$. As before, we assume that the utility of consumption is additively separable in time and that the utility of bequests enters as a linear function with the marginal utility out of bequests parameter $\beta$. Defining $\rho$ as the instantaneous utility of consumption discount rate and using the instantaneous real interest $r$ as the discount rate for computing the present discounted value (PDV) of bequests, we can write the utility function as

$$EU = \int_0^\infty u(C_t)e^{-(\rho+r)t}dt + \frac{p}{1-p}\beta \int_0^\infty B_te^{-(\rho+r)t}dt$$

We assume that the individual's instantaneous utility function is of the CRRA type with a coefficient of relative risk aversion equal to $(1- \alpha)$. We can hence rewrite the instantaneous utility function as $u(C_t) = \frac{1}{\alpha}C_t^\alpha$.\(^4\)

Using the notational shortcut $b = \frac{p}{1-p}\beta$, we can rewrite the objective function as

$$EU = \frac{1}{\alpha} \int_0^\infty C_t^\alpha e^{-(\rho+r)t}dt + b \int_0^\infty B_t e^{-(\rho+r)t}dt$$

\(^4\)Using a utility function with bounded marginal utility at the zero consumption point, combined with exponential discounting and a potentially infinite lifetime, would give us the rather unappealing result that optimal consumption levels reach zero in finite time, and stay at zero forever after.
As Bernheim(1987), we assume that the individual owns an initial financial wealth $W_0$ and is entitled to an income flow under the form of an annuity stream $Y_t$ that starts at time $E$ and that grows at rate $g$

$$Y_t = \begin{cases} 
0 & \text{for } t < E \\
Y_E e^{g(t-E)} & \text{for } t \geq E
\end{cases} \tag{2.7}$$

which allows us to write the budget constraint as

$$\int_0^\infty (C_t - Y_t) e^{-rt} dt \leq W_0 \tag{2.8}$$

The annuity income can be thought of as pension payouts or social security benefits that start at age of first eligibility $E$. In this interpretation, it is most plausible to consider cases where $g \leq 0$. $g = 0$ represents the case of a real annuity flow that is indexed for changes in the consumer price index, such as for example U.S. social security benefits. $g < 0$ represents the case of a nominal annuity payout stream, such as they are more common in private pension contracts. More generally, as already noted in the introduction, we prefer to think of the present model as a model of consumption and bequests after retirement. Several assumptions in our setup reflect this interpretation. First, we assume a constant bequest parameter $b$ over the entire life-span we consider. Some people may argue that the bequest motive probably varies over the life-cycle, with a stronger bequest motive when old than when young. By renormalizing time 0 to be the age of (early) retirement, we implicitly take this possible criticism into account. Second, we do not explicitly allow for labor income in the present model. This does not mean that our setup is incompatible with a model of labor income. In fact, we can reinterpret time 0 as the beginning of the working life and time $E$ as the moment the individual both retires and starts claiming benefits. Assuming additive disutility of work and inelastic labor supply before retirement, we can view $Y_t$ as taking into account both retirement income, labor income and disutility of work. Third, our model is well suited for studying the consumption and wealth decumulation behavior of early retirees. Thinking of time 0 as the time of retirement, and of time $E$ as the time of first benefit claiming, it becomes clear that our income process $Y_t$ allows

---

\(^5\)In the present paper, we do not consider the issue of measurement error in the CPI and its implications for social security benefits, and hence assume it measures the changes in prices correctly.
for the possibility of having an early retirement period from 0 to $E$ during which the old-age income level is zero.

In our analysis we assume that the individual takes the annuity income stream as exogenous and that it is impossible for the individual to vary annuity wealth holdings on the margin. This assumption, even though it may look very restrictive is in our opinion quite close to reality. Indeed, annuity holdings are very often largely composed of social security payments and pension benefits. These payments are rather lumpy for the individual as he has only a very limited ability to adjust his holdings on the margin once he is enrolled in a particular pension or social security system. An example of a possibility for individuals to have some flexibility on the margin is discussed in chapter 4: The U.S. social security system allows individuals to adjust annuity wealth holdings on the margin through strategic benefit claiming delays. These claiming delays may even imply better than actuarially fair "prices" depending on marital status and life expectancy. But apart from this type of exceptions, there is a rather limited potential for marginal variations in annuity wealth holdings.

The analysis of the present section addresses two questions: What is a general characterization of the consumption profile? What comparative statics results can we derive on the effect of the bequest parameter $b$ on the consumption and wealth levels at any time $t$?

As in the two-period certainty model, there are different types of solutions depending on which constraints are binding at the optimum. For our analysis, we can group these different solution types into two big categories depending on the presence or absence of explicit liquidity constraints.

### 2.3.1 No liquidity constraints

We start by analyzing the case of no liquidity constraints. The easiest way to illustrate the impact of the bequest parameter $b$ on the optimum, in the absence of credit constraints, is to assume that there are no annuities, i.e., $Y_E = 0$. Defining wealth and bequest levels at any time $t$ as

$$W_t \equiv B_t \equiv e^{rt}W_0 - \int_0^t C_r e^{r(t-r)}dr$$

---

6By delaying claiming optimally, the individual can increase his social security wealth by up to 5%, depending on his marital status and the life-expectancy of the household members.
we can rewrite the optimization problem in the absence of annuity income as

\[ \max_{C_t} EU = \frac{1}{\alpha} \int_0^\infty C_t^\alpha e^{-(p+\rho)t} dt + b \int_0^\infty W_t e^{-(p+r)t} dt \]  
\hspace{10cm} (2.9)

s.t.

\[ \int_0^\infty C_t e^{-rt} dt \leq W_0 \]  
\hspace{10cm} (2.10)

It is easy to see that there are two possible solutions, depending on whether \( \lim_{t \to \infty} W_t e^{-rt} \) is zero or positive. Notice that the transversality condition (TVC) plays the role the bequest parameter \( B_2 \) played in the two-period certainty model of the previous section.

**Interior solution**

Deriving the optimal consumption profile

\[ C_t^{opt} = \left( \frac{p}{b} \right)^{\frac{1}{1-\alpha}} e^{\frac{p-r}{\alpha-1} t}, \ \forall t \in [0, \infty[ \]  
\hspace{10cm} (2.11)

we find that the consumption process is only a function of the growth parameters \( p \) and \( r \) as well as of the linear bequest parameter \( b \), and that it is completely independent of the precise wealth level. Further notice that the rate of change of consumption is constant, i.e., \( \frac{\dot{C}_t}{C_t} = \frac{\rho - r}{\alpha - 1} \) and that it is higher than what it would be for the case of no bequest motive when it would be \( \frac{\rho + p - r}{\alpha - 1} \).\footnote{See table 2.2 below, where this latter result becomes obvious when setting \( b = 0 \).}

The wealth process follows

\[ W_t = e^{rt} W_0 - e^{rt} e^{\left( \frac{\rho-r}{\alpha-1} \right) t} - 1 \left( \frac{p}{b} \right)^{\frac{1}{1-\alpha}} \]

which implies that \( \dot{W}_t \) can be positive or negative, but that \( \frac{dW_t}{dt} > 0 \). To have an interior solution, we need \( \lim_{t \to \infty} W_t > 0 \). This requires \( \frac{\rho - r}{\alpha - 1} < r \), as well as the quite intuitive restriction.
that total initial wealth is bigger than total expected consumption $W_0 > \frac{1}{r - \frac{\beta}{\alpha - 1}} \left( \frac{\rho}{\gamma} \right)^{\frac{1}{1-\alpha}}$. Notice that having an interior solution means that the person does not find it optimal to run down his wealth expressed in period zero equivalents to zero, i.e., $\lim_{t \to \infty} W_t e^{-rt} > 0$. This finding has to be seen in contrast to the finding of the simple life-cycle framework where we have $\lim_{t \to \infty} W_t e^{-rt} = 0$.

Concerning the effect of the bequest parameter $b$ on the consumption level $C_t$, we can easily differentiate expression 2.11 to find that the effect of an increase in the bequest parameter decreases consumption, and hence increases wealth holdings.

$$\frac{dC_{t}^{opt}}{db} = \frac{1}{(\alpha - 1)b} \left( \frac{p}{b} \right)^{\frac{1}{1-\alpha}} e^{\frac{r}{\alpha - 1}t} = \frac{C_t}{(\alpha - 1)b} < 0 \tag{2.12}$$

This finding is not too surprising, as the individual uniformly attaches more value to bequests over the entire lifetime. This increased value decreases the relative value of consumption over the entire lifetime, hence pushing consumption levels down over the entire interval. Notice that the slope of the log consumption profile is unchanged as we vary $b$.

**Budget constraint binding**

The optimal consumption profile can be characterized by the following equation

$$C_t^{\alpha-1} = \left( C_0^{\alpha-1} e^{(\rho+p-r)t} - \frac{b}{p} \left( e^{(\rho+p-r)t} - e^{(\rho-r)t} \right) \right)$$

and the binding budget constraint.

$$\int_0^{\infty} C_t e^{-rt} dt = W_0$$

The wealth profile over time follows trivially from the above consumption process.

The continuous lines in figures 2-2 and 2-3 illustrate consumption and wealth profiles for the case of $g < 0$ and a value $b_1$ of the bequest parameter. The dotted line in the consumption graph represents the interior solution $C^{opt}$ of the previous section for the same value of the bequest parameter.

Notice that $C_0$ is smaller than what it would have been in the unconstrained model. Indeed,
Figure 2-2: $C_t$ Profile, Budget Constraint binding

Figure 2-3: $W_t$ Profile, Budget Constraint binding
\[
\begin{array}{ll}
t = 0 & \frac{r - p - E}{1 - \alpha} + \frac{bc_{1}^{1 - \alpha}}{1 - \alpha} \\
t > 0 & \frac{r - p - E}{1 - \alpha} + \frac{b\epsilon(p-r)C_{1}^{1 - \alpha}}{1 - \alpha} \\
t \to \infty & \frac{r - p - E}{1 - \alpha}
\end{array}
\]

Table 2.2: Time pattern of \( \frac{C_{1}}{C_{t}} \)

\( \frac{C_{0}}{(\frac{p}{b})^{1 - \alpha}} \) because of the lower bound that the bequest motive imposes on the marginal utility of wealth. Further notice the time pattern of \( \frac{C_{1}}{C_{t}} \) as displayed in table 2.2. In the limit as \( C_{0} \to \frac{p}{b} (1 - \alpha) \), \( \frac{C_{1}^{1 - \alpha}}{1 - \alpha} \) which corresponds to the pattern it displays in the fully optimal solution. More generally, both consumption levels and the rate of growth of consumption at any point in time are increasing in \( W_{0} \) as \( \frac{dC_{1}}{dW_{0}} > 0 \) and \( \frac{d\left( \frac{C_{1}}{C_{t}} \right)}{dW_{0}} > 0 \).

Totally differentiating the rate of growth of consumption with respect to the bequest parameter \( b \), we find \( \frac{d\left( \frac{C_{1}}{C_{t}} \right)}{db} > 0 \). This implies that if we are confronted with a case other than the interior solution of the previous section, the bequest motive does not only have an impact on the level of the consumption profile such as illustrated by the decrease in \( C_{0} \), but also on the slope.

**Proposition 1** Suppose the consumption process is determined by the equation

\[
C_{t}^{\alpha - 1} = \left( C_{0}^{\alpha - 1}e^{(\rho + p - r)t} - \frac{b}{p} \left( e^{(\rho + p - r)t} - e^{(p - r)t} \right) \right), \quad \forall t \in [0, \infty[ \quad \text{and by the budget constraint} \\
\int_{0}^{\infty} C_{t}e^{-rt}dt = W_{0} \quad \text{where} \quad W_{0} > 0. \quad \text{Then} \quad \exists \quad t^{*} \in ]0, \infty[ \quad \text{such that} \quad \frac{dC_{1}}{db} = 0, \quad \frac{dC_{1}}{db} < 0 \quad \forall t < t^{*} \quad \text{and} \quad \frac{dC_{1}}{db} > 0 \quad \forall t > t^{*}.
\]

**Proof.**

It is easy to derive \( \frac{dC_{1}}{db} < 0 \) using the two equations that determine the consumption path.

Further, we can use the results from table 2.2 and show \( \frac{d\left( \frac{C_{1}}{C_{t}} \right)}{dt} < 0 \) holds over the entire interval considered. Using these results together with \( \frac{d\left( \frac{C_{1}}{C_{t}} \right)}{db} > 0 \), we are able to establish single crossing of the two consumption profiles in the interval \([0, \infty[\). Denoting the time when these two consumption profiles cross \( t^{*} \), we find the desired result.

**Corollary 2** At any time \( t \in [0, \infty[ \), \( W_{t} \) is strictly increasing in \( b \).
This latter property trivially holds for \( t < t^\star \) as consumption is lower at all times. After \( t^\star \),
given that the present discounted value of consumption from any time \( t > t^\star \) until infinity is
bigger after the increase than before, the wealth level at \( t \) also has to be bigger accordingly.

The result that consumption will actually rise over a positive interval of time may seem
somewhat surprising. But when thinking a little more carefully about the problem, it becomes
less puzzling. Indeed, the increase in the bequest parameter corresponds to a stronger desire
to leave money. This is exactly what is going on here. Corollary 2 shows that, at any point in
time, the wealth is strictly bigger after the increase in \( b \). In expectation, the person thus leaves
a bigger bequest.

Another way of thinking about the result of proposition 1 is consider the constrained opti-
mization problem the individual faces over the time interval \([0, \infty)\). An increase in \( b \) corresponds
to a decrease in the marginal utility of consumption, net of its impact on wealth levels. This is
particularly true in the limit at time 0. Optimal consumption will thus be lower at time 0. Now
thinking at the other extreme of the impact of an increase in \( b \) on the optimal consumption
level at some time close to infinity, the impact is relatively speaking much smaller because the
transversality condition \( \lim_{t \to \infty} W_t e^{-rt} = 0 \) binds at infinity, both before and after the change
in \( b \). Hence the consumption profile is tilted towards later periods.

Notice that these findings nuance Bernheim, Skinner and Weinberg(1997)'s claim that be-
quost motives give rise to systematic correlations between accumulated wealth and the level of
consumption, but not between accumulated wealth and changes in consumption. Within our
setup, the claim is correct for an interior solution such as described in the previous section.
But, as soon as we allow for a binding budget constraint, we find that the rate of change of
consumption is correlated with wealth levels. Consumption levels are, on the other hand, not
monotonic in the strength of the bequest motive.

It may be instructive to look at a graphical representation of the findings. The dashed lines
in figures 2-2 and 2-3 represent the consumption and wealth profiles after an increase in the
bequest parameter from \( b_1 \) to \( b_2 \). These profiles contrast with the initial situation represented
by the continuous lines.
2.3.2 Liquidity constraints

Now we study the more general case of having positive annuity income levels over an interval \([E, \infty]\). Given our definition of the income process in equation 2.7, this assumption can be summarized by \(Y_E > 0\). In the present section, we explicitly allow for liquidity constraints. The reason for doing so is that in many countries, retirees are restricted in their use of old-age income as collateral in credit contracts. Hence borrowing on survival contingent claims on future resources, as represented by the annuity flow \(Y_t\), is limited. In the U.S., for example, using Social Security entitlements as collateral is prohibited. Hausman and Paquette (1987) find empirical evidence that these legal restrictions are actually not fully undone by the markets. These authors find that for involuntary early-retirees, there are consumption jumps at the age of first eligibility for social security benefits. Given that life-cycle savers would prefer to smooth consumption over the life-cycle and hence would prefer to borrow on future social security payments, this finding can be interpreted as evidence for the presence of liquidity constraints.

