ESSAYS ON OPTIMAL TAXATION AND THE LIFE CYCLE

A dissertation presented by

SEAMUS JOHN SMYTH

to

The Department of Economics in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the subject of Economics.

Harvard University
Cambridge, Massachusetts
May, 2006
ESSAYS ON OPTIMAL TAXATION AND THE LIFE CYCLE

Abstract

Chapter 1 studies optimal taxation of capital and labor in general equilibrium using a calibrated overlapping generations (OLG) economy. Allowing for very general non-linear separate schedules of capital and labor income taxes, I quantitatively solve for the optimal taxes. The results show that capital and labor should be taxed quite differently. While all labor is taxed at close to a 12 percent flat rate, relative to marginal rates of over 30 percent today, most capital income is exempt from taxation. The optimal marginal tax rates on capital holdings above $100,000 rises quickly from 0 to over 20 percent. I then identify the exact economic forces driving these results.

Chapter 2 shows that, allowing for permanent and transitory income shocks, mean VSLs follow an inverted-U shape over the life cycle for most of the population. Substantial heterogeneity, however, exists across agents with different productivities and assets. At age 45, the mean VSL is one-third larger than the median, and the 95th percentile VSL is about 12 times the 5th percentile VSL. The model provides a unified theoretical justification for a variety of recent empirical findings: the VSL-income elasticity, the black-white VSL gap and the male-female VSL gap.
Chapter 3 solves for optimal non-linear taxes on consumption in an OLG model with idiosyncratic risk. Consumption taxes, as well as income and labor taxes, can be progressive. I find the optimal progressive consumption tax is very redistributive. Marginal rates of taxation on additional consumption rise until over $200,000 and reach a maximum of almost 200% - it takes $3 to buy $1 extra of consumption. These high marginal rates on consumption greatly reduce the cross sectional variance of consumption and transfer consumption across the life cycle from the middle aged and old to the young.
# Contents

Abstract .................................................. iii  
Table of Contents ........................................ v  
List of Figures ........................................... ix  
List of Tables ............................................ xi  
Acknowledgements ........................................ xii

1 A Balancing Act: Optimal Nonlinear Taxation in Overlapping Generations Models 1
  1.1 Introduction ........................................... 2  
  1.2 Overlapping Generations Model .......................... 6  
      1.2.1 Household ......................................... 7  
      1.2.2 Production ........................................ 13  
      1.2.3 Government ........................................ 14  
          1.2.3.1 Form of Tax Function ....................... 15  
          1.2.3.2 Government Budget Constraints .......... 16  
      1.2.4 Social Planners Objective ....................... 18  
      1.2.5 Equilibrium Definition and Computation ........ 19  
  1.3 Parametrization ....................................... 19  
  1.4 Model Properties ..................................... 23
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.4.1</td>
<td>Aggregate Statistics</td>
<td>24</td>
</tr>
<tr>
<td>1.4.2</td>
<td>Life Cycle Patterns of Consumption and Wealth</td>
<td>25</td>
</tr>
<tr>
<td>1.4.3</td>
<td>Variances of Consumption and Wealth</td>
<td>27</td>
</tr>
<tr>
<td>1.5</td>
<td>Steady State Optimality Results</td>
<td>29</td>
</tr>
<tr>
<td>1.5.1</td>
<td>Differences from Initial System</td>
<td>32</td>
</tr>
<tr>
<td>1.5.2</td>
<td>Life Cycle Tax Burden Distribution</td>
<td>32</td>
</tr>
<tr>
<td>1.5.3</td>
<td>Increase in Average Consumption</td>
<td>34</td>
</tr>
<tr>
<td>1.5.4</td>
<td>Labor Tax Balancing</td>
<td>35</td>
</tr>
<tr>
<td>1.5.5</td>
<td>Capital Tax Balancing</td>
<td>40</td>
</tr>
<tr>
<td>1.5.6</td>
<td>Discussion of Optimum</td>
<td>41</td>
</tr>
<tr>
<td>1.6</td>
<td>Adding a Lump Sum Levy</td>
<td>43</td>
</tr>
<tr>
<td>1.6.1</td>
<td>Adding a Lump Sum Levy</td>
<td>44</td>
</tr>
<tr>
<td>1.6.2</td>
<td>Changes in Tax Functions Induced by Lump Sum Component</td>
<td>45</td>
</tr>
<tr>
<td>1.6.3</td>
<td>Impacts Over Life Cycle</td>
<td>47</td>
</tr>
<tr>
<td>1.7</td>
<td>Transition</td>
<td>50</td>
</tr>
<tr>
<td>1.7.1</td>
<td>Evolution of Aggregates</td>
<td>51</td>
</tr>
<tr>
<td>1.7.2</td>
<td>Welfare gains</td>
<td>53</td>
</tr>
<tr>
<td>1.8</td>
<td>Conclusion</td>
<td>56</td>
</tr>
</tbody>
</table>

2 Value of a Statistical Life Over the Life Cycle

2.1 Introduction                                         60
2.2 Model                                                63
2.2.1 Household                                          64
2.2.2 Wage Process                                       66
2.2.3 Recursive Representation                           67
2.2.4 Defining the Value of a Statistical Life            68
2.3 Calibration .................................................. 70
   2.3.1 Agent’s Utility ........................................ 70
   2.3.2 Life Cycle Pattern and Variance of Wages .......... 73
   2.3.3 Survival Probabilities ................................ 75
   2.3.4 Social Security Benefits Function .................. 75
2.4 Base Case .................................................. 75
   2.4.1 Entire Population ..................................... 77
   2.4.2 Without Income Shocks ............................... 82
2.5 Individual Characteristics ................................. 84
   2.5.1 Income ................................................ 84
   2.5.2 Race .................................................. 90
   2.5.3 Gender ............................................... 93
2.6 Conclusion .................................................. 95

3 Optimal Consumption Taxes ................................. 99
   3.1 Introduction ............................................ 100
   3.2 Model and Calibration ................................... 102
      3.2.1 Household .......................................... 102
      3.2.2 Firm ............................................... 104
      3.2.3 Government ........................................ 105
      3.2.4 Steady State of Model ............................. 106
      3.2.5 Social Welfare .................................... 107
   3.3 Optimal Consumption Taxes ............................. 107
      3.3.1 Form of Optimal Consumption Tax ................. 108
      3.3.2 Life Cycle Effects ................................ 110
      3.3.3 Macroeconomic Impacts ............................ 114
3.4 Efficiency and Redistribution Compared to Linear Consumption Taxes 117
  3.4.1 Linear Consumption Tax ............................................. 117
3.5 Conclusion ................................................................. 121

Bibliography ................................................................. 123

A Appendices to Chapter 1 .................................................. 130
  A.1 Solving the Life Cycle Problem ...................................... 130
  A.2 Steady State Solution .................................................. 132
    A.2.1 Steady State Equilibrium Definition .......................... 132
    A.2.2 Steady State Numerical Solution .............................. 133
  A.3 Transition Path Solution ............................................. 134
    A.3.1 Transition Path Equilibrium Definition ....................... 134
    A.3.2 Transition Path Numerical Solution ........................... 135
  A.4 Estimation of Wage Process ......................................... 136

B Appendices to Chapter 2 .................................................. 137
  B.1 Solving the Life-Cycle Problem .................................... 137

C Appendices to Chapter 3 .................................................. 140
  C.1 Estimation of Wage Process .......................................... 140
  C.2 Solving the Life Cycle Problem .................................... 141
  C.3 Steady State Solution ................................................ 143
    C.3.1 Steady State Equilibrium Definition ........................ 143
    C.3.2 Steady State Numerical Solution .............................. 144