Summarizing, we can rewrite the optimization problem as

\[
\begin{align*}
\max_{C_t} & \; EU = \frac{1}{\alpha} \int_0^\infty C_t^\alpha e^{-(\delta+r)t} dt \\
& \quad + b \int_0^\infty W_t e^{-(\delta+r)t} dt
\end{align*}
\]

s.t. conditions 2.8 and

\[
W_t \equiv B_t \equiv e^{rt} W_0 - \int_0^t (C_\tau - Y_\tau) e^{r(t-\tau)} d\tau \geq 0 \quad \forall t \in [0, \infty[ \tag{2.14}
\]

Given the form of the income process and the presence of explicit liquidity constraints, there is a multitude of scenarios. We can classify these scenarios into three big groups using the TVC and the binding of the liquidity constraints after \(E\): The first group consists of solutions that display a non-binding budget constraint at the infinite horizon, i.e., \(\lim_{t \to \infty} W_t e^{-rt} > 0\). The second group corresponds to scenarios where wealth is run down to zero in finite time and stays at zero forever after. The third and last group consists of cases that display a binding budget constraint at infinity but where at any given time \(t\) the wealth level is positive, hence implying that the liquidity constraints never bind.
<table>
<thead>
<tr>
<th>Scenario</th>
<th>Group</th>
<th>$\lim_{t \to \infty} W_t e^{-rt}$</th>
<th>$W_E$</th>
<th>$W_t \forall t \in [E, \infty[$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>$&gt; 0$</td>
<td>$&gt; 0$</td>
<td>$&gt; 0, \forall t$</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>$&gt; 0$</td>
<td>$= 0$</td>
<td>$&gt; 0, \forall t$</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>$&gt; 0$</td>
<td>$&gt; 0$</td>
<td>$= 0, \forall t \in [t', t'']$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$&gt; 0, otherwise$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$*$</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>$&gt; 0$</td>
<td>$= 0$</td>
<td>$= 0, \forall t \in [E, t']$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$&gt; 0, \forall t \in [t', \infty[$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$*$</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>$= 0$</td>
<td>$&gt; 0$</td>
<td>$&gt; 0, \forall t \in [E, t'$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$= 0, \forall t \in [t', \infty[$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$*$</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>$= 0$</td>
<td>$= 0$</td>
<td>$= 0, \forall t$</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>$= 0$</td>
<td>$&gt; 0$</td>
<td>$&gt; 0, \forall t$</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>$= 0$</td>
<td>$= 0$</td>
<td>$&gt; 0, \forall t$</td>
</tr>
</tbody>
</table>

Note: $* t', t'' \in [E, \infty[$

Table 2.3: Summary of the different possible scenarios

<table>
<thead>
<tr>
<th>Group 1</th>
<th>$\lim_{t \to \infty} W_t e^{-rt} &gt; 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 2</td>
<td>$\lim_{t \to \infty} W_t e^{-rt} = 0$, $W_t = 0$ for some $t &gt; E$</td>
</tr>
<tr>
<td>Group 3</td>
<td>$\lim_{t \to \infty} W_t e^{-rt} = 0$, $W_t &gt; 0 \forall t &gt; E$</td>
</tr>
</tbody>
</table>

Within every one of these groups, we have different scenarios depending on when precisely liquidity constraints bind in the interval $[0, \infty[$. We consider different parameter values giving us the different scenarios. We summarize the different possible scenarios in table 2.3. In the next four subsections, we present scenarios 1, 2, 5 and 7 successively. Cases 1, 5 and 7 are the simplest possible illustrations of the three big solution groups. Case 2 allows us to illustrate the impact of having liquidity constraints bind over a first retirement period before eligibility for social security benefits. All other scenarios are discussed in appendix A.

Case 1: Interior solution, $\frac{\theta - t}{a-1} < r$, $g < r$, budget constraint not binding

This case is the logical extension of the scenario we discussed in section 2.3.1 to the case of a non-zero retirement income flow. Exactly like in the case of section 2.3.1, the optimal consumption process is

$$C_t^{opt} = \left(\frac{p}{b}\right)^{\frac{1}{\alpha}} e^{\frac{\theta - t}{a-1}}$$

(2.15)

Notice that the comparative statics results with respect to $b$ are also unchanged.
The wealth process, on the other hand, now follows a somewhat different process

\[ W_t + V_t = e^{rt} (W_0 + V_0) - e^{rt \frac{(\frac{r}{\alpha-1} - r)t}{\frac{r}{\alpha-1} - r}} - 1 \left( \frac{p}{b} \right)^{\frac{1}{1-\alpha}} \]

where \( V_t \) represents the annuity wealth computed using simple discounting.

\[ V_t = \begin{cases} \frac{V_0}{r-g} & \text{for } t \geq E \\ V_0 e^{r(t-E)} & \text{for } t < E \end{cases} \]

We know that for this program, the financial wealth level at infinity needs to be positive. A necessary condition for this to be true is that

\[ \lim_{t \to \infty} (W_t + V_t) > 0. \]

Checking this condition, we find that we have to rule out \( \frac{r}{\alpha-1} \geq r \). Furthermore, for \( \frac{r}{\alpha-1} < r \), we have to impose that total initial wealth is bigger than total expected consumption \( W_0 + V_0 > \frac{1}{r - \frac{r}{\alpha-1}} \left( \frac{p}{b} \right)^{\frac{1}{1-\alpha}} \).

As opposed to the no-annuity case of section 2.3.1, we now have to take the rate of growth of the annuity stream into account. To have \( \lim_{t \to \infty} W_t > 0 \), we need \( r > g \). This implies \( \lim_{t \to \infty} W_t e^{-rt} > 0 \), which clearly shows that the present scenario is part of the first group of scenarios.

Figure 2-4 illustrates the consumption pattern and figure 2-5 the wealth pattern for \( \rho > r > g = 0 \).

Case 2: Liquidity constraints binding at time \( E \), \( \frac{r}{\alpha-1} < r \), \( g < r \), budget constraint not binding

A second possible solution is the one where liquidity constraints only bind at time \( E \), but that thereafter the fully optimal consumption profile is possible. This means that for \( t, \tau \in [0, E] \) consumption is determined by

\[ C_t^{\alpha-1} = \left( C_0^{\alpha-1} e^{(\rho+p-r)(t-\tau)} - \frac{b}{p} \left( e^{(\rho+p-r)t} e^{-p\tau} - e^{(\rho-r)t} \right) \right) \quad (2.16) \]
Figure 2-4: $C_t$ Profile, Interior solution

Figure 2-5: $W_t$ Profile, Interior solution
and

\[ \int_0^E C_t e^{-rt} dt = W_0 \]

For \( t \geq E \), the consumption pattern 2.15 is still attainable and optimal. The limiting conditions on total wealth \( W_t + V_t \) are still the same.\(^8\)

To illustrate this scenario, figures 2-6 and 2-7 present the consumption and wealth profiles.

**Figure 2-6: \( C_t \) Profile, Constraint binding at \( E \)**

![Diagram of \( C_t \) Profile]

**Figure 2-7: \( W_t \) Profile, Constraint binding at \( E \)**

![Diagram of \( W_t \) Profile]

Notice that the equation determining the consumption profile over the interval \([0, E]\) is the

---

\(^8\) The conditions are \( r > \frac{\sigma - \gamma}{\sigma - 1} \) and total initial resources bigger than total expected consumption.
same as in section 2.3.1 when we discussed the case of a binding budget constraint in the absence of liquidity constraints. Hence, the time pattern of \( \frac{\dot{C}_t}{C_t} \) is unchanged from table 2.2.

The comparative statics analysis of the impact of an increase in the bequest parameter \( b \) is somewhat more complicated than for the interior solution. Result 2.12 obviously still holds for all \( t \geq E \), but for \( t < E \) the analysis differs and goes into the direction of proposition 1 of section 2.3.1.

**Proposition 3** Suppose we have a bounded interval of time \([t_1, t_2]\) where the consumption process is determined by the equation

\[
C_t^{t_1} = \left(C_{t_1}^{t_2} e^{(\rho + p - r)(t - r)} - \frac{b}{p} \left(e^{(\rho + p - r)t} e^{-pr} - e^{(\rho - r)t}\right)\right), \quad \forall t, r \in [t_1, t_2] \text{ and by the budget constraint } \int_{t_1}^{t_2} C_t e^{-rt} dt = W_0 \text{ where } W_0 > 0.
\]

Then \( \exists t^* \in [t_1, t_2] \) such that \( \frac{dC_t}{db} = 0 \), \( \frac{dC_t}{db} < 0 \) \( \forall t < t^* \) and \( \frac{dC_t}{db} > 0 \) \( \forall t > t^* \).

**Proof.**

It is easy to derive \( \frac{dC_t}{db} < 0 \) and \( \frac{dC_t}{db} > 0 \) using the two equations that determine the consumption path. Further, using the expression for \( \frac{\dot{C}_t}{C_t} \) from table 2.2 we know \( \frac{d\left(\frac{\dot{C}_t}{C_t}\right)}{dt} < 0 \) and \( \frac{d\left(\frac{\dot{C}_t}{C_t}\right)}{db} > 0 \). Hence, we are able to establish single crossing of the two consumption profiles in the interval \([t_1, t_2]\). Denote \( t^* \) the time when these two consumption profiles cross. □

**Corollary 4** At any time \( t \in [t_1, t_2] \), \( W_t \) is strictly increasing in \( b \).

The intuition for the tilt in the consumption profile is easy to understand. As \( b \) increases, the value the individual attaches to bequests increases much more at time \( t_1 \) rather than at some time \( t \) close to \( t_2 \). The reason is that the individual knows, that he runs down wealth to zero until time \( t_2 \). Hence, to make the budget constraint hold, a decrease in consumption early in retirement has to be compensated by an increase in consumption later in life.

Using proposition 3 for the special case where \( t_1 = 0 \) and where \( t_2 = E \), we can summarize our findings by

\[
\begin{align*}
\frac{dC_t}{db} &< 0 \quad \text{for } t < t^* \\
\frac{dC_t}{db} &> 0 \quad \text{for } E > t > t^* \\
\frac{dC_t}{db} &< 0 \quad \text{for } t \geq E
\end{align*}
\]
or maybe even more instructively in figure 2-8, where the dashed line describes the consumption profile under the initial value of the bequest parameter $b_1$ and the continuous line describes the consumption profile under the new increased value of the parameter $b_2$.

Assuming identical initial wealth, our model predicts that stronger bequest motives should imply higher wealth levels. This reinforces the findings of section 2.3.1 that Bernheim, Skinner and Weinberg’s claim only holds as long as both liquidity constraints and the budget constraint are slack over the entire life-span we consider.

![Figure 2-8: Effect of $b$ on consumption profile](image)

The present scenario is an interesting illustration of the impact of liquidity constraints on the consumption and wealth accumulation behavior of early retirees. Recalling our interpretation of the time $E$ as the time the (early) retiree becomes eligible for social security benefits, we find that binding credit constraints, which are especially stringent for involuntary early retirees, decrease both the level and the rate of change of consumption during periods of early retirement. Further, we find that the jump in consumption at the age of eligibility should be stronger for people with smaller bequest motives.

**Case 5: Budget constraint binding at infinity, \( \frac{\rho - r}{\alpha - 1} > g > \frac{\rho + \rho - r}{\alpha - 1} \)**

After these two cases where \( \lim_{t \to \infty} W_t e^{-rt} > 0 \), the next two sections contain scenarios where the latter expression is 0. It is trivial to show that for this to be true, we need \( \frac{\rho - r}{\alpha - 1} > g \).

---

*Suppose the latter relation does not hold. Then there exists a time \( t \) such that the fully optimal consumption profile and the annuity income profile will cross, hence implying for \( t > \tilde{t} \) a strictly positive level for the above*
In the present section, we consider a growth rate of the annuity stream that is bigger than the limiting growth rate of the constrained consumption path as \( t \to \infty \), but smaller than the growth rate of the fully optimal consumption profile. Denoting \( T \) the time when the liquidity constraints start to bind, the optimal consumption path is determined by

\[
C_t = \begin{cases} 
C_0^{\alpha-1}e^{(\rho+p-p-r)t} - \frac{b}{p} \left(e^{(\rho+p-r)t} - e^{(\rho-r)t}\right)^{\frac{1}{\alpha-1}} & \text{for } t < T \\
Y_t & \text{for } t \geq T
\end{cases}
\]

and the resource constraint

\[ W_0 = \int_0^T (C_t - Y_t)e^{-rt}dt \]

We assume that the liquidity constraints do not bind on the subinterval \([0, E]\). For a discussion of the case of having liquidity constraints bind on this subinterval, we refer the reader to appendix A.3.

It is easy to see that consumption will be continuous in time, so we have

\[
\left(C_0^{\alpha-1}e^{(\rho+p-r)T} - \frac{b}{p} \left(e^{(\rho+p-r)T} - e^{(\rho-r)T}\right)^{\frac{1}{\alpha-1}} \right) = e^{\theta(T-E)}Y_E
\]

The consumption and wealth pattern are illustrated in figures 2-9 and 2-10. They are characterized by a period, where the first-order conditions determine the consumption profile, and then afterwards by a period, when the consumption pattern is determined by the binding liquidity constraints.

Using a simple extension of proposition 3 it is easy to derive that the comparative statics of the consumption path as we vary \( b \): There is a time \( t^{**} \in [0, E] \) and a point in time \( M' > M \) such that

\[
\begin{cases} 
\frac{dC_t}{db} < 0 & \text{for } [0, t^{**}] \\
\frac{dC_t}{db} > 0 & \text{for } [t^{**}, M'] \\
\frac{dC_t}{db} = 0 & \text{for } [M', \infty[\
\end{cases}
\]

limit.
Figure 2-9: $C_t$ Profile, Constraint binding at $t > E$

Figure 2-10: $W_t$ Profile, Constraint binding at $t > E$
Case 7: Budget constraint binding at infinity, $\frac{\rho + p - r}{a - 1} > g$

Considering the case of $\frac{\rho + p - r}{a - 1} > g$ with no binding liquidity constraint at $E$, we find that the optimal consumption profile is determined by the optimality condition equation 2.16 as well as the binding budget constraint

$$\int_0^\infty (C_t - Y_t)e^{-rt}dt = W_0$$

Notice that the present scenario is most plausible for a constant nominal annuity stream ($g < 0$). Indeed, for a constant real annuity stream ($g = 0$) we would have to impose the rather implausible requirement $r > \rho + p$ to be in the present scenario. For $g < 0$, on the other hand, the condition $\frac{\rho + p - r}{a - 1} > g$ becomes much more plausible, especially for risk averse individuals.

Figures 2-11 and 2-12 illustrate an example of the consumption and wealth profiles for $g < 0$. All of the results we derived in the absence of a retirement income system go through to the case of positive income given that the present scenario satisfies all conditions of scenario 2.3.1.

Figure 2-11: $C_t$ Profile, $\frac{\rho + p - r}{a - 1} > g$

2.3.3 Comments on the uncertainty case

We conclude this section by summarizing our results. First, the consumption level at any time $t$ can determined by three regimes: The fully optimal consumption level, a first-order condition combined with a binding budget constraint and the exogenous retirement income level.
Second, a finding that the reader may not have suspected at the outset is that consumption levels actually increase with the strength of the bequest motive over some subintervals in all cases but the one when the fully optimal consumption path can be attained. The idea behind this finding is that the consumption profile changes slopes on constrained intervals, and therefore forces a reallocation of resources from earlier periods to later periods, as consumption in earlier periods becomes relatively speaking less interesting due to the increased opportunity cost in terms of foregone bequests. By doing so, the slope of the consumption profile approaches the slope of the fully optimal consumption profile \( \frac{r-a}{1-\alpha} \), a fact that is captured by the expression \( \frac{d}{db} \left( \frac{c_t}{c_{t+1}} \right) > 0 \).

Combining this finding of a reallocation of resources due to binding liquidity constraints with the strictly decreasing effect of the bequest parameter on consumption levels in periods when the consumption process is determined by the fully optimal path, it is easy to see that increases in the bequest parameter \( b \) render liquidity constraints less binding. An intuitive way to see this general result is to look at the special case of section 2.3.2. There was a jump in consumption at the age of first eligibility for retirement benefits because of binding borrowing constraints over the early retirement period. As we increase the value of the bequest parameter \( b \), the discontinuity in consumption shrinks, as can easily be seen in figure 2-8. The mechanism behind this reduction is that the person attaches a smaller value to consumption early in life, hence reducing the implicit cost of not being able to borrow on the future. Therefore, for a
sufficiently high \( b \), we will end up in the case of the fully optimal consumption profile determining consumption over the entire life-span.

Notice that the effect on the wealth profile as we increase the bequest parameter is exactly the expected one, as the wealth and thus bequests grow in expectations and even more strongly it is non-decreasing at any given point in time.

### 2.4 Annuity valuation

Using the setup of section 2.3, we analyze the question of annuity valuation. The previous literature has largely ignored the impact of bequest motives on the value of annuities. There have been two common approaches to computing the value of an annuity payout stream. A first strand of the literature relies on simple actuarial calculations. It is easy to show that this method is correct in the case of either a risk averse consumer or a competitive annuity markets. The second approach is due to Bernheim(1987). Using a standard life-cycle model with a risk averse individual and a marginal unavailability of annuities, Bernheim showed that the value of an annuity is close to the simple financial value of the annuity payout stream. The reason for this finding is easiest to understand when comparing a survival contingent annuity contract to a simple financial asset. Without binding liquidity constraints, an individual attaches equal value to an annuity contract paying out \$1\ every period conditional on survival and a financial asset paying out \$1\ per period independently of survival. Hence, annuities should be valued exactly in the same way as a financial asset.