viii
# List of Figures

1.1 Various Tax Functions ............................................ 17
1.2 Approximation of U.S. Statutory Tax Code ...................... 22
1.3 Social Security Benefits Function. ............................... 23
1.4 Average Consumption and Wealth over the Life Cycle ............ 26
1.5 Variances of Consumption and Labor Income Over the Life Cycle . 28
1.6 Optimal Labor and Capital Taxes .................................. 31
1.7 Life Cycle Pattern of Taxes Paid Under Optimal Tax System .... 33
1.8 Average Consumption Profiles Under Current and Optimal Tax Systems 36
1.9 Labor Supply Elasticity of Agent ................................... 38
1.10 Marginal Value of Assets .......................................... 41
1.11 Optimal Taxation when the Government has Access to Capital, Labor and Lump Sum Taxes .......................... 46
1.12 Optimal Taxation with Capital, Labor and Lump Sum Taxes .... 48
1.13 Evolution of Aggregate Variables Along Transition Path ........ 52
1.14 Percent of Initially Alive with Welfare Gains .................... 54

2.1 VSL Over Life Cycle .............................................. 76
2.2 Mean, Median, 5th and 95th Percentiles of VSL Over Life Cycle ... 79
2.3 Distribution of VSL at Age Forty ................................. 80
2.4 Variance of VSL Over Life Cycle ............................................. 81
2.5 Impact of Income Shocks on Mean and Median VSL .................. 83
2.6 VSL by Race ........................................................................... 90
2.7 Breakdown of Sources of Differences in VSL between Blacks and Whites 91
2.8 VSL by Sex ............................................................................. 93
2.9 Breakdown of Sources of Differences in VSL between Males and Females 94

3.1 Optimal Tax on Consumption .................................................. 108
3.2 Mean Consumption ................................................................. 110
3.3 Variance of Consumption ......................................................... 111
3.4 Life Cycle Pattern of Mean Taxes Paid ..................................... 112
3.5 Life Cycle Patterns - Linear Consumption Tax Compared to Linear Consumption Tax ......................................................... 118
## List of Tables

1.1 Summary of Parameter Values, Sources and Main Economic Interpretation .................................................. 21  
1.2 Steady State Aggregate Quantities of the Model Parametrized to Current U.S. Data ........................................... 24  
1.3 Comparison of Initial and Optimal Steady States ............................................................................................. 30  
1.4 Comparison of Optimal Steady States Under Different Tax Regimes ................................................................. 50  

2.1 Summary of Parameter Values, Sources and Main Economic Interpretation .................................................. 72  
2.2 Permanent and Transitory Variances of Wage Shocks ....................................................................................... 74  
2.3 Lifetime Value and Value of Statistical Life ......................................................................................................... 84  
2.4 Elasticity of VSL ............................................................................................................................................... 85  
2.5 Elasticity of VSL at Different Ages ................................................................................................................... 87  

3.1 Summary of Parameter Values, Sources and Main Economic Interpretation .................................................. 106  
3.2 Comparison of Initial and Optimal Steady States ............................................................................................. 115
Acknowledgements

I would like to thank my advisers Aleh Tsyvinski, Jeffrey Liebman and David Cutler. Their guidance and comments have improved this work greatly over where it would be without them. Aleh particularly helped me to shape my ideas into a coherent paper, for which I am grateful.

I owe an immense debt of gratitude to my family: my brother Alex and my parents Jane and David Smyth. From a young age my parents worked, and it sometimes was hard work, to instill a love of learning and intellectual curiosity in me. My mother was especially successful at doing this, despite any feelings to the contrary I might have had at the time. My father also was instrumental, passing on his love of economics to me without pushing me or forcing me. Without them, and their huge investments in my human capital, I would not be where I am today. I thank them all.

Lastly, but most certainly not leastly, I would like to thank my fiancee Monica. Her love, support and patience through the, rather longer than hoped for, process helped inspire and motivate me. I am extremely grateful, and lucky, she thought “I had it in me” and decided to wait me out.
Chapter 1

A Balancing Act: Optimal Nonlinear Taxation in Overlapping Generations Models
1.1 Introduction

A classical question in economics is how to structure the tax code. How should capital be taxed? How should labor be taxed? How do we trade off between taxing the two? Should taxes apply to total income or separately to capital and labor income? How does changing the shape of one set of taxes impact the optimal shape of the other? The recent presidential commission on fundamental tax reform shows the continuing significance and salience of these questions.

This paper examines the optimal mix between non-linear labor and non-linear capital taxes in a model with incomplete markets, idiosyncratic wage risk, a leisure-labor trade-off and age specific mortality. Solving a model with all these features necessitates a quantitative approach. Separately looking at capital and labor taxes, a treatment novel to the literature, allows us to differentiate between their effects. By using flexible functional forms for taxation, we can investigate a wide range of possible taxes. The combination of the above allows us to simultaneously examine both the efficiency and distributional consequences of the tax system. In doing so we gain important insights about the trade-offs between protection and distortion in raising revenue through taxation.

To assess risk sharing, labor supply and capital formation effects our model contains several important features. First, the agents face uninsurable idiosyncratic wage risk and age specific mortality over the life cycle. This departure from the complete markets framework has important implications for the tax code as the insurance aspect of taxes now becomes important. Second, an overlapping generations framework with age-specific pattern of productivity and a life cycle for agents allows modeling inter-generational trade-offs. Third, the general equilibrium nature of the model links the aggregate statistics in the economy to the individuals’ decisions so
changes in the tax code feed through to economy-wide quantities such as wages, interest rates and the capital stock. Finally, using both a labor-leisure choice and a wage process that generates realistic income, wealth and consumption life cycle variance patterns allows examination of the distributional aspects of the current system and how these would change in response to shifts in the tax system.

The main contribution of this paper is the determination of the exact form and magnitude of optimal non-linear taxes on capital and labor. A key result is that both the level and shape of the tax on capital significantly differ from the level and shape of the tax on labor. Labor is taxed at a flat rate of around 12 percent - a stark contrast to the current system which has marginal rates that start low and increase to over 30 percent. Meanwhile, a large deduction for capital income exempts most agents from paying any taxes on their asset income. Above a deduction of about $100,000, however, capital taxes quickly rise to marginal rates of over 20 percent.

The second important contribution is that we determine the precise economic forces shaping these non-linear taxes. The labor tax results from the social planner balancing her conflicting desires. On one hand, she wants to take money from agents around retirement, who have low marginal values of assets, and give it to the young, who have high marginal values. On the other hand, she wishes to tax inelastically supplied labor. The young have very low labor supply elasticity which rises as the agents age becoming quite high as they approach retirement. So the social planner chooses the optimal labor tax by weighing these two effects against each other resulting in a close to linear tax on labor income.

Choosing the optimal capital tax also entails balancing distortions and benefits. Older, richer agents hold most of the capital. The planner would like to transfer some of this to poorer and younger agents. The familiar trade-off between taxing capital income and the aggregate level of capital, however, arises. High taxes on capital
push down the amount of capital in the economy. Reducing the amount of capital pushes down the capital-labor ratio and lowers wages. These lower wages hurt the young and poor who rely on wage income. Leaving most capital untaxed thus helps raise average wages. Some capital can still be taxed for redistributional purposes without severe adverse effects on incentives. Near retirement agents save mainly to prevent a consumption drop in retirement. Given why they save, their asset holdings are not very responsive to higher marginal tax. Since they are relatively insensitive to taxation, taxing their capital raises aggregate welfare. Thus, even without age-specific taxation, non-linear taxation allows differentially taxing agents of different ages.

Finally, the third important economic contribution is that we show there are aggregate economic gains from switching from the status quo to the optimal system. Fully accounting for the transition, a sudden unanticipated switch to the optimal tax system increases aggregate welfare. The gain is equivalent to making transfers to agents alive at the transition that sum to 10 percent. However, a majority of agents alive would oppose the the shift. The opposition is especially concentrated among the old.

Two distinct strands of the literature focus on issues of taxation related to our question. One strand of the literature primarily concerns itself with the efficiency and macroeconomic effects of taxation. Dating back to Ramsey (1927), this strand has examined the trade-offs between capital and labor taxation. The theme of capital taxes inhibiting the formation and capital, resulting in a lower capital stock, capital-labor ratio and hence lower wages flows through this literature. The classic result, due to Chamley (1986) and Judd (1985), finds that the optimal capital tax features extremely high rates for a short time followed by a zero tax on capital thereafter. Auerbach and Kotlikoff (1987) extend the analysis of taxation to realistic overlapping
generations (OLG) economies while Aiyagari (1995) shows how idiosyncratic risk and borrowing constraints imply a positive capital tax.