Following the typology of table 2.3 of section 2.3.2, we can once again discuss annuity valuation within every one of the eight possible scenarios. In the present section we only analyze the same four representative scenarios that we presented in section 2.3.2.\(^{10}\)

\(^{10}\)The results for the remaining scenarios are straightforward extensions of the four scenarios we present.
2.4.1 Case 1: Interior solution, \( \frac{r-\delta}{\alpha-1} < r, g < r \), budget constraint not binding

In this particular setup, it is easy to determine the value of the marginal dollar of annuity payments \( Y_E \) in terms of financial wealth at time 0 \( (W_0) \).

\[
\left. \frac{dW_0}{dY_E} \right|_{U^*} = \frac{e^{-(r+p)E}}{r + p - g}
\]

Recalling the definition of the simple discounted value of annuity wealth \( V_t \) from section 2.3.2, we can also rewrite the present finding as

\[
\left. \frac{dW_0}{dV_0} \right|_{U^*} = \frac{r - g}{r + p - g} e^{-pE}
\]

Similarly, defining the actuarially correctly discounted value of annuity wealth \( V_t^{act} \) as the present discounted value, taking both survival probabilities and interest rates into account

\[
V_t^{act} = \begin{cases} 
\frac{V_t}{r+p-g} & \text{for } t \geq E \\
V_E^{act} e^{(r+p)(t-E)} & \text{for } t < E
\end{cases}
\]

and noticing that \( V_0^{act} = V_0 \frac{r-g}{r+p-g} e^{-pE} \), we can rewrite the expression for the marginal value of annuity payouts as

\[
\left. \frac{dW_0}{dV_0^{act}} \right|_{U^*} = -1
\]

This result stands in stark contrast with Bernheim's finding that the correct way to discount future annuity payouts is to use simple discounting and not the actuarially correct counterpart for the case of no binding liquidity constraints. The reason for this quite dramatic change is that the individual no longer only attributes a certain utility value to the periods when alive, but also to the periods when dead. This implies that annuities and financial assets are no longer perfect substitutes and he values a financial asset more since it pays out a return in every future state of the world, whereas the annuity only pays out a return in the states when the individual is alive.

The result also means that we do not need to have perfect annuity markets or a coefficient of relative risk aversion equal to zero to validate actuarially fair valuation. It suffices to have an interior solution for a risk averse consumer with a bequest motive, and this independently
of the existence of a market in marginal annuity claims. The reason for this finding is that the
equalization of marginal utilities of consumption and bequests in all possible states of the world
makes the individual de facto risk neutral.

Notice that our finding of actuarial valuation is independent of time. For example, another
natural way to look at the same problem of marginal annuity valuation is to look for the change
in period E financial wealth required to keep the person indifferent after a marginal change of
annuity holdings. Notice that to make this comparison meaningful, we need to assume that the
individual can’t borrow against period E wealth. More formally we are interested in the value
of \( \frac{dW_E}{dV_0^{act}} \bigg|_{U^*} \).

For the case of our interior solution, it is trivial to show that \( \frac{dW_E}{dV_0^{act}} \bigg|_{U^*} = \frac{dW_0}{dV_0^{act}} \bigg|_{U^*} = -1 \),
and hence that the two approaches prove to be totally identical. Further, the restriction of
considering wealth at period E that the person cannot borrow against is not a real restriction
for the case an interior solution since wealth levels are anyway positive at all times.

2.4.2 Case 2: Liquidity constraints binding at time E, \( \frac{p-r}{s-1} < r, g < r \), budget
constraint slack

As noted in section 2.3.2, the present case best describes the situation of an early retiree whose
consumption is constrained before the time of first entitlement (E) to retirement benefits, but
who attains the fully optimal consumption levels after E. The value of an increase of a marginal
dollar of annuity payout \( Y_E \) in terms of the period 0 concepts is

\[
\frac{dW_0}{dV_0^{act}} \bigg|_{U^*} = -\frac{1}{(1-e^{-pE}) + pX_1} > -1
\]

where \( X_1 = \frac{1}{b} \left( \int_0^E e^{-(p+r)t} C_t^{\alpha-1} C'_t dt \right) - \left( \int_0^E e^{-(r+p)+t} \int_0^t e^{r(t-s)} C'_s ds dt \right) \), which implies \( C'_t \equiv C'_t(W_0) \). But \( X_1 > \frac{1}{p} > \frac{e^{-pE}}{p} \), which implies that the value that the individual attributes
to the annuity payout stream \( Y_t \) is inferior to the actuarial value.

This latter result is rather interesting as it shows that for people facing binding liquidity
constraints in their early periods of life, even an actuarially fair insurance system would not
be sufficient to have them participate in the annuity market. But the finding is not surprising:
First the payout stream does not have an annuity value, i.e., there is no bigger utility attached
to receiving income when alive rather than having heirs receive the income when dead because it is an interior solution. Second, given that he is liquidity constrained over the first interval from early retirement until retirement, he prefers to have a higher $W_0$ rather than have additional resources in the unconstrained interval after initial entitlement $E$. Similarly, we can say that any deterministic future payment displays the same property. Indeed, comparing a $1$ state-of-the-nature independent payout in period $E$ to its compensating variation in terms of $W_0$, we also find that it is smaller than $e^{-rE}$ because of the binding liquidity constraint for the time period $[0, E]$. To separate out the part of the effect that is due to the inherent survival-contingent characteristics of annuities, we use our second measure of annuity valuation, which is based on the wealth concepts as of age $E$, supposing it is impossible to borrow on $W_E$. Computing the latter, we find

$$\frac{dW_E}{dV^\text{act}} \bigg|_{U^*} = -1$$

which means that the net effect due to the annuity's inherent characteristics are not changed with respect to the previous scenario of a fully interior solution.

2.4.3 Case 5: Budget constraint binding at infinity, $\frac{\rho - r}{a - 1} > g > \frac{\rho + p - r}{a - 1}$

Recalling our findings from section 2.3.2 on the consumption and wealth profiles, we know that liquidity constraints bind from some finite time $T$ onwards. Hence, after time $T$, the value of the annuity stream will be smaller than the actuarially fair value.

It is difficult to derive the value of a marginal annuity stream explicitly. Nonetheless, we can derive bounds on its value

$$0 > \frac{dW_0}{dV_0} \bigg|_{U^*} > -1$$

Without any additional information on the precise parameter values of the problem, it is difficult to determine whether the value of the marginal annuity is bigger or smaller than the actuarial value, as there are two effects playing in opposite directions: Prior to time $T$, there is a clear annuity value to having this payout stream rather than a financial asset of the same actuarial value. After $T$, the annuity payout stream generates a utility which is lower than the one generated by a financial asset, as the individual would prefer to reallocate annuity income which is possible in the case of a financial wealth.
2.4.4 Case 7: Budget constraint binding at infinity, \( \frac{\theta + p - r}{a - 1} > g \)

Lastly, for the case of no binding liquidity constraints but a binding budget constraint at infinity, we find

\[
0 > \left. \frac{dW_0}{dV_0} \right|_{U^*} > -1 \\
-1 > \left. \frac{dW_0}{dV_0^{act}} \right|_{U^*}
\]

or combining these two inequalities

\[
-1 < \left. \frac{dW_0}{dV_0} \right|_{U^*} < -\frac{r - g}{p + r - g} e^{-pE}
\]

The value of an annuity is hence bigger than its actuarial value, but still smaller than the purely financial value. This result is quite intuitive: Because of the binding budget constraint, there is a clear annuity value attached to the retirement income stream. The above finding is reconfirmed when we look at the bounds that we can impose on the expression \( \left. \frac{dW_E}{dV_E} \right|_{U^*} \). We can easily determine that

\[
-1 < \left. \frac{dW_E}{dV_E} \right|_{U^*} < -\frac{r - g}{p + r - g}
\]

2.4.5 Summary of the annuity results

Our results show that Bernheim(1987)'s claim that the value of an annuity is equal to the simple discounted value of future payouts has to be nuanced. Indeed, for a sufficiently strong bequest motive, even in the presence of a marginal unavailability of annuity contracts and without perfect competition, actuarial valuation is the correct method for valuing future annuity claims. Our analysis shows that using actuarially correct valuation implicitly assumes an active bequest motive if we are confronted with imperfect annuity markets. Simple financial valuation on the other hand implicitly assumes no bequest motive.

When allowing for an early retirement phase from time 0 to \( E \), we show that to separate the value of the annuity stream that is due to the survival-contingency from the value that is due to the binding of the liquidity constraints during the early retirement period, we should use a second indicator. We use wealth at the time of first entitlement \( E \) that we cannot borrow against
in addition to the standard period 0 wealth equivalence measure such as used by Bernheim.

2.5 Conclusion

In the present paper we set out to answer two big questions that arise when we introduce a bequest motive into the standard life-cycle model. These two questions were: First, a characterization of the consumption and wealth profiles both under perfect and imperfect capital and annuity markets. Second, the issue of the correct valuation of marginal annuity claims. Using a quasi-linear approach to analyze these questions we successively studied the simplest two period certainty case as well as the infinite horizon life-span uncertainty model.

Within the framework of the first model, we have a clear result that consumption is monotone non-increasing in the linear bequest parameter, and that wealth is hence monotone non-decreasing. Switching over to the life-span uncertainty case, we showed that the above result no longer holds, except in the case of a strictly interior solution. Indeed, as soon as liquidity constraints or the budget constraint bind, there is a strictly positive interval of time during which consumption actually rises! As surprising as this result may seem, we show that it is just the natural complement of the finding that an increased bequest motive generates higher wealth levels at any point in time.

We further show that as long as liquidity constraints or the budget constraint are binding at some point in time, a bequest motive does not only affect consumption levels, but also the slope of the consumption profile. The latter effect of a bequest motive on the slope has generally been ignored.

As for annuity valuation, our analysis shows that there is no single criterion that completely summarizes the value of a marginal annuity payout stream to the individual. Using two different, but closely related concepts of valuation, we are able to separate out the pure annuity effect from the additional effect due to the timing of the onset of the annuity payout stream. We show that as soon as we deviate from the simplest life-cycle framework with uncertainty such as the one used by Bernheim(1987), the value of a marginal increase payouts is no longer equal or very close to the present discounted value computed using simple discounting, but tends towards the direction of its actuarially correct equivalent. We find that for the case of a sufficiently
strong bequest motive, actuarial valuation is closest to the true economic value of a marginal annuity stream. For a weak or non-existent bequest motive, simple financial discounting best approximates the true value. These findings are important for some of today's most acute policy questions, particularly the evaluation of reforms to the present-day old-age income systems. To evaluate and understand the changes to the system, there is a need to correctly measure the value of future annuity payouts. Our findings indicate that depending on the scenario we are in, using simple financial discounting for valuing future social security and pension may introduce a major measurement error into the analysis.
Chapter 3

Timing of Gifts and Bequests

3.1 Introduction

We extend the standard life-cycle model of consumption to include a joy-of-giving motive. In contrast to the previous literature on joy-of-giving bequest motives, we explicitly allow for the second form of giving that is inter-vivos gifts. The complete absence of inter-vivos gifts from the previous literature on joy-of-giving motives is quite surprising, as it is widely recognized that inter-vivos gifts are by no means a negligible phenomenon. For example, Cox (1987) computes a ratio of gifts to bequests of 1.5 using estimates from Kotlikoff and Summers (1981) and Kurz (1984). The only area in the economic literature where gifts and bequests have been studied in an integrated framework is the estate taxation literature: simple present discounted values (PDV) are computed to derive which time pattern for the wealth transfer from one generation to the other has the highest payoff. In the present paper we use a more general approach as we consider the life-cycle pattern of gifts and bequests that a liquidity constrained utility-maximizer would choose. We discuss different setups for the utility of gifts and bequests function.

For the special case of a linear utility of giving functions, we analyze the impact of choosing

---

1See, for example, Fischer (1973), Friedman-Warshawsky (1988) and Hurd (1989) for earlier formulations of joy-of-giving motives.

2Rivers and Crumbley (1979) analyze the optimal timing of gifts and bequests after the 1976 tax reform using a simple money's worth calculation. Similarly, Kuehlwein (1994) uses simple PDV calculations to show that the true gift and bequest tax were not equalized by the same reform.
different discount rates and characterize the pattern of gifts and bequests over the life-cycle. The question of which utility of giving discount rate to choose has widely been ignored even though it has strong implications for the level and the timing of gifts and bequests. The choice of the discount rate is by no means neutral with respect to either the timing and the level of gifts and bequests, or the interpretation of the linear utility parameters. Our analysis shows that even within the class of linear utility of gifts and bequests functions, the timing of the wealth transfer from one generation to the next one does not have to be characterized by a trivial corner solution, mostly because of the interaction of the motive of giving with the liquidity constraints.

The structure of the paper is as follows. In section 3.2, we discuss different ways of setting up the utility of gifts and bequests function. We then proceed to characterize the optimum for a baseline case with a linear specification for the utility of gifts and bequests function in section 3.3. We find that using the interest rate as the discount rate for gifts and bequests has two advantages: first, it simplifies the problem substantially as the timing of gifts and bequests becomes irrelevant; second, it allows the interpretation of the linear utility of gifts and bequests parameter as a summary measure of the strength of the motive of giving. Section 3.4 contains some concluding comments.

### 3.2 The framework

We assume that the individual lives a two-period life \( t = \{0, 1\} \) and faces no uncertainty about his life-span. Consumption, gifts and bequests in period \( i \) are demoted \( C_i, G_i \) and \( B_i \) respectively. We suppose that the utility of consumption is additively separable over time and satisfies strict concavity assumptions, as well as \( u'(0) \to \infty \). We define \( \rho \) to be the time preference rate for the discounting of utility derived out of consumption. Further, we define the real interest rate the individual faces on the capital markets as \( r \). Writing the utility of gifts and bequests function in a general form \( f = f(G_0, G_1, B_2) \), where \( f \) is increasing and concave in the three arguments, we can state the individual's utility function as

\[
U(C_0, G_0, C_1, G_1, B_2) = u(C_0) + \frac{1}{1 + \rho} u(C_1) + f(G_0, G_1, B_2) \tag{3.1}
\]
The individual maximize this objective function 3.1 subject to a budget constraint, to liquidity constraints as well as to gift and bequest non-negativity constraints. Denoting initial wealth $W_0$, period 1 income $Y_1$ and savings at time zero $S$, we can write this set of constraints as

$$ (W_0 - C_0 - G_0)(1 + r)^2 + (Y_1 - C_1 - G_1)(1 + r) - B_2 = 0 $$  \hspace{1cm} (3.2)

$$ S = W_0 - C_0 - G_0 \geq 0 $$  \hspace{1cm} (3.3)

$$ G_0 \geq 0, G_1 \geq 0, B_2 \geq 0 $$  \hspace{1cm} (3.4)

It is difficult to say anything about the optimum using such a general specification of $f(G_0, G_1, B_2)$. Therefore, we now discuss different plausible formulations for $f$. Two major elements play an important role in determining what is a reasonable specification of the function $f$. First, there is the question of the time discounting of gifts and bequests in different periods of life. Second the issue of whether there is any relative preference for either $G_0, G_1$ or $B_2$ beyond the preceding sheer time-discounting element.

We start by considering the case of an individual who only values the total amount of money he gives away as either gifts or bequests, regardless of timing. Under this mental setup, the utility function $f$ can be written as a simple function $f^1$ of the PDV of gifts and bequests

$$ f(G_0, G_1, B_2) = f^1 \left( G_0 + \frac{G_1}{1 + r} + \frac{B_2}{(1 + r)^2} \right) $$  \hspace{1cm} (3.5)

The use of the interest rate for the discounting of gifts and bequests does not look unreasonable, since that both gifts and bequests are pure wealth transfers.

The assumption of perfect substitutability of equal monetary values is probably somewhat extreme. It is hence useful to generalize expression 3.5 to allow for less than perfect substitutability between equal financial values of the different forms of giving. We do so in expression 3.6 where the utility function $f$ remains a function of some linear combination of $G_0, G_1$ and $B_2$, with linear parameters $\beta_1$ and $\beta_2$.

$$ f(G_0, G_1, B_2) = f^2 (G_0 + \beta_1 G_1 + \beta_2 B_2) $$  \hspace{1cm} (3.6)
We can think of parameter interpretations of the linear parameters $\beta_i$ depending on the mental setup that underlies the model. On the one hand, we can think of a simple generalization of the setup described in expression 3.5: the individual uses a separate discount factor $\xi$, instead of the interest rate $r$, for the utility discounting of gifts and bequests from different periods within the function $f$. This translates into $\beta_i = \frac{1}{1+\xi}$ for $i = \{1,2\}$ in expression 3.6. Indeed, there is no reason that excludes any given discount rate from being applied to the utility discounting in $f$ instead of the interest rate $r$.

On the other hand, differences in individual preferences may take various forms beyond the previously discussed simple time discounting effects. To discuss this idea, we define linear utility parameters $\beta_i^*$ that are net of the time discounting effect. We can hence write the utility function $f$ as

$$f(G_0, G_1, B_2) = f^3 \left( G_0 + \frac{\beta_1^*}{1+\xi}G_1 + \frac{\beta_2^*}{(1+\xi)^2}B_2 \right)$$

(3.7)

The setup allows us to study the implications of a wide variety of issues: first, the preference for giving under the form of either inter-vivos gifts or bequests at death; second, the question of changing preferences for gifts and bequests over the different stages in the life-cycle; third, the implications of estate taxation rules for gifts and bequests at different times over the life-cycle.