The other strand focuses on the distributional aspects of progressive labor income taxation. Following Mirrlees (1971), these papers examine how much a planner with a redistributive goal can actually shift the income distribution in the presence of informational asymmetries. More recent examples in this strain have further examined the static properties of this model. For instance, Saez (2001) attains additional analytic characterizations of the optimum tax system in terms of elasticities as well as providing numerical simulations of the optimum tax code given the U.S. income distribution. (2003) extend the analysis of Mirrlees to a dynamic setting. In doing so, they show that the intertemporal margin should be distorted implying a positive capital tax.

Recently some authors have begun quantitatively exploring the trade-offs where both efficiency and equity are involved. Domeij and Heathcote (2004) look at a Ramsey type model, but following Aiyagari (1995) and Hugget (1993), include idiosyncratic risk in their model of infinitely lived agents and linear taxes. Within their model, the capital tax acts as a risk-sharing mechanism as agents who receive positive income shocks save more and thus pay more in capital taxes. Reducing the capital tax leads to welfare losses from the reduction in risk sharing stemming from the lower capital tax. Nishiyama and Smetters (2005) build an OLG model with idiosyncratic risk and a redistribution authority to examine the efficiency gains of a move from our current system towards taxation of consumption when agents are compensated for any gains or losses. Again, moving away from taxation of capital causes losses from the reduction in insurance provided by the tax system. Conesa and Krueger (2005) return to looking for the best possible tax system, having added the presence of idiosyncratic risk, but concern themselves only with income taxes.
Their work does not examine the trade-offs in choosing between taxation of capital and labor.

Recent theoretical work on optimal taxation in OLG economies relates to our results. Erosa and Gerivas (2002) and Garriga (2003) study linear taxation of capital and labor in overlapping generations models without individual level heterogeneity. Age-dependant taxes are generally optimal. This result stems from optimally chosen leisure not being constant over the life cycle. In the absence of age dependent taxes, a tax on capital partially proxies for age-dependent taxes by differentially affecting the trade-offs at different ages. Thus a non-zero tax on capital improves welfare. This same effect helps explain our results on the optimal capital tax. The non-linear capital tax gives the planner an instrument to differentially tax the old and the young. Since the old have substantial asset holdings while the young do not, the substantial exemption on capital means that only the old pay the capital tax.

The remainder of the paper takes the following form: Section 1.2 presents the details of the quantitative model while Section 1.3 contains the details of parametrization. Section 1.4 discusses the properties of the model and compares them to the results in the literature. The heart of the paper is in Sections 1.5 and 1.6 which discuss the tax system that maximizes steady state welfare and Section 1.7 on the transition to the optimal system. Finally, Section 1.8 concludes.

1.2 Overlapping Generations Model

The model is designed to capture some features of the U.S. economy that are very important for assessing the welfare effects of different tax policies. Modeling detail focuses primarily on the household. Households are heterogeneous and make both savings and labor supply decisions. A realistic, idiosyncratic, uninsurable process
generates household wages. Importantly, there are both non-linear taxes levied on the agent’s income from various sources and a realistic Social Security system. A perfectly competitive representative firm and a government levying non-linear taxes and running a balanced budget close the model.

### 1.2.1 Household

Households optimally choose their consumption and leisure to maximize the discounted sum of future utility:

$$\max_{c, l} E \sum_{t=0}^{T} \beta^t \phi_t u(c, l)$$

Agents discount the future at a constant, time-invariant rate $\beta$. Their mortality differs by age with $\phi_t$ being the age-specific mortality. Especially for older agents, this age-dependent mortality, which rises as agents get older, shortens their effective time horizon. Older agents behave as if they have substantially higher discount factors because of the risk of not surviving to the next period.