The question of a preference for either inter-vivos gifts or bequests, net of the time discounting effect, is best thought of in the framework of the period 1 decision problem while conditioning on some given period 0 outcome. The individual has to allocate a given amount of resources to either consumption $C_1$, gifts $G_1$ or bequests $B_2$. One might argue that the utility a person derives from the fact of giving depends on whether he gives while alive, or he leaves some bequest at his death. But in our opinion, there is no clear indication as to which way of giving generates the biggest utility. Some people may prefer to give while living, so as to see the heirs take pleasure in receiving the gift. Others may consider giving at their death the better solution so as not to be confronted with disputes between heirs for the different parts of the estate. Hence, a natural candidate for the gift and bequest parameters in period 1 is $\beta_1^* = \beta_2^*$. It is important to notice that $\beta_1^* = \beta_2^*$ only means that net of time discounting, the person is indifferent between the two forms of giving. Only for the case of $\xi = r$, we have a strict indifference between the two forms of giving. The reason for this finding is that wealth accumulates at rate $r$ (see the budget constraint 3.2) but that gifts and bequests are discounted

49
at rate $\xi$. Hence, for $\xi > r$ (resp. $\xi < r$), setting $\beta_1^* = \beta_2^*$ implies a strict preference for gifts (resp. bequests).

It is easy to think of reasons why people may have preference parameters $\beta_1^* \neq 1$ for giving wealth away in different periods of their life. One stylized fact is that people prefer to make descending intergenerational transfers rather than ascending ones. In a two-period overlapping generations world, a young person's gift can only be ascending. Hence, it would be of smaller utility value than an old person's gift or bequest as the latter transfers are directed towards the next generations, which implies $\beta_1^* > 1, \beta_2^* > 1$. At the opposite extreme, it is also possible to think of the case of an individual who derives less utility from gifts or bequests later in life rather than earlier in life.

Within this same framework described by the function $f^3$, we can address another question, namely the impact of estate taxation rules on the timing of gifts and bequests. The standard approach of analyzing the impact of the estate tax is to use the gift and bequest variables net of taxes in the utility function, while taking taxes into account in the individual's budget constraint. But equivalently, we can study the same question by expressing all gift and bequest variables as pre-tax variables and by allowing for different linear parameters $\beta_i^*$ in our utility function of gifts and bequests.

### 3.3 Characterization of a baseline model

Our focus in the present section is on the impact of the discount rate $\xi$ on the timing of the act of giving, as well as its interaction with the strength of the motive of giving and the liquidity constraints. We present a characterization of the gifts and bequests process over the life-cycle for a baseline model. We analyze the gift and bequest pattern for the case of a linear specification of the utility function of the PDV of gifts and bequests, with linear parameter $b$. We allow the discount rate $\xi$ to be different from either $r$ or $\rho$. We suppose that the utility derived out of the act of giving is realized at the moment when the gift or bequest leaves the estate of the giver. Hence, we can rewrite the utility function $f^2$ as:

$$f(G_0, G_1, B_2) = b \left( G_0 + \frac{G_1}{1 + \xi} + \frac{B_2}{(1 + \xi)^2} \right)$$

(3.8)
The linear structure of the utility of gifts and bequest needs some explanation. Fischer (1973) for example uses a CRRA specification for both utility out of consumption and out of bequests, while imposing both the same risk aversion parameter and the same discount rate. Hurd (1989) on the other hand uses the linear specification for the utility out of bequests, while still keeping the same discount rate for utility out of consumption and bequests.

Quasi-linearity allows for increased tractability while preserving the features of more complicated utility functions.\(^3\) In real life, people tend to be less risk averse with respect to gifts and bequests than with respect to their own personal consumption. This implies some restrictions on the concavity of the function \(f\). Quasi-linear preferences fully capture this idea, as they imply a risk aversion parameter of 0 with respect to gifts and bequests, whereas in all generic cases the risk aversion parameter is positive for consumption. But even though we restrict our attention to a quasi-linear setup, it is conceptually easy to generalize to other functional forms, as long as the utility of gifts and bequests is less concave than the utility of consumption.

A further important remark relates to the assumption that the utility of giving was generated at the time the gifts or bequests leave the estate of the giver. For simplicity reasons, we carry this assumption throughout our model. However, allowing for a non-zero interval between the time the money leaves the estate of the giver and the utility is realized is not as unreasonable as it may sound. For example, it is quite common that people set up trust funds that allow them to control their wealth beyond the sheer limits of their own life. Abstracting away from pure tax reasons, trust funds are mainly used to delay the wealth transfer to the heirs.\(^4\) Allowing for for the existence of trust funds in this framework, the question of when the utility out of the act of giving is realized becomes non-trivial: is it at the time the trust fund gets set up, or rather when the wealth effectively starts arriving in the estate of the heirs?\(^5\)

Suppose that the individual has a possibility of introducing a uniform \(n\)-period delay to the utility realization independent on whether he gives as \(G_0, G_1\) or \(B_2\). Further assume that while the money is in the trust fund, wealth still accumulates at the same rate \(r\) and utility is still

---

\(^3\)Our findings easily extend to more general setups allowing for continuous time and life-span uncertainty. Further, we can generalize to utility functions other than quasi-linear as long as the utility of gifts and bequests is less concave than the utility of consumption.

\(^4\)Trust funds may be subject to legal restrictions such as, for example, the U.S. laws against perpetuities.

\(^5\)Nothing conceptually excludes the possibility of having a person derive utility at a time which is before the actual act of giving.
discounted from the future using the same discount rate \( \xi \). It is easy to generalize the previous no-delay setup to incorporate the possibility for the individual to introduce an \( n \)-period delay and to optimize over \( n \). We can hence write the linear utility of gifts and bequests function as

\[
 f(G_0, G_1, B_2) = b\left(\frac{(1 + r)^nG_0}{(1 + \xi)^n} + \frac{(1 + r)^nG_1}{(1 + \xi)^{n+1}} + \frac{(1 + r)^nB_2}{(1 + \xi)^{n+2}}\right) = b^{\text{trust}} \left(\frac{G_0}{1 + \xi} + \frac{G_1}{1 + \xi} + \frac{B_2}{(1 + \xi)^2}\right)
\]

where \( b^{\text{trust}} = b\left(\frac{1 + \xi}{1 + \xi}\right)^n \). Therefore, allowing for the possibility of an \( n \)-period delay only has the effect of rescaling the bequest parameter from \( b \) to \( b^{\text{trust}} \).

Getting back to the original quasi-linear model without the possibility of setting up a trust fund, we solve the individual’s constrained optimization problem using the Kuhn-Tucker method. We find the following set of optimality conditions in terms of consumption, gifts, and bequests:

\[
\begin{align*}
 u'(C_0) &= \lambda_1 + \frac{1 + \xi}{1 + \rho} u'(C_1) \\
 b &= \lambda_1 + \frac{1 + \xi}{1 + \rho} u'(C_1) - \mu_0 \\
 \frac{b}{1 + \xi} &= \frac{1 + \rho}{1 + \rho} u'(C_1) - \mu_1 \\
 \frac{b}{(1 + \xi)^2} &= \frac{1}{(1 + \rho)(1 + \rho)} u'(C_1) - \mu_2 \\
 W_0 - C_0 - G_0 &\geq 0, \quad \lambda_1 \geq 0, \quad \lambda_1(W_0 - C_0 - G_0) = 0 \\
 W_0 + \frac{Y_1}{1 + r} &= C_0 + G_0 + \frac{C_1 + G_1}{1 + r} + \frac{B_2}{(1 + r)^2} \\
 G_0 &\geq 0, \quad \mu_0 \geq 0, \quad \mu_0 G_0 = 0 \\
 G_1 &\geq 0, \quad \mu_1 \geq 0, \quad \mu_1 G_1 = 0 \\
 B_2 &\geq 0, \quad \mu_2 \geq 0, \quad \mu_2 B_2 = 0
\end{align*}
\]

where \( \lambda_1 \) is the Kuhn-Tucker multiplier associated with the liquidity constraint 3.3. The multipliers \( \mu_0, \mu_1 \) and \( \mu_2 \) are associated with the gift and bequest non-negativity constraints 3.4. To interpret these optimality conditions, we distinguish two cases depending on whether \( \xi > r \) or \( \xi \leq r \).

---

\( ^6 \)We can extend the present analysis to the case of different durations of trust funds for the different forms of giving. Instead of a simple rescaling of the parameter \( b \) to \( b^{\text{trust}} \), we would have a rescaling to three separate parameters \( b_0^{\text{trust}}, b_1^{\text{trust}} \) and \( b_2^{\text{trust}} \) for \( G_0, G_1 \) and \( B_2 \) respectively.
3.3.1 The case of $\xi > r$

The individual discounts utility derived out of gifts and bequests at rate $\xi$ which is bigger than the rate of return $r$ he earns on the financial markets by delaying gifts and bequests. Hence, in a world without liquidity constraints, there would be a clear corner solution where the individual makes all wealth transfers early in his life under the form of gifts $G_0$. But in the presence of potentially binding liquidity constraints, the ability to give in period 0 is limited to period 0 resources on hand. However, even in the presence of such constraints, we have $B_2 = 0$ at any optimum. The reason for this finding is that the individual can always make any planned wealth transfers under the form of period 1 gifts rather than as bequests without violating the liquidity constraints.

To characterize the optimum, we derive equations for the expressions that represent the borders between the different scenarios with respect to the binding of the liquidity and gift non-negativity constraints. We express the equations in terms of the variables $b$ and $Y_1$, keeping total income $R \equiv W_0 + \frac{Y_1}{1+r}$ constant and represent them graphically in $(b,Y_1)$ space.

The expression separating the binding and slackness of the credit constraint, can be written as

$$\begin{cases} Y_1 = Y_1^* & \text{for } b < b^* \\ b = \frac{1+r}{1+\rho} u'(Y_1) & \text{for } b \geq b^* \end{cases}$$

(3.9)

where $Y_1^*$ is the solution to $u'(R - \frac{Y_1}{1+r}) = \frac{1+r}{1+\rho} u'(Y_1)$ and $b^*$ is the solution to $b = \frac{1+r}{1+\rho} u'(Y_1^*)$. Notice that $b$ is non-increasing in $Y_1$ because of the sign of $u''(.)$ and that the concavity of the expression depends on the sign of the third derivative of the utility function. For example, for the case of a CRRA utility function $u(.)$, the expression is non-concave.

The equation for the non-negativity constraint on $G_0$ is

$$\begin{cases} b = b^* & \text{for } Y_1 \leq Y_1^* \\ b = u'(R - \frac{Y_1}{1+r}) & \text{for } Y_1 > Y_1^* \end{cases}$$

(3.10)

Similarly, the expression for the non-negativity constraint on $G_1$ is

$$b = \frac{1+\xi}{1+\rho} u'(Y_1)$$

(3.11)
Figure 3-1 presents the three borders as well as the regions delimited by them. The continuous line represents equation 3.9 (credit constraints), the dotted line equation 3.10 (gifts in period 0) and the dash-dotted line equation 3.11 (gifts in period 1). The dashed vertical line at $R(1+r)$ corresponds to the total lifetime resources expressed in terms of period 1 dollars. One noteworthy finding is that, for a constant $Y_1$, increases in the gift and bequest parameter $b$ do not make the liquidity constraints less binding. This is rather surprising as it is generally thought that higher value of the bequest parameter increase the incentives to save, hence alleviating liquidity constraints. In our model of gifts and bequests, this logic does not apply. For $\xi > r$, increases in $b$ have two effects that work in opposite directions. First, we have the standard effect of increases in $b$ implying a stronger motive of giving and hence smaller consumption. Second, increases in $b$ also affect the timing of the act of giving. More precisely, higher values of $b$ strengthen the desire to give early in life under the form of gifts in period 0, rather than leave the money as gifts in period 1. The latter timing effect clearly worsens the impact of the liquidity constraints. For $b$ sufficiently big, this second effect dominates the first one.

Further, within the present scenario of $\xi > r$, the possibility of setting up a trust fund to delay the time of utility realization by $n$ periods is not of a particular interest to the individual. Any additional $n$-period delay to the time the individual perceives the utility of his gift or bequest would decrease his well-being, as the negative effects of the wedge between $\xi$ and $r$ are amplified. Hence, at an optimum, $n = 0$ for $\xi > r$.

A last remark on the timing implications of the present scenario relates to Hurd(1989).
The author assumes \( \xi = \rho \) as well as \( \rho > r \) and uses a simple bequest motive, rather than a more complete gift and bequest motive. The analysis of the present section has shown that individuals characterized by a general gift and bequest motive would prefer to make wealth transfers early in life rather than at the end of their life through bequests. Hence, assuming our model is a better representation of reality, the setup used in Hurd\textsuperscript{(1989)} forces people to make wealth transfers through a very unattractive instrument. Individuals are implicitly forced to keep the resources they want to give to their heirs until the end of their lives, even though they would clearly prefer to give them as soon as possible. Further, even within the bequest-only setup of Hurd\textsuperscript{(1989)} the interpretation of the linear bequest parameter \( b \) as the marginal utility of bequests is incorrect. Rather, the marginal utility of bequests in period \( t \) is correctly summarized by \( b \times \left( \frac{1+r}{1+\rho} \right)^t \).

### 3.3.2 The case of \( \xi \leq r \)

We start our discussion of the scenario \( \xi \leq r \) by stating some basic results: First, notice that for \( \xi = r \), the optimality conditions imply that \( \mu_1 \) and \( \mu_2 \) are either both zero or both positive. This indicates an indifference between gifts in period 1 and adequately discounted bequests in period 2. Hence, the individual only cares about \( G_1 + \frac{B_2}{1+r} \) and not about its components. If \( \lambda_1 = 0 \), the same property extends to gifts in period 0 and the relevant decision variable becomes \( G_0 + G_1 + \frac{B_2}{1+r} \). For \( \lambda_1 > 0 \) on the other hand, we have \( G_0 = 0 \). The mechanism underlying these results is simple. Since wealth accumulates at the interest rate \( r \), and future gifts and bequests are discounted back to the present at the exact same rate \( \xi = r \), the two effects cancel each other. Second, for \( \xi < r \), the optimality conditions imply that \( G_0 = G_1 = 0 \) at any possible optimum. These two observations can be summarized by the following property:

*In a certainty framework, we can w.l.o.g. simplify the optimization problem for \( \xi \leq r \) by dropping all gift variables.*

This result is of substantial practical importance as it easily generalizes to continuous time as well as to the case of life-span uncertainty. The ability to reduce the number of variables is of a substantial interest for the analyst using numerical methods to derive the optimal life-cycle consumption and wealth profiles for an individual with a motive of giving.\textsuperscript{7}

\textsuperscript{7}See appendix B for one illustration of a life-span uncertainty model.
In the same spirit as in the previous section 3.3.1 where we discussed the case of $\xi > r$, we now derive the borders delimiting the different scenarios in terms of the binding of the non-negativity constraints. The preceding analyses allows us to focus on the liquidity and bequest non-negativity constraints. The expression separating the binding and slackness of the credit constraint, can be written as

$$
\begin{align*}
Y_1 &= Y_1^* & \text{for } b < b^{**} \\
\frac{1}{1+\xi} \left(1+\xi\right)^2 &= u'(R - \frac{Y_1}{1+r}) & \text{for } b \geq b^{**}
\end{align*}
$$

(3.12)

where $Y_1^*$ is the solution to $u'(R - \frac{Y_1}{1+r}) = \frac{1+\xi}{1+r} u'(Y_1^*)$ and $b^{**}$ is the solution to $b = \frac{(1+\xi)^2}{(1+r)(1+\rho)} u'(Y_1^*)$. Notice that the $Y_1^*$ is the same as in the previous section for $\xi > r$, but that $b^{**}$ is smaller.\(^8\) Further notice that $Y_1^*$ is independent of $\xi$, but $b^{**}$ is increasing in $\xi$.

Similarly, the function separating the binding and slackness of the bequest non-negativity constraint is defined by

$$
\begin{align*}
b &= b^{**} & \text{for } Y_1 \leq Y_1^* \\
b &= \frac{(1+\xi)^2}{(1+r)(1+\rho)} u'(Y_1) & \text{for } Y_1 > Y_1^*
\end{align*}
$$

(3.13)

This functional form implies that $b$ is decreasing in $Y_1$, and the function is either convex or non-convex depending on the sign of $u''''(\cdot)$. For the special case of a CRRA utility function, the latter function is non-convex.

It is instructive to represent these equations and the scenarios they delimit graphically. In figure 3-2, the continuous curve represents expression 3.12 (credit constraints) while the dotted curve represents expression 3.13 (bequests). Looking at this figure, the comparative statics results with respect to variations in the parameter $b$ become obvious. If the initial optimum is characterized by either a binding credit or bequest constraint, sufficiently strong increases in the bequest parameter $b$ will make the new optimum be part of the unconstrained scenario. Similarly, if both constraints are initially binding, a continuous increase in $b$ first makes the optimum be part of the scenario where only the credit constraint binds and ultimately into the

\(^8\)More precisely, $b^{**} = \left(\frac{1+\xi}{1+r}\right)^2 b^*$. The intuitive reason for this finding is that in the present case of $\xi \leq r$, the solution is anchored around the period 2 variable $B_2$, whereas for $\xi > r$ it is anchored around the period 0 variable $G_0$. 

56
unconstrained scenario. This finding of increases in the linear bequest parameter $b$ making the liquidity constraints less binding has to be seen in stark contrast to the results for $\xi > r$. The reason is that increases in $b$ now have an unambiguous effect on the willingness to save in period 0 because the two effects attached to variations in $b$ play in the same direction: First increases in $b$ increase the desire to leave gifts and bequests. This first mechanism has a non-increasing effect on consumption, and hence a non-decreasing effect on period 0 savings. Second, since $\xi \leq r$, increases in $b$ reinforce the desire to leave late in the life. It is interesting to notice that the above comparative statics findings are completely independent of the value of the utility of consumption discount rate $\rho$.

For $\xi = r$, the individual's problem basically reduces to the choice between consumption on the one hand and the PDV of gifts and bequests on the other. Therefore, the linear parameter $b$ can be seen as a summary statistic for the strength of the gift and bequest motive. For $\xi \neq r$, this property does not hold true. For example, for $\xi < r$ the strength of the bequest motive is effectively measured by $b \left( \frac{1+r}{1+\xi} \right)^2$. The reason for this complication is that for $\xi \neq r$, the parameter $b$ no longer only determines the choice between consumption on the one side and gifts and bequests on the other, but also the choice between gifts and bequests in different periods through its interaction with the interest and discount rates.

The above findings indicate that for $\xi \leq r$ the individual tends to delay the act of giving as long as possible. The same logic also applies to a possible delay in utility realization through the setting up of a trust fund (i.e. choosing $n > 0$). It is trivial to see that at an optimum for
$\xi < r$, in addition to making all transfers to his heirs under the form of bequests, the individual delays the utility realization to a maximum by setting up a trust fund for the longest duration authorized by law. For the special case of $\xi = r$, the possibility of timing the wealth realization through a trust fund has no impact on the level of the bequest as $b = b^{trust}$.

Summarizing the findings of the present section for $\xi = r$, we have the following property: *For $\xi = r$ we can abstract away from both the timing of gifts and bequests problem as well as the question of when the utility of gifts and bequests is realized.*

### 3.4 Conclusion

In the present paper we generalize the standard joy-of-giving model to include both inter-vivos gifts and bequests at death. We discuss different setups that are all characterized by some form of substitutability of gifts and bequests. Using a quasi-linear model in consumption on the one hand, and gifts and bequests on the other, we find that the choice of the discount rate has important implications on the timing of gifts and bequests. We find that the interpretation of this parameter crucially depends on the choice of the utility discount rate for gift and bequest. We show that for the analyst interested in inserting an easily interpretable motive of giving into his model, using the interest rate as the utility discount rate for gifts and bequests is the best choice. The use of interest rate discounting of gifts and bequests allows an interpretation of the bequest parameter as a summary statistic of the strength of the motive of giving. Further, the ability to drop the inter-vivos gift variables from the problem without any loss of generality is obviously of quite some practical interest, as it simplifies analytical and numerical solutions to the individual optimization problem considerably. For all other choices of the discount rate, we find that these properties do not hold. If the discount rates is bigger than the interest rate, liquidity constraints have an impact on the PDV of gifts and bequests. If it smaller than the interest rate, the timing of the utility realization becomes an important issue, and hence the possibility of setting up a trust fund has a strong influence.
Chapter 4

Delays in claiming Social Security benefits

4.1 Introduction

Social Security (SS) is the largest entitlement program in the United States today, providing income support for retired and disabled workers and their families. SS benefit payments in 1996 amounted to $350 billion, which is over 22% of the federal budget and nearly 5% of GDP; this represents a doubling of payments as a share of GDP over the past 35 years. The concurrent growth in this program and decline in the labor force participation of older men has motivated an extensive literature investigating how SS influences retirement behavior. There is another large literature investigating the transfers induced by SS across and within cohorts.

While the work in this area encompasses a range of empirical strategies and data sets, one common feature has been the assumption that dates of retirement and benefits claiming are identical. That is, estimates of the retirement incentives of this system have been computed assuming that individuals claim SS benefits as soon as they are eligible. However, as with other social insurance programs, there is a take-up decision associated with claiming SS benefits. Individuals need not claim their benefits immediately upon retirement. By delaying claiming, workers increase the benefits paid to them and their spouses, through the actuarial adjustment. As we demonstrate below, it is optimal in a wide variety of cases to delay claiming benefits for a period of time after retirement.
This *dynamic take-up* consideration suggests that standard computations of both the retirement incentives of SS and the redistribution through SS may be biased. Moreover, we are not aware of any in-depth analysis of this take-up behavior. An examination of whether observed claiming patterns are consistent with rational choice theory may have important implications for aspects of SS design and reform.

The purpose of our paper is to investigate delays in SS benefits claiming and to explore their implications. We do so in two steps. First, in section 4.2, we provide relevant institutional background on the SS program. We highlight the fact that retirement provides only a necessary, and not a sufficient, condition for claiming SS benefits. We briefly review the SS literature, emphasizing areas where realistic consideration of claiming behavior can affect analysis.

In section 4.3, we turn to a theoretical examination of claiming delays. We begin with a discussion of the benefit rules to explore how worker characteristics such as mortality expectations, wealth, age difference with spouse, and relative earnings of spouses may influence claiming delays. Then we use simulations of financial gains from delay to generate cross-sectional predictions that can be tested in our empirical analysis. Finally, we also present simulation results based on an expected utility maximization model with liquidity constraints, as we recognize that financial calculations in general understate the incentives to delay relative to the optimization of a risk averse utility function. This is because SS provides a real annuity valued by risk averse individuals with an uncertain date of death; individuals buy more of this annuity by delaying, so delays are more attractive with risk aversion.

We note that while it is easy to focus on delays, as we frequently do in the paper, our analysis is in fact based on a theory of claiming, not a theory of delays. This becomes obvious in section 4.3 when we look at the effect of varying the retirement age; as we show, retiring one year later does not affect the optimal age of claiming, but it does shorten the optimal delay by one year, since the delay is the period of time after retirement and before claiming.

Section 4.4 summarizes our finings and considers the implications for previous research on SS.
4.2 Background

4.2.1 Institutional Features

Understanding the motivation for our analysis requires a brief overview of how benefits are determined. Individuals are fully insured for retired worker benefits once they have worked 40 quarters in the covered sector. Benefits are computed as follows: nominal taxable annual earnings before age 60 are converted into age 60 dollars using a wage index, the 35 highest years of indexed earnings (indexed before age 60 but not after) are averaged and divided by 12 to generate the Average Indexed Monthly Earnings (AIME), and a non-linear formula is applied to the AIME to generate the Primary Insurance Amount (PIA) on which monthly benefits are based.

Fully insured individuals can claim retired worker benefits if they meet two criteria. First, they must be at least age 62. Second, individuals must pass an earnings test: if their earnings exceed a ceiling amount, benefits are reduced by 50 cents for each additional dollar of earnings (if age 62-64) or 33 cents for each additional dollar of earnings (if age 65-69). In 1995, this ceiling was $8,160 for 62-64 year olds, and $11,280 for 65-69 year olds.\(^1\)

Monthly benefits also depend on age at claiming. If individuals claim at the normal retirement age (NRA) of 65, the monthly benefit equals 100% of the PIA.\(^2\) If they claim between age 62 and age 65, there is an actuarial reduction in the benefit of 5/9% for each month of claiming before age 65 up to 36 months and 5/12% for every month of claiming beyond 36 months. Thus workers claiming on their 62nd birthdays currently have a benefit equal to 80% of PIA. If they claim after age 65, there is a delayed retirement credit. For a pensioner turning 65 in 1998, the credit is of 5.5% per year. This is significantly smaller than the actuarial reduction before age 65, generating a kink at age 65 in the schedule of benefits as a function of claim date.\(^3\)

A key feature of this institutional structure is that retirement need not be concurrent with

\(^1\)There is also a monthly earnings test that individuals may use for one year only, usually the year of retirement. In 1995, the monthly earnings test ceiling was $680 for 62-64 year olds and $940 for 65-69 year olds. In the year that the monthly earnings test is applied, individuals may have annual earnings above the annual earnings test ceiling but may still receive full benefits for any months in which they earned less than the monthly earnings test ceiling.

\(^2\)The NRA is scheduled to rise in a series of steps, reaching 67 for workers attaining age 62 in the year 2022 or later.

\(^3\)The delayed retirement credit is rising over time, and is scheduled to reach 8% per year for those attaining age 62 in the year 2005 or later.
claiming. For example, if individuals retire at age 62, they need not claim on their 62nd birthday. As we document below, in many cases it may be optimal for such individuals to delay, as total benefits received may be increased by delaying for some months thereby raising the benefit level.

Calculating the advantages of delaying is complicated by the family benefits structure of SS. Spouses age 62 and above of fully insured workers are eligible for dependent spouse benefits and may also be entitled to retired worker benefits; however, spouses receive only the larger of the two amounts. The dependent spouse benefit is 50% of the retired worker's PIA, can be claimed once the dependent spouse is 62 and the worker has claimed, and is subject to an actuarial reduction if the dependent spouse claims before 65. Surviving spouses of retired workers are entitled to a survivor benefit of 100% of the retired worker's PIA; the benefit can be claimed once the survivor is 60 and may be reduced depending on the survivor's age when benefits begins and the worker's ages at claiming and at death. Claiming the survivor benefit implies foregoing the survivor's retired worker or dependent spouse benefit. We return to the question of how family benefits affect incentives for delays below.

4.2.2 Previous Literature

The concern that this complicated benefits structure might have important implications for retirement incentives has motivated an enormous literature on the effect of SS on retirement, reviewed in Hurd (1990) and Diamond and Gruber (1997). The first strand of this literature uses aggregate information on the labor force behavior of workers at different ages over time to infer the impact of SS. Hurd (1990) and Ruhm (1994) find a spike in the age pattern of retirement at 62 and show that this peak has grown over time as SS benefits have increased; Burtless and Moffitt (1984) show that there was no peak before claiming at 62 became an option. Of course, for workers for whom the actuarial adjustment, additional tax, and AIME recomputation is fair on average, there is no reason for SS to induce a spike at age 62; Diamond and Gruber (1997) note that the implicit tax rate on work at age 62 is negative for the median worker. But the spike could be driven by an interaction of SS with liquidity constraints; indeed, Kahn (1988) finds a spike in retirement at age 62 for low wealth workers, but not for high wealth workers.
There is also a spike in retirement at age 65, which is consistent with the unfair actuarial adjustment for work beyond age 65. Blau (1994) finds that nearly 25% of the men in the labor force on their 65th birthday retire in the next quarter; this hazard rate is 2.5 times as large as the rate in surrounding quarters. However, Lumsdaine, Stock, and Wise (1990) document that this penalty alone cannot account for "excess" retirement at age 65; nor can incentives in private pension plans or the availability of health insurance through Medicare. This does not rule out a role for Social Security; by setting up the "focal point" of a normal retirement age, the program may be a contributing causal factor in explaining this spike.

The second strand of this literature uses micro-data sets with SS benefit determinants or ex-post benefit levels to measure the incentives to retire across individuals, then estimates retirement models as a function of these incentives.\(^4\) While the techniques differ across papers, the conclusions are similar: SS has large effects on retirement, but they are small relative to the time trend.\(^5\) For example, Burtless (1986) found that the 20% benefit rise between 1969 and 1973 raised the probability of being retired at 62 and at 65 by 2 percentage points; however, labor force participation fell by 6 points during this period, so SS explains only 1/3 of the change.\(^6\)

This second strand of the literature has been criticized on two grounds. The first is that the key regressor, SS benefits, is a non-linear function of past earnings, and retirement propensities are correlated with earnings. This criticism is raised most compellingly by Krueger and Pischke (1992), who use a natural experiment provided by the end of double-indexing for the "notch generation" retiring in the late 1970s and early 1980s. For this cohort, SS benefits were greatly reduced relative to what they would have expected, yet the fall in labor force participation continued unabated. This raises important concerns about the identification of earlier estimates.

\(^4\)The data used are generally the Retirement History Survey (Boskin and Hurd, 1978; Burtless, 1986; Burtless and Moffitt, 1984; Hurd and Boskin, 1984; Fields and Mitchell, 1984; Blau, 1994), though some authors have used the National Longitudinal Survey of Older Men (Diamond and Hausman, 1984) or the Survey of Consumer Finances (Samwick, 1993).

\(^5\)The earliest studies (Boskin and Hurd, 1978; Fields and Mitchell, 1984) used standard linear or non-linear regression techniques. Later research (Burtless, 1986; Burtless and Moffitt, 1984) used non-linear budget constraint estimation to capture the richness of Social Security's effects on the opportunity set. The most recent work (Diamond and Hausman, 1984; Hausman and Wise, 1985; Samwick, 1993; Blau, 1994) uses dynamic estimation of the retirement transition.

\(^6\)One exception is Hurd and Boskin (1984), who claim that the large benefits increases of the 1969-1973 period can explain all of the change in labor force participation in those years.
The second problem with much of this literature is that it focuses on only one of the two key SS variables: it includes SS benefits or wealth but ignores the implicit SS tax/subsidy rate on further work. In theory, both factors affect retirement. Studies including the accrual rate have found it to be significant (Fields and Mitchell, 1984; Samwick, 1993; Krueger and Pischke, 1992). Stock and Wise (1990) note that the correct variable is not the accrual rate, but the return to working this year relative to retiring at some future optimal date.

There is a third potential weakness of previous studies that has been ignored thus far: the possible endogeneity of the timing of SS benefits claiming and therefore of the benefit level. The key independent variable in cross-sectional estimation, SS benefits, may confound potentially exogenous characteristics which determine benefits, such as lifetime earnings, with the endogenous take-up decision. For example, consider two individuals who retire at the same point (their 61st birthday) and are identical in every respect except time preference. Impatient individual B claims benefits at 62, while patient individual A delays until age 65 and receives a higher benefit. Regression analysis would show that higher SS benefits do not cause earlier retirement. But in fact these two individuals have the same PIA\(^s\) and thus face the same menu of retirement benefit choices. This suggests that by using actual SS benefits received rather than PIA to model retirement incentives, previous studies may have misstated the incentives.

The literature which has used accrual rates or option values is also affected by ignoring the endogenous claiming decision. These accrual/option value measures universally assume that retirement and claiming are on the same date. But if claiming can be delayed, it affects the accrual rate or option value of SS benefits, leading again to mismeasurement of the key regressor.

Another strand of the literature stresses redistribution within and across generations arising from SS (Hurd and Shoven 1985, Boskin et al. 1987, Steuerle and Bakija 1994). This literature has found significant redistribution from recent cohorts to older cohorts, from low earners to high earners (in previous cohorts) and high earners to low earners (in current and future cohorts), from short lived to long lived, and from single to married. But this literature has also ignored delayed claiming. The ability to delay claiming increases the redistribution of the system, for example, from short lived to long lived and from singles to married with a dependent spouse; SS also differentially affects those who are liquidity constrained and those who are not.
A full analysis of the problem of delayed claiming would model jointly the retirement and the claiming decision. Such a model is beyond the scope of the current effort. Rather, our goal is to demonstrate theoretically that delayed claiming is often optimal.

4.3 When is it Optimal to Delay Claiming?

In this section, we illustrate the incentives for claiming delays under the US Social Security System by presenting two simulation approaches. The first technique is a purely financial calculation of the expected present discounted values (EPDV) of future net benefit streams for a single worker and for a married couple. We examine the variation in incentives for claiming delays among subgroups of the population with different characteristics.

The second technique is expected utility maximization under liquidity constraints. This technique has the advantage of capturing the value of SS as a real annuity to a risk averse person with an uncertain date of death. However, due to computational complexity, we estimate the expected utility maximization model for a single worker only, leaving a full household optimization model for future work. Before turning to the simulations, we review the benefit rules to explain how factors such as mortality expectations affect incentives for claiming delays.

4.3.1 Benefit Rules

Consider a single person who is fully insured for retired worker benefits, has just turned 62, and has stopped working. He could claim benefits immediately and begin receiving a monthly benefit of .8*PIA.\(^7\) Alternatively, he could delay claiming for some period of time. Consider the effect of waiting one year and claiming on his 63rd birthday: he forgoes one year of benefits, but receives a monthly benefit of .867*PIA for the rest of his life, an increase of 8.33%. Thus, claiming delays involve the sacrifice of current benefits for a higher future benefit level.

To evaluate this tradeoff, he considers his life expectancy and discount rate. A longer life expectancy creates a stronger incentive to delay because the higher future benefit level is expected to last longer. A lower discount rate creates a stronger incentive to delay because

\(^7\)The rules for male and female retired workers are the same; we refer to a male for ease of exposition. The rules are slightly different with birthdays on the first and second days of the month. Benefits are adjusted annually to reflect cost-of-living adjustments.
future benefits are valued more highly. The discount rate, in turn, depends at least in part on wealth and bequest motives. A person with low wealth has a higher short-run discount rate due to the difficulty in maintaining a reasonable consumption level while delaying benefits. A person with moderate wealth has a short-run discount rate roughly equivalent to the market rate (adjusted for taxes on the return on wealth). However, his long-run discount rate is lower because SS benefits are paid as a real annuity, a form unavailable in the market, which makes for higher valuation of distant benefits; this effect is weaker if he has a bequest motive, assuming that the utility from bequests shows less risk aversion than the utility from consumption. For those with high wealth and linear utility of bequests, there is no valuation of the annuity aspect of SS, since consumption is never reduced to just SS benefits and variation in bequests provides length-of-life insurance. Therefore, we expect an inverse u-shaped pattern of claiming delays as wealth rises: those with low wealth have short delays, those with moderate wealth have long delays, and those with high wealth have medium delays.  

The question of how to value a marginal annuity is controversial. Bernheim (1987) argues that only the discount rate, and not survival probabilities, should be used to value a marginal annuity stream for a risk averse individual with an uncertain life span. However, Bernheim assumes that annuities are not available on the margin and that there are no bequest motives. In this paper, we recognize that individuals can vary their annuity holdings on the margin by delaying social security claiming. In addition, we showed in chapter 2 that for a sufficiently strong bequest motive, the correct way to value a marginal annuity is actuarial valuation, which takes into account both survival probabilities and the discount rate.

A couple's claiming delays are affected by two additional factors: the age difference between the spouses and the relative PIAs of the spouses. To understand the effects of these factors, we first consider variation in the relative ages of a husband and wife in a one earner couple and then consider variation in the relative PIAs of a husband and wife of the same age.

Consider a retired, 62 year old husband with a 62 year old wife who has never worked. If they both claim at 62, he receives a retired worker benefit for life and she receives a dependent spouse benefit for as long as they both are alive. If he dies first, she receives a survivor benefit

---

*We ignore possible complications from the tax treatment of benefits; benefits are untaxed for low enough income and taxes only affects the analysis if applicable tax rates change over time.

*In this case, the dependent spouse benefit is .75*5*PIA; the benefit is .5 *PIA if she claims at 65 and the
equal to his PIA, though it is reduced if she receives it before age 65 and is limited to the benefit he received.\textsuperscript{10} If he delays until 63, she must also delay; they forego worker and spouse benefits for one year and then receive a worker benefit, dependent spouse benefit, and eventually a survivor benefit which are increased by 8.33\%, 11.11\%, and 5.05\%, respectively.\textsuperscript{11}

Now consider the effect of his delaying until 63 when the wife is one year younger. The increase in the survivor benefit is the same, 5.05\%, but it is more valuable in EPDV terms with one year age difference because there is a longer expected time of receipt. In general, men with a larger positive age difference have a stronger incentive to delay.\textsuperscript{12} We also note that there are life expectancy difference across women other than those associated with age and that men whose wives have a higher life expectancy have a stronger incentive to delay.

Now we examine a couple with two fully insured workers of the same age to see how the ratio of their PIAs affects delays. The story is complicated since there are four alternatives: both might claim at 62, both might delay, or one might claim at 62 and the other might delay. Here we focus on the husband’s decision only. However, the claiming decision may be made jointly and it may be optimal, for example, for the wife to claim immediately and the husband to delay.\textsuperscript{13}

Raising the ratio PIAW/PIAH has two effects on the husband’s incentives to delay. First, there is a survivor benefit effect. With a ratio below some cutoff, the wife receives a survivor benefit based on his PIA and his delay, as in the one-earner case; with a ratio above this cutoff, actuarial reduction is 25/36\% per month before the NRA.

\textsuperscript{10}In this case, the survivor benefit is .825*PIA. The actuarial reduction is 19/40\% per month before the NRA; benefits can first be claimed at 60. The benefit is limited to the amount of the husband’s benefit, but can be no less than .825*PIA. If the worker dies before the NRA, the adjustment limiting the survivor benefit is based on the number of months he received benefits.

\textsuperscript{11}She is required to delay because she is claiming dependent spouse benefits; this would not be the case if she were claiming her own retired worker benefits. The new values of the three benefits are .867*PIA, .833*.5*PIA, and .867*PIA.

\textsuperscript{12}There may be a second effect due to the fact that a delay by the husband forces the wife to delay in the same age case, raising her benefit, but does not force her to delay in the other case. In both the simulations and the empirical work, this effect seems not to matter. Therefore, we leave a discussion of it to the section on the PIA ratio, where the effect seems more important.

\textsuperscript{13}If she claims at 62 and he delays until 63, they forego his benefit for one year and raise his benefit and her survivor benefit, assuming her PIA is small enough relative to his. If he claims at 62 and she claims at 63, they forego her benefit for one year and raise her benefit, assuming his PIA is large enough relative to hers. Since the latter option raises only her benefit (received as long as they both are alive), while the former option raises his benefit (received as long as he lives) and her survivor benefit (received for the time she outlives him), it may well be optimal for him to delay and her to claim immediately.
the wife receives her own retired worker benefits after his death.\textsuperscript{14} Therefore, the incentive to delay claiming to raise the survivor benefit decreases as the ratio rises, but not evenly. This is relevant for high ratios but not low ones.

Second, there is a dependent spouse benefit effect. With a ratio below some cutoff, the wife claims dependent spouse benefits; in this case, a delay by the husband forces her to delay, raising her benefit.\textsuperscript{15} This effect is not relevant for every woman: it is binding only for women who prefer not to delay, it does not apply to couples with a large age difference, and it is possible for women to claim retired worker benefits at 62 and to receive an additional benefit once eligible for dependent spouse benefits.\textsuperscript{16} With a ratio above the cutoff, the wife claims retired worker benefits, so a delay by the husband does not force her to delay or affect her benefit. Thus, the incentive to delay to raise the wife's benefit ends when the ratio rises above the cutoff.

Both effects become less important as the PIA ratio increases. This suggests that delays should decrease as the PIA ratio increases. In fact, this is not necessarily correct. For a given PIA ratio, each effect taken on its own leads to a different choice of the optimal claiming delay. As the PIA ratio increases, both effects lessen in importance, but since this may happen at different rates, it is possible for the optimal delay to switch from being driven by one effect to the other. Since we have no general predictions about which effect leads to a longer delay and which is decreasing faster as the PIA ratio increases, we cannot simply predict how delays will change as the ratio rises. The simulation results and empirical work can shed some light on the question.

Having discussed the factors that affect a couple's claiming decisions, we can consider how delays differ across single and married men. In general, having a wife provides a greater incentive to delay, as a delay raises the benefits the wife receives. However, in some special cases, having a wife may lead to a shorter delay. For example, if the wife is older and the PIA ratio is low, the husband may claim benefits early so that the wife can also claim benefits.

To recap, our discussion of the benefit rules has clarified how the incentive to delay is

\textsuperscript{14} This cutoff depends on when both spouses claim. If both spouses claim at the same time, the cutoff is 1; if the wife claims at 62 and the husband at 65, the cutoff would be 1/8 or 1.25, since the survivor benefit is $\text{max}(.8\text{PIAW}, \text{PIAH})$ in this case.

\textsuperscript{15} Cutoff is .5 if both spouses claim at 65.

\textsuperscript{16} If, for example, she claimed retired worker benefits at 62 then dependent spouse benefits at 63, she would receive $\text{.8PIAW}$ during age 62 and $.833\times .5\times \text{PIAH} - .067\times \text{PIAW}$ thereafter; an adjustment is made to the dependent spouse benefit because she received benefits at age 62.
affected by various factors. First, the incentive to delay is stronger if the claimant has a longer life expectancy. Second, delays follow an inverse u-shaped pattern as wealth rises. Third, the incentive to delay is stronger if the claimant has a larger positive age difference with his wife. Fourth, the effect of a higher PIA ratio on delays is variable. Fifth, married men generally have a stronger incentive to delay claiming than single men. The simulations in the next two subsections allow us to further develop these points and to measure the gain from choosing the optimal delay.

4.3.2 Financial calculations

We begin with financial calculations which measure how the EPDV of SS benefits varies with months of delay. The EPDV is defined as the discounted flow of future potential benefits paid to the family minus social security taxes paid by the family. The program we developed first computes, for every future month at which a family member may be alive, the benefits corresponding to all possible survival and death patterns in the family, then adjusts them for survival probabilities and inflation and discounts them back to the base year. This computation is repeated for each possible month of benefit claim by the prime earner. Appendix C provides a more detailed explanation of these calculations.

We focus on a household whose prime earner is a male born on January 2, 1930 and alive at age 62. The base year for the simulations is 1992. We make the following assumptions in all cases unless otherwise noted. We assume that the wife was born on January 2, 1932 and that the couple has no dependents. We assume that this is a one-earner couple, that the husband stops work on his 62nd birthday, and that the husband's wage history corresponds to the economy-wide median earnings profile for his age cohort until from age 20 to age 50 and is constant in real terms thereafter. We assume that the wife claims benefits as soon as possible. Finally, we assume that the household's discount rate is 3% and that mortality risks correspond to the Social Security Administration's sex- and cohort-specific survival tables.

Table 4.1 presents the EPDV calculations. Column (b) shows the optimal delay in months.

---

17 Deaths may occur on a monthly basis. We use all Social Security rules including future planned adjustments as of 1996.
18 A wife claiming retired worker benefits claims at age 62. A wife claiming dependent spouse benefits claims at the later of age 62 or her husband's date of claim.
<table>
<thead>
<tr>
<th>Case</th>
<th>Optimal Delay (months)</th>
<th>EPDV Zero Delay</th>
<th>EPDV Optimal Delay</th>
<th>Change (e)</th>
<th>PIA</th>
<th>Change in EPDV/PIA (g)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>One-Earner Couple</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base Case</td>
<td>36</td>
<td>206,268</td>
<td>212,538</td>
<td>6,270</td>
<td>963</td>
<td>6.51</td>
</tr>
<tr>
<td>High Mortality Risk</td>
<td>34</td>
<td>196,370</td>
<td>200,204</td>
<td>3,834</td>
<td>963</td>
<td>3.98</td>
</tr>
<tr>
<td>Low Mortality Risk</td>
<td>36</td>
<td>217,369</td>
<td>226,381</td>
<td>9,012</td>
<td>963</td>
<td>9.36</td>
</tr>
<tr>
<td>High Discount Rate</td>
<td>0</td>
<td>151,247</td>
<td>151,247</td>
<td>0</td>
<td>963</td>
<td>0.00</td>
</tr>
<tr>
<td>Low Discount Rate</td>
<td>36</td>
<td>263,513</td>
<td>280,320</td>
<td>16,807</td>
<td>963</td>
<td>17.45</td>
</tr>
<tr>
<td>High Earnings</td>
<td>36</td>
<td>229,067</td>
<td>236,043</td>
<td>6,976</td>
<td>1,069</td>
<td>6.53</td>
</tr>
<tr>
<td>Low Earnings</td>
<td>36</td>
<td>110,809</td>
<td>114,184</td>
<td>3,375</td>
<td>518</td>
<td>6.52</td>
</tr>
<tr>
<td>Retire at 63</td>
<td>24</td>
<td>212,906</td>
<td>217,073</td>
<td>4,167</td>
<td>969</td>
<td>4.30</td>
</tr>
<tr>
<td>Retire at 64</td>
<td>12</td>
<td>218,850</td>
<td>220,556</td>
<td>1,706</td>
<td>919</td>
<td>1.86</td>
</tr>
<tr>
<td>Retire at 65</td>
<td>0</td>
<td>222,757</td>
<td>222,757</td>
<td>0</td>
<td>898</td>
<td>0.00</td>
</tr>
<tr>
<td><strong>Single</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base Case</td>
<td>10</td>
<td>117,924</td>
<td>118,126</td>
<td>202</td>
<td>963</td>
<td>0.21</td>
</tr>
<tr>
<td>High Mortality Risk</td>
<td>0</td>
<td>102,113</td>
<td>102,113</td>
<td>0</td>
<td>963</td>
<td>0.00</td>
</tr>
<tr>
<td>Low Mortality Risk</td>
<td>23</td>
<td>134,487</td>
<td>136,473</td>
<td>1,986</td>
<td>963</td>
<td>2.06</td>
</tr>
<tr>
<td>High Discount Rate</td>
<td>0</td>
<td>81,884</td>
<td>81,884</td>
<td>0</td>
<td>963</td>
<td>0.00</td>
</tr>
<tr>
<td>Low Discount Rate</td>
<td>36</td>
<td>143,339</td>
<td>146,346</td>
<td>3,007</td>
<td>963</td>
<td>3.12</td>
</tr>
<tr>
<td>High Earnings</td>
<td>10</td>
<td>130,955</td>
<td>131,161</td>
<td>206</td>
<td>1,069</td>
<td>0.19</td>
</tr>
<tr>
<td>Low Earnings</td>
<td>9</td>
<td>63,352</td>
<td>63,455</td>
<td>103</td>
<td>518</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Assumptions for Calculations: (1) Household’s prime earner is a male born January 2, 1930 and alive at age 62. (2) Base year for simulations is 1992; all values are in $1992. (3) Wife is born January 2, 1932; wife has never worked; couple has no dependents. (4) Worker retires on 62nd birthday; wage history corresponds to economy-wide median earnings profile for worker’s cohort from ages 20 to 50 and is constant in real terms thereafter. (5) Wife claims as soon as possible (the later of her 62nd birthday or husband’s date of claim). (6) Household discount rate is 3%; mortality risks correspond to Social Security Administration’s sex- and cohort-specific survival tables.

Table 4.1: Financial Calculations: One-Earner Couple and Single Worker
Column (c) shows the EPDV with no delay, column (d) shows the EPDV at the optimal delay, and column (e) shows the difference between these two, which is the value of delayed claiming. Column (g) presents the change in EPDV scaled by PIA. As the monthly retired worker benefit is equal to the PIA if the worker claims at age 65, the number in column (g) can be loosely interpreted as the number of additional months of retired worker benefits received in expectation over the worker’s lifetime as a result of choosing the optimal delay.\textsuperscript{19}

The first row shows the results for the base case with a one-earner couple. In this case, it is optimal for the husband to delay claiming by 36 months to age 65. The delay raises the EPDV of benefits by $6,270, or 651\% of PIA. This result and those that follow suggest that optimal claiming delays are frequently long and that gains from delay are moderate for a one-earner couple. In the base case, a delay of 36 months would result in an increase of $232 in the couple’s monthly benefit check, from $1132 to $1364. Figure 4-1 illustrates the EPDV of benefits as a function of delay for a one-earner couple in the base case.

The next six rows of Table 4.1 show the effect of varying the mortality risk, discount rate, and earnings level. We leave the discussion of these factors to the single worker case; due to a kink in the actuarial adjustment schedule at age 65, optimal delays in the one-earner couple cases bunch up at 36 months, making it difficult to see the effect of these factors. Note that the gains from delay are significant in many of these cases.

The final one-earner couple cases illustrate the effect of varying the age of retirement. We do not present results for the single worker case because optimal claiming age is independent of retirement age for a single individual (for a given PIA); when the worker retires one year later, his optimal delay is one year shorter (if positive). For the case of married couples, the relationship may not be exact, though it is in the simulations: the optimal delay is 24 months if he retires at 63, 12 months if he retires at 64, and 0 months if he retires at or after 65.

Next we examine results for the single worker cases. In general, delays are shorter and gains from delays are much smaller compared to the one-earner couple cases. This is due to the fact that with couples, delays raise not only the retired worker benefit, but also the survivor benefit and potentially the dependent spouse benefit; this is consistent with our earlier statement that married men have a stronger incentive to delay than single men. In the single worker base case,

\textsuperscript{19}Of course, this relationship is not exact unless the worker claims at age 65.
the optimal delay is 10 months and the gain from delay is $202, or 21% of PIA. The change in the EPDV as delay increases for a single worker in the base case is shown in Figure 4-2.

The next two rows explore the effect of varying mortality risk.\(^{20}\) As expected, we find that a longer life expectancy leads to longer delays. With increased mortality risk, the optimal delay falls to 0 months, while with decreased mortality risk, delay rises to 23 months. The gain from delay in the low mortality risk case is $1,986, or 206% of PIA.

\(^{20}\)Mortality risk is altered by multiplying the number of deaths per period by a constant: 0.84 for low mortality, 1.16 for high mortality. Starting with equal numbers of age 62 high and low mortality types, the age 70 population would be 53% low mortality types, 47% high mortality types. With multipliers of 0.70 and 1.30, the age 70 population would be 55% low mortality types, 45% high mortality types. Note that calculations are for given mortality expectations at age 62 with no later health news. News might lead an individual to change his planned delay and a full model would include these option values. Our calculations are only suggestive of the dynamic problem of mortality expectations, which is presumably reflected in the empirical work.
The following two rows of the table show the effect of varying the discount rate.\footnote{The discount rate is 1\% in the low discount rate case and 6\% in the high discount rate case.} In the preceding subsection explaining the benefit rules, we described how a lower discount rate leads to longer delays.\footnote{The second prediction is that delays follow an inverse u-shaped pattern as wealth rises; as discussed above, this depends on the discount rate, the valuation of annuities, and bequest motives. However, simple financial calculations only allow for the effect of the discount rate.} The simulations confirm this: in the high discount rate case, delay drops to 0 months, while in the low discount rate case, delay rises to 36 months. The gain from delay in the low discount rate case is $3,007, or 312\%$ of PIA.

The last two rows in the table illustrate the irrelevance of the earnings level for the optimal delay. The optimal delay is approximately 10 months whether we consider a person at the 10th
percentile of earnings, at the median, or at the 90th percentile. This is because the PIA scales the expression without changing the shape of the time pattern, apart from rounding.

Table 4.2 presents results for married couples varying the age difference and the PIA ratio. As expected, we find that an increase in the age difference between the spouses leads to longer delays. For example, with a PIA ratio of .5, the optimal delay is 0 months if the wife is five years older, 12 months if she is two years older, and 36 months if she is the same age or younger. The gain from delay when the wife is the same age and the PIA ratio is .5 is $5,833, or 606% of PIA.

To explore the role of the PIA ratio, we examine the cases with an older wife, since the optimal delay is usually 36 months in the simulations where the wife is the same age or younger. As discussed above, we have no simple analysis of the effect of a rising PIA ratio on delays. We find that a higher PIA ratio leads to longer delays in the simulations we have performed. For example, with a wife who is two years older, delay is 12 months if the ratio is 0 or .5 and 36 months if the ratio is 1. The gain from delay with a ratio of 1 is $5,251, or 545% of PIA.

For a two-earner couple, the pattern of EPDV as a function of delay may be bimodal, as shown in Figure 4-3. As explained above, the husband's delay affects the survivor benefit and potentially the dependent spouse benefit, depending on the age difference and PIA ratio. A delay of x months may be best when looking only at its effect on the survivor benefit, while a delay of y months may be best when looking only at its effect on the dependent spouse benefit. Variables such as life expectancy and age difference determine which effect dominates in each case.

Before moving to utility maximization, we make one additional calculation. As there is substantial claiming at 62 in the data, we calculate what combination of mortality multiplier and discount rate just makes such behavior optimal for the base case one-earner couple and single worker. The results are shown in Table 4.3. As we would expect, with a lower mortality risk, a higher discount rate is required to make claiming at 62 optimal. For a couple, claiming at 62 is optimal with a mortality multiplier of 0.70 and a discount rate of 6.3 percent or more.

---

23 We construct earnings histories for the 10th and 90th percentile earnings level by taking the relative position in the last year of earnings (year worker turns 61) to fix the level of the earnings profile and then copying the shape of the earnings profile from the baseline scenario. Delay would be 10 months in all cases except for rounding in the benefit rules.
<table>
<thead>
<tr>
<th>Case: Wife's PIA / Husband's PIA (a)</th>
<th>Optimal Delay (months) (b)</th>
<th>EPDV</th>
<th>Change in EPDV /PIA (f)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Zero Delay (c)</td>
<td>Optimal Delay (d)</td>
<td>Change (e)</td>
</tr>
<tr>
<td><strong>Wife 5 years older</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>207,472</td>
<td>0</td>
</tr>
<tr>
<td>0.5</td>
<td>0</td>
<td>207,472</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>36</td>
<td>240,404</td>
<td>3,982</td>
</tr>
<tr>
<td><strong>Wife 2 years older</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>12</td>
<td>212,513</td>
<td>1,634</td>
</tr>
<tr>
<td>0.5</td>
<td>12</td>
<td>212,513</td>
<td>1,634</td>
</tr>
<tr>
<td>1</td>
<td>36</td>
<td>253,483</td>
<td>5,251</td>
</tr>
<tr>
<td><strong>Wife same age</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>35</td>
<td>208,602</td>
<td>5,525</td>
</tr>
<tr>
<td>0.5</td>
<td>36</td>
<td>211,712</td>
<td>5,833</td>
</tr>
<tr>
<td>1</td>
<td>36</td>
<td>260,993</td>
<td>5,832</td>
</tr>
<tr>
<td><strong>Wife 2 years younger</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>36</td>
<td>206,268</td>
<td>6,270</td>
</tr>
<tr>
<td>0.5</td>
<td>36</td>
<td>206,141</td>
<td>6,022</td>
</tr>
<tr>
<td>1</td>
<td>36</td>
<td>242,497</td>
<td>6,021</td>
</tr>
<tr>
<td><strong>Wife 5 years younger</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>36</td>
<td>203,563</td>
<td>5,720</td>
</tr>
<tr>
<td>0.5</td>
<td>36</td>
<td>198,902</td>
<td>6,017</td>
</tr>
<tr>
<td>1</td>
<td>36</td>
<td>217,843</td>
<td>5,643</td>
</tr>
</tbody>
</table>

Assumptions for Calculations: (1) Household’s prime earner is a male born January 2, 1930 and alive at age 62. (2) Base year for simulations is 1992; all values are in $1992. (3) Wife is born on January 2 of 1925, 1928, 1930, 1932, or 1935 (depending on case); couple has no dependents. (4) Worker retires on 62nd birthday; wage history corresponds to economy-wide median earnings profile for worker’s cohort from ages 20 to 50 and is constant in real terms thereafter. (5) Wife claims as soon as possible (the later of her 62nd birthday or husband’s date of claim if claiming dependent spouse benefits, her 62nd birthday if claiming retired worker benefits). (6) Household discount rate is 3%; mortality risks correspond to Social Security Administration’s sex- and cohort-specific survival tables. (7) PDV of wife’s SS contributions for work past January 2, 1992 is $8,320 for a wife born in 1932 and $20,122 for a wife born in 1935 if the PIA ratio is 1, 1/3 of these values if the ratio is 0.5.

Notes: (1) Husband’s PIA is $963. (2) Bold row indicates base case from table 4.1.

Table 4.2: Financial Calculations: Two-Earner Couple
or with a multiplier of 1.30 and a discount rate of 4.3 percent or more; for a single worker, claiming at 62 is optimal with a multiplier of 0.70 and a discount rate of 5.7 percent or more, or with a multiplier of 1.30 and a discount rate of 0.2 percent or more.

4.3.3 Expected Utility Maximization

The EPDV results in Tables 4.1 and 4.2 show that in many cases it is optimal to delay claiming and that the gains from delays can be large in some case. However, if individuals are risk averse, these calculations understate the gains from delays. SS provides a real annuity valued by risk
<table>
<thead>
<tr>
<th>Mortality Multiplier</th>
<th>Discount Rate</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>One-earner Couple</td>
<td>Single Worker</td>
<td></td>
</tr>
<tr>
<td>0.70</td>
<td>6.30</td>
<td>5.70</td>
<td></td>
</tr>
<tr>
<td>0.84</td>
<td>5.80</td>
<td>4.90</td>
<td></td>
</tr>
<tr>
<td>1.00</td>
<td>5.30</td>
<td>3.70</td>
<td></td>
</tr>
<tr>
<td>1.16</td>
<td>4.70</td>
<td>1.90</td>
<td></td>
</tr>
<tr>
<td>1.30</td>
<td>4.30</td>
<td>0.20</td>
<td></td>
</tr>
</tbody>
</table>

Assumptions for Calculations: (1) Household’s prime earner is a male born January 2, 1930 and alive at age 62. (2) Base year for simulations is 1992; all values are in $1992. (3) Wife is born January 2, 1932; wife has never worked; couple has no dependents. (4) Worker retires on 62nd birthday; wage history corresponds to economy-wide median earnings profile for worker’s cohort from ages 20 to 50 and is constant in real terms thereafter. (5) Wife claims as soon as possible (the later of her 62nd birthday or husband’s date of claim). (6) Household discount rate is 3%; mortality risks correspond to Social Security Administration’s sex- and cohort-specific survival tables.

Note: Results are similar if use other years of wife’s birth (1925, 1928, 1930, 1935).

Table 4.3: Mortality Risk and Discount Rate Required for Optimal Claiming at 62

averse individuals with an uncertain date of death.\textsuperscript{24} Individuals are able to purchase more of this real annuity by delaying, so delays are more attractive under risk aversion.

We present simulations from an expected utility maximization model with liquidity constraints to show how the inclusion of risk aversion affects the length of optimal delays and the gains from delay. While liquidity constraints are irrelevant in the financial calculations, since the timing of benefit receipt does not matter except through the discount rate, liquidity constraints are key here. In order to purchase more of the real annuity, the individual must delay the onset of the annuity stream. Assuming that the individual has no other income and cannot borrow against SS, he must consume from financial wealth during his delay.\textsuperscript{25} An individual with high wealth will delay longer, since he can better afford to consume out of wealth during the delay.

For the simulations, we restrict our attention to the case of a single individual. This is sufficient to illustrate the difference between this model and the financial calculations and avoids the computational burden of a full household optimization model. We use a CRRA specification

\textsuperscript{24}Crawford and Lillien (1981) model the incentive to work longer because of the increased value of a real annuity.

\textsuperscript{25}We assume that liquidity constraints are in the form of wealth non-negativity constraints. This seems reasonable for this cohort, especially since it is illegal to use SS as collateral for a loan.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No Bequest Motive, CRRA = 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>11</td>
<td>177,796</td>
<td>179,493</td>
<td>179,527</td>
<td>1.80</td>
</tr>
<tr>
<td>20</td>
<td>10</td>
<td>18</td>
<td>179,824</td>
<td>182,143</td>
<td>182,842</td>
<td>3.13</td>
</tr>
<tr>
<td>40</td>
<td>10</td>
<td>27</td>
<td>182,885</td>
<td>185,790</td>
<td>187,746</td>
<td>5.05</td>
</tr>
<tr>
<td>80</td>
<td>10</td>
<td>34</td>
<td>187,090</td>
<td>190,583</td>
<td>194,414</td>
<td>7.61</td>
</tr>
<tr>
<td>120</td>
<td>10</td>
<td>36</td>
<td>190,019</td>
<td>193,886</td>
<td>198,926</td>
<td>9.25</td>
</tr>
<tr>
<td>200</td>
<td>10</td>
<td>36</td>
<td>194,026</td>
<td>198,315</td>
<td>204,858</td>
<td>11.25</td>
</tr>
<tr>
<td>Bequest Motive, CRRA = 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>11</td>
<td>143,331</td>
<td>144,285</td>
<td>144,522</td>
<td>1.24</td>
</tr>
<tr>
<td>40</td>
<td>10</td>
<td>18</td>
<td>146,850</td>
<td>147,674</td>
<td>147,681</td>
<td>0.86</td>
</tr>
<tr>
<td>120</td>
<td>10</td>
<td>11</td>
<td>132,195</td>
<td>132,701</td>
<td>132,702</td>
<td>0.53</td>
</tr>
<tr>
<td>No Bequest Motive, CRRA = 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>10</td>
<td>32</td>
<td>203,859</td>
<td>208,198</td>
<td>212,665</td>
<td>9.14</td>
</tr>
</tbody>
</table>

Assumptions for Calculations: (1) Household’s prime earner is a male born January 2, 1930 and alive at age 62. (2) Base year for simulations is 1992; all values are in $1992. (3) Worker retires on 62nd birthday; wage history corresponds to economy-wide median earnings profile for his age cohort from ages 20 to 50 and is constant in real terms thereafter. (4) Household discount rate is 3%; mortality risks correspond to Social Security Administration's sex- and cohort-specific survival tables. (5) Bequest motive is in the form of a linear utility of bequests term. The linear parameter is set equal to 5.5 * 10^{-5}.

Note: Worker’s PIA is $963 in all cases.

Table 4.4: Expected Utility Maximization: Single Worker

of the instantaneous utility function of consumption. There are three new parameters: the utility discount rate, the coefficient of relative risk aversion and the initial wealth level. We assume that the utility discount rate is equal to the market interest rate of 3%. In the base case, we use log utility, corresponding to a CRRA of one, and financial wealth of $40,000. In some simulations, we introduce a linear utility of bequests term. The full model is presented in Appendix D.

Table 4.4 presents results using expected utility maximization. Columns (b) and (c) show optimal delays under financial calculation and expected utility maximization. The following three columns report wealth equivalents at zero delay, the financial optimal delay, and the

---

26The parameter on the linear utility of bequests term is 5.5 x 10^{-5}.
expected utility maximizing delay; the wealth equivalent is the amount of wealth an individual requires today to be made as well off as he is by his entitlement to the stream of SS benefits. The final column contains the change in wealth equivalent from choosing the expected utility maximizing delay rather than zero delay divided by the PIA.

We can compare Table 4.4 to the single worker base case from Table 4.1. Several pieces of evidence support our prediction that SS benefits are valued more highly under risk aversion. First, the optimal delay is longer using expected utility maximization than financial calculation for any wealth level. Second, for any given delay, the wealth equivalent is higher than the EPDV.27 Third, the increase in the wealth equivalent from choosing the expected utility maximizing delay rather than zero delay is much larger than the increase in EPDV from choosing the financial optimal delay rather than zero delay reported in Table 4.1.

The first six rows show the effect of varying the wealth level when there is no bequest motive. The expected utility maximizing delay increases monotonically with wealth: the optimal delay is 11 months when wealth is $10,000, 27 months when wealth is $40,000, and 36 months when wealth is $120,000. This is consistent with our explanation that delay is less costly for high wealth individuals. To underpin the theoretical result of an inverse u-shaped pattern, we add a linear bequest motive. Now we find the expected inverse u-shaped pattern: the optimal delay is 11 months when wealth is $10,000, 18 months when wealth is $40,000, and 11 months when wealth is $120,000. The final row shows that a larger CRRA leads to a longer delay and a larger wealth equivalent for any given delay.

To recap, in the financial calculations we estimate, we find that optimal delays are often lengthy and that the gain from delays are moderate for couples and small for single workers. The calculations also illustrate five cross-sectional predictions for empirical testing. The expected utility maximization simulations suggest that gains from delays may be ten times larger or more when we include risk aversion. In the following sections we examine the extent of delays in the data and test the cross-sectional predictions.

27Although the EPDV at zero delay and at the financial optimal delay are not shown, they are the same as those in Table 4.1 for the single worker base case. The EPDV and wealth equivalent are equal only in the case of a linear utility function with equal utility discount and interest rates.
4.4 Conclusion

While there is a large literature on take-up decisions for programs such as UI and AFDC, we are not aware of any previous analysis of SS claiming behavior as a take-up decision, despite the fact that SS dwarfs these programs in terms of annual expenditures and beneficiaries.\(^{28}\) Each year, roughly one million male and 750,000 female fully insured individuals reach age 62 and face choices of when to retire and claim SS benefits.\(^{29}\) In addition, the SS take-up decision differs from that for JI or AFDC because it is a dynamic decision; the question is not whether to take up benefits, but when.

In this paper, we use financial calculations and simulations of an expected utility maximization model to estimate optimal delays and the gains from delay. We find that delays are optimal in a wide variety of cases and that gains are often significant. In the financial calculations, gains from delay are around 600% of PIA for married couples in many of the simulations we perform, though they are much smaller for a single worker. The simulations of the expected utility maximization model for a single worker suggest that optimal delays are longer and that gains from delay may be ten or more times larger when risk aversion is incorporated.

Tables 4.1, 4.2, and 4.4 indicate that immediate claiming is almost never optimal, except in the cases of a much older wife, high mortality risk, or high discount rate. Table 3 shows that with average mortality risk, a couple must have a real discount rate of 5.3% or higher for zero delay to be optimal.

Our research has implications for the large literature on SS, in particular the estimation of retirement responses to SS and the computation of the distributional effects of the program. As we have discussed, the SS benefit level may be endogenous. Theoretically, claiming behavior appears to be influenced by factors such as health, wealth, wife’s age, and wife’s earnings which may also affect retirement propensities. Even more serious, claiming may be influenced by the choice of retirement date. To avoid this endogeneity problem, researchers studying the impact of benefits on retirement should use the PIA rather than the benefit level. Assuming that

\(^{28}\)For example, in 1992 there were 41.5 million SS beneficiaries (of which 25.8 million were retired worker beneficiaries), versus 9.6 million UI recipients and 4.8 million AFDC families (or 13.6 million AFDC recipients). In 1992, SS benefits paid were $280 billion, compared to UI benefits of $25.8 billion and AFDC benefits of $22.2 billion. All figures are from the 1993 Green Book, Committee on Ways and Means, U.S. House of Representatives.

\(^{29}\)Figures are unpublished data from the Office of the Chief Actuary, Social Security Administration.
individuals claim benefits as soon as they retire overstates the benefit of continued work, as part of the benefit is available to those who retire by delaying claiming.

Researchers studying the distributional effects of the program may want to estimate such effects using both the PIA and the actual benefit level; the former shows patterns of redistribution inherent in the system conditional on everyone claiming at age 65, and comparing the latter to the former shows to what extent these patterns are altered by claiming behavior. We feel these issues are particularly relevant now, as the release of the HRS will certainly lead to a new round of research on SS.
Appendix A

Other consumption and wealth profiles in the presence of a bequest motive

This appendix contains a rather brief discussion of the remaining scenarios of table 2.3 describing the consumption and wealth profiles in the presence of a linear bequest motive. Below, we present the assumptions giving us these different scenarios and present graphical illustrations of both consumption and wealth profiles over the considered life-cycle.

A.1 Case 3: Liquidity constraints binding for $t' \in \{E, \infty\}, g > \frac{p-r}{\alpha-1}$, budget constraint not binding

This case is probably the most complicated one as can easily be seen from a quick glance at figures A-1 and A-2. To be in this scenario, we need to have $g > \frac{p-r}{\alpha-1}$.\(^1\) Denote by $M$ the time when the income stream equals the fully optimal consumption path $C_i^{opt}$. $M$ is therefore a natural cutoff point for separating the problem into two independent subproblems. Indeed, given that we suppose that liquidity constraints hit sometime between time $L$ and infinity, we know that this moment has to be before $M$ as thereafter the fully optimal consumption profile is

\(^1\)Indeed, for $g \leq \frac{p-r}{\alpha-1}$, we have that if $W_{t^*} = 0$, then $W_t = 0 \forall t > t^*$. 

82
once again feasible without binding liquidity constraints, as income is bigger than consumption. Hence this scenario belongs to the first group of scenarios having $\lim_{t \to \infty} W_t e^{-rt} > 0$.

For $t < M$ we are basically doing a constrained optimization of

$$\max_{C_t} EU = \frac{1}{\alpha} \int_0^M C_t^{\alpha} e^{-(p+\delta) t} dt + b \int_0^M B(t) e^{-(p+\xi) t} dt$$  \hspace{1cm} (A.1)$$

under the binding budget constraint

$$\int_0^M C_t e^{-rt} dt = W_0 + \int_E^M Y_t e^{-rt} dt$$  \hspace{1cm} (A.2)$$

and wealth non-negativity constraints $W_t \geq 0$.

On intervals where the liquidity constraints are not binding, the first-order condition

$$C_t^{\alpha-1} = \left( C_t^{\alpha-1} e^{(\rho+p-r)t} - \frac{b}{p} (e^{(\rho+p-r)t} - e^{(\rho-r)t}) \right)$$  \hspace{1cm} (A.3)$$

determines the optimal consumption pattern. Given the pattern of $\frac{C_t}{C_0}$ illustrated in table 2.2, and given the implications for $C_0$ of having a binding budget constraint, we know that the slope and the level of the consumption profile lie strictly below the fully optimal consumption pattern 2.15. Now suppose that the first-order equation holds over the entire interval $[0, M]$. Then it is impossible for $C_M$ to reach the unconstrained optimal level $(\frac{p}{\delta})^\frac{1}{1-\delta} e^{\frac{p-r}{\delta-1} M}$. Given the functional form of the objective function, it is easy to see that at any optimum, we have a continuous consumption profile. Therefore the first-order condition cannot determine consumption on the entire interval $[0, M]$. We thus have a positive period of time denoted $[N, M]$, during which the liquidity constraints do strictly bind. For $t \geq M$, the fully optimal consumption pattern 2.15 is once again feasible and hence optimal.

Summarizing, we have three different regimes on the optimal path of the consumption process: First a period $[0, N]$ where a first-order condition determines the consumption levels, then for $[N, M]$ the consumption level is exactly equal to current income, and then lastly the interval $[M, \infty]$ where the fully optimal consumption levels are achievable. A graphical illustration of these findings can be found in figures A-1 and A-2 that show the consumption

83
and wealth pattern over time. Considering the case where \( g \equiv 0 \), we find that consumption is decreasing over time, whereas the wealth process follows a rather unusual pattern of a first period of decumulation, then zero wealth and then ultimately increasing wealth. Hence we find that even for life-cycle savers that are liquidity constrained over part of their life, it may still be optimal to have wealth accumulation over the last part of their life. This is obviously an interesting result, as it gets away from the finding of the pure life-cycle framework were we have that whenever liquidity constraints bind once, they will do so forever after. Hence, our model may shed some new light on some of the observed patterns of wealth accumulation of the elderly: People may find it optimal to consume less than their income because of some motive of giving.

Figure A-1: \( C_t \) Profile, Constraint binding at \( t > E \)

After this determination of the optimal consumption process and implicitly of the wealth profile, we now turn to the comparative statics for the consumption process as we vary the bequest parameter \( b \). It is easy to show that there is a point \( t^{**} \in [0, E] \) as well as times \( N' \) and \( M' \) both part of the interval \( ]N, M[ \) with \( N' < M' \) such that

\[
\begin{align*}
\frac{dC_t}{db} &< 0 \text{ for } [0, t^{**}] \\
\frac{dC_t}{db} &> 0 \text{ for } [t^{**}, N'] \\
\frac{dC_t}{db} &= 0 \text{ for } [N', M'] \\
\frac{dC_t}{db} &< 0 \text{ for } [M', \infty[ \\
\end{align*}
\]
A.2 Case 4: Liquidity constraints bind for $t = [E, t'[, for some $t' \in ]E, \infty[,$ budget constraint not binding

In this case we basically combine the effects of the previous scenarios. The implications for the annuity valuation are a straightforward extension of cases 2 and 3 as discussed in sections 2.3.2 and A.1.

A.3 Case 6: Liquidity constraints binding at $t = E,$ $\frac{\rho - r}{\alpha - 1} > g > \frac{\rho + p - r}{\alpha - 1}$

This scenario is the extension of scenario 5 to the case of a binding liquidity constraint at time $E.$ Graphically, we can easily see the effect of this additional constraint on the consumption and wealth profiles, with consumption tracking current income after $E$ and wealth staying at 0 after $E.$ (see figures A-3 and A-4) Using proposition 3, we can also easily derive the comparative statics results with respect to the impact of an increase in the bequest parameter $b.$

A.4 Case 8: Liquidity constraints binding at $t = E,$ $\frac{\rho + p - r}{\alpha - 1} > g$

Lastly, we turn to the scenario where the budget constraint is binding at infinity and where we also have a binding liquidity constraint at $t = E,$ which can be seen as an extension of scenario
Figure A-3: $C_t$ Profile, Constraint binding at $E$

Figure A-4: $W_t$ Profile, Constraint binding at $E$
Figure A-5: $C_t$ Profile, $\frac{\rho+p-r}{\alpha-1} > g$, Constraint binding at $E$

Figure A-6: $W_t$ Profile, $\frac{\rho+p-r}{\alpha-1} > g$, Constraint binding at $E$

7 that we discussed in section 2.3.2. In this scenario, the first-order condition 2.16 will still hold on the two subintervals $[0, E]$ and $[E, +\infty]$. Figures A-5 and A-6 present the time pattern of consumption and wealth.
Appendix B

Life-span uncertainty model of gifts and bequests

We outline one possible setup of a continuous time life-span uncertainty model for which the properties derived in section 3.3 still hold. We consider an individual who faces life-span uncertainty as the only form of uncertainty. As before, we allow for the presence of potentially binding borrowing constraints under the form of weak non-negativity constraints.

We assume that mortality risk is known as of the initial time \( t = 0 \) and that there is no updating about health outcomes as time passes. We summarize the mortality risk by the density function \( p_t \) (the distribution function is \( P_t = \int_0^t p_s \, ds, \ P_T = 1 \)) that is defined over the maximum life-span \( T \). We suppose that the hazard rate \( \frac{P_t}{1 - P_t} \) is increasing over the entire interval \( [0, T] \) and that the hazard rate attains 1 for \( t = T \).\(^1\) We define \( r, \rho \) and \( \xi \) as the continuous time equivalents of the previously discussed parameters. \( \{Y_t\} \) denotes the exogenous income stream of the individual.

Consider the following individual optimization problem:

\[
\max_{C_t, G_t} EU = \int_0^T u(C_t)(1 - P_t)e^{-\rho t}dt + b \int_0^T \left( G_t + B_t \frac{P_t}{1 - P_t} \right)(1 - P_t)e^{-\xi t}dt
\]

\(^1\)This pattern is consistent with real world life tables as those compiled by the U.S. Social Security Administration
subject to the set of constraints B.2 and B.3.

\[ B_T e^{-rT} + \int_0^T (C_t + G_t) e^{-rt} dt = W_0 + \int_0^T Y_t e^{-rt} dt \quad \text{(B.2)} \]

\[ B_t = W_t \equiv W_0 + \int_0^t Y_{\tau} e^{-r\tau} d\tau - \int_0^t (C_\tau + G_\tau) e^{-r\tau} d\tau \geq 0 \quad \forall t \in [0, T] \]
\[ G_t \geq 0 \quad \forall t \in [0, T] \quad \text{(B.3)} \]

It is easy to see that the results of section 3.3 generalize to the present setup.
Appendix C

EPDV formula

C.1 One-earner couple

In this appendix we provide the formulae for our computations of the expected present discounted value (EPDV) of future benefits for the case of a one-earner household.

Notation:

- $AM$ denotes the age of a the worker.

- $AF$ denotes the age of a potential spouse at the time the worker is aged $AM$.

- $YM$ denotes the year of birth of the worker.

- $YF$ denotes the year of birth of the partner.

- $Maxage$ represents the maximum potential age that we consider for both the worker and the spouse.

- $\rho$ denotes the real discount factor used.

- $t$ denotes the number of months after attaining the age of $AM$, that the worker decides to wait before first claiming social security benefits.

- $s$ denotes the number of months that a partner, older than 62 years but younger than normal retirement age, decides to wait till starting to claim benefits. We restrict our
attention to delays in claiming spousal benefits up to the normal retirement age, as the social security law does not provide any financial incentives to wait beyond this age.

- Similarly define \( \{s_k\}, \forall k = 1, \ldots, 12 * (Maxage - AM) \) as being the number of months that a widow, older than 60 but younger than her normal retirement age, decides to wait before claiming her survivor benefits, conditional on her partner having died \( k \) months after attaining age \( AM \). Once again, there is no financial incentive to delay claiming of survivor benefits beyond the normal retirement age.

- \( BEN_x \) represents the amount of benefits the worker is entitled to in the month \( x \) after retirement, expressed in dollars of the time when the worker retired at age \( AM \).

- \( DEP_x \) is the amount of benefit the spouse is entitled to claim in month \( x \) after the worker's retirement date, expressed in dollars of the time when then worker was aged \( AM \).

- \( SU1_{x,y} \) is the amount of benefits the surviving spouse is entitled to claim in month \( x \) after the worker's retirement date, in case the worker dies in month \( y \) after his retirement before first claiming benefits. (expressed in dollars of the time when then worker was aged \( AM \))

- \( SU2_{x,y} \) is the amount of benefits the surviving spouse is entitled to claim in month \( x \) after the worker's retirement date, in case the worker dies in month \( y \) after his retirement, which is after first claiming benefits. (expressed in dollars of the time when then worker was aged \( AM \))

- \( \Theta_1^1 \) is a dummy variable which is 1 if \( 12 * AF + i \geq 12 * 62 + s \), and which is 0 otherwise.

- \( \Theta_{i,k}^2 \) is a dummy variable that exists for every \( k \) and which is 1 if \( 12 * AF + i \geq 12 * 60 + s_k \), and which is 0 otherwise.

- \( p_x(.\mid YM, \text{sex}) \) is the cohort and sex specific conditional probability measure expressing the probability that the worker is still alive in month \( x \) after the his retirement date, conditional on that the worker was alive in month \( x - 1 \). By definition, \( p_0 = 1 \).

- \( q_x(.\mid YF, \text{sex}) \) is the cohort and sex specific conditional probability measure expressing the probability that the partner is still alive in month \( x \) after the worker's retirement date,
conditional on that the partner was alive in month \( x - 1 \). By definition, \( q_0 = 1 \).

- \( i, j \) are simple counting variables.

For the purpose of the present paper, we assume that \( s = 0 \) and that \( s_k = 0, \forall k \), i.e. that the partner claims as soon as possible. Hence, the benefit amounts \( BEN, DEP, SU1 \) and \( SU2 \) are only a function of the retirement age and the claiming delay \( t \) of the worker.

The concept of interest is the expected present discounted value of Social Security benefits (EPDV) for the household computed at the time the worker is aged \( AM \) conditional on delaying claiming by \( t \) months.

The EPDV is composed of three components:

\[
PB \equiv \sum_{i=t}^{12\text{*(Maxage-AM)}} \{ BEN_i \ast (\prod_{j=0}^{i}(p_j(.|YM, sex))) \ast (1 + \rho)^{-i}\}
\]

\[
SpB \equiv \sum_{i=t}^{12\text{*(Maxage-AF)}} \{ DEP_i \ast \Theta_i^{1} \ast (\prod_{j=0}^{i}(q_j(.|YF, sex) \ast p_j(.|YM, sex))) \ast (1 + \rho)^{-i}\}
\]

\[
SuB \equiv \sum_{k=0}^{t-1} \left\{ \sum_{i=k}^{12\text{*(Maxage-AF)}} \{ SU1_{i,k} \ast \Theta_{i,k}^{2} \ast (\prod_{j=0}^{k-1}(p_j(.|YM, sex))) \ast (1 - p_{12\text{*(Maxage-AF)+j}}(.|YM, sex))) \ast (1 + \rho)^{-i}\} \right\}
\]

\[
+ \sum_{k=t}^{12\text{*(Maxage-AM)}} \sum_{i=k}^{12\text{*(Maxage-AF)}} \{ SU2_{i,k} \ast \Theta_{i,k}^{2} \ast (\prod_{j=0}^{k-1}(p_j(.|YM, sex))) \ast (1 - p_k(.|YM, sex)) \ast (\prod_{j=0}^{i}(q_j(.|YF, sex))) \ast (1 + \rho)^{-i}\}
\]
The expected present discounted value of social security benefits for a one-earner household is thus:

\[ EPDV = PB + SpB + SuB \]

C.2 Two-earner couple

The extension of the previous \( EPDV \) calculations to the case of a two earner couple is straightforward. In our simulations, we assume that the spouse retires as soon as possible after the primary worker turns \( AM \). Conceptually the various components of the \( EPDV \) can easily be generalized to the case of a two-earner couple.

Notation:

\( PB \) continues to denote the expected present discounted value of the prime earner's benefit.

\( SpB \) continues to denote the expected present discounted value of spousal benefits, conditional on the spouse claiming benefits as a dependent of the prime earner.

\( SuB \) continues to denote the expected present discounted value of survivor benefits, conditional on the surviving spouse claiming benefits as a dependent of the prime earner.

\( SB \) is defined as the expected present discounted value of benefits the spouse (secondary earner) can claim on her own earnings record. The definition of \( SB \) parallels that of \( PB \).

\( AB \) is defined as the expected present discounted value of the spousal and survivor benefits the secondary earner is entitled to claim based on the prime earner's earnings history in addition to the benefits the secondary earner claims on her own earnings record.

\( SSC \) denotes the present discounted value of social security contributions in the case that the secondary earner continues to work after the date when the prime earner turns \( AM \).

The concept of interest is as before the \( EPDV \) for the household. It is composed as follows

\[ EPDV = PB - SSC + \max[SpB + SuB, SB + AB] \]
Appendix D

Expected utility maximization model

In this appendix we provide an outline of the structure of the expected utility maximization model underlying our simulations for a single worker.

- We use a utility function that is quasi-linear in consumption and bequests.

- In our optimization, we use years as the base unit for survival, consumption and wealth periods. For claiming behavior, we allow for monthly claiming delays. These assumptions are made for pure reasons of tractability.

- Social Security benefits in period $i$ conditional on claiming $t$ months after the age of 62 are denoted $B_{i,t}$.

- Consumption in year $i$ conditional on the income path attached to a $t$ month delay on claiming is denoted $C_{i,t}$.

- Wealth in period $i$ is similarly denoted $W_{i,t}$.

- The survival probability of living till period $i$ conditional on having lived till $i - 1$ is denoted $p_i$, with $p_{62} = 1$ and $p_{Maxage+1} = 0$.

- The cumulative survival probability till period $i$ is $P_i = \Pi_{j=0}^{i} p_i$, with $P_{Maxage+1} = 0$. 
The subjective rate of time preference is $\rho$ and the real interest rate is $r$.

We proceed in two steps.

1. First, for any given claiming delay of $t$ months, we write the optimization problem as

$$\max_{C_{i,t}} EU_t \equiv \sum_{i=62}^{Maxage} \left( \frac{P_i}{(1+\rho)^{i-62}}u(C_{i,t}) + b \frac{(1-p_{i+1})P_i}{(1+r)^{i-62}} W_{i+1,t} \right)$$

We maximize the above objective function subject to several constraints

$$W_{i+1,t} = (1+r)W_{i,t} + B_{i,t} - C_{i,t}$$

$$W_{62,t} + \sum_{i=62}^{Maxage} \frac{B_{i,t}}{(1+r)^{i-62}} = \sum_{i=62}^{Maxage} \frac{C_{i,t}}{(1+r)^{i-62}} + \frac{W_{Maxage+1,t}}{(1+r)^{Maxage-61}}$$

$$W_{i,t} \geq 0$$

and find an optimal level of expected utility $EU_t^*$ conditional on $t$.

2. In a second stage, we then maximize $EU^*$ over $t$

$$\max_t EU_t^*$$

and find the optimal claiming delay $t$. 

95
Bibliography