Agents face incomplete markets in that no private markets provide insurance against future wage shocks. Agents can, however, self-insure through savings. All saving is risk-less and pays the market rate of return. This ability to partially self-insure, as well as conduct life cycle savings, provides a very important means of smoothing consumption in the face of the idiosyncratic shocks. There are two uses of an agent’s resources: consumption and saving for the future. The agent’s resources consist of prior savings, $a_{it}$, interest on those savings, $R_{it} a_{it}$ and work in the market. The agent allocates one unit of time between leisure and market work.\(^1\) Work in the market brings in additional resources of $w_{it} (1 - l_{it})$. Subtracted from these resources

\(^1\)Periods are one year.
are taxes levied by the government.

\[ c_t + a_{t+1} = a_t (1 + R_t) + w_t (1 - l_t) \]

\[ -\tau_k (R_t a_t) - \tau_l (w_t (1 - l_t)) - \tau_{ss} (w_t (1 - l_t)) \]

\[ -\tau_i (R_t a_t + w_t (1 - l_t)) \]

(1.2)

All of the \( \tau(\cdot) \) represent non-linear tax functions of the quantity being taxed. The main results focus on when the agent pays taxes, \( \tau_l \) and \( \tau_k \), on their labor earnings, \( w_t (1 - l_t) \), and capital earnings, \( R_t a_t \), respectively. The other possible taxable quantity is total income, denoted by \( \tau_i (R_t a_t + w_t (1 - l_t)) \). Generally not all forms of taxation are active at the same time. In addition, the Social Security tax is always active.

The base case is a borrowing constraint of zero. A borrowing constraint of zero ensures that no agents die in debt:²

\[ a_{t+1} > a_{min} \]

\[ a_{min} = 0 \]

(1.3)

We also consider cases where agents can borrow. In these cases the age-dependent borrowing constraint starts fairly loose and rises to zero as the agent ages. The tightening borrowing constraint prevents agents from dying in debt with certainty.³

²If agents can borrow some agents who die young will owe money. These agents’ debts must be dealt with somehow. Setting a borrowing constraint of zero ensures that this case never happens. In cases where agents borrow, the government covers the debts of agents who die owing money. This can come either from general revenues or from revenues of taxation of unintended bequests.

³If agents were allowed to borrow at older ages they would have incentives to borrow, consume what they borrowed, and die in debt with certainty. Making agents pay back all debts before retirement ensures that this does not occur.
The base model does not include a bequest motive. The size and importance of 
the bequest motive generates significant debate in the literature. In addition, fully 
specifying it requires including inter-generational linkages. Modelling these would 
greatly increase the complexity of the problem.\(^4\)

Mean wages, the life cycle pattern of wages and the idiosyncratic dispersion of 
wages over the agent's lifetime all play key roles in the model. To capture this, we 
use a common wage process that has both a permanent and transitory component. 
A similar process has been frequently estimated from microeconomic data. This 
process has been successful in matching several important empirical features of the 
U.S. economy. Storesletten, Telmer, and Yaron (2004) show that a general equilib-
rium economy, broadly similar to the one in this paper, with an income (not wage) 
generating process very similar to this one generates consumption inequality profiles 
that match the stylized facts about the data.

This wage process requires tracking both the permanent component of wages and 
a transitory component. Together, these two determine the agent's realized wage. 
Denoting the permanent component by \( P_t \), we write:

\[
P_{it} = P_{i,t-1} \left( \frac{M_t}{M_{t-1}} \right) \eta_{it} 
\]

\( M_t \) is the mean wage in period \( t \) while \( M_{t-1} \) was the mean wage the previous period. 
Their ratio, \( \frac{M_t}{M_{t-1}} \), gives the deterministic element of wage growth. The permanent 
shock, \( \eta_{it} \), persists into future periods through \( P_{it} \) depending on the previous per-
manent component of wages, \( P_{i,t-1} \).

\(^4\)DeNardi (2004) builds a model with inter-generational linkages and bequests. She focuses on the 
linkages between generations and for computational tractability studies a reduced number of gen-
erations. Our focus is instead on the tax system so we abstract from the inter-generational linkages.
The actual wage received by an agent is:

\[ w_{it} = w_t^p P_{it} \varepsilon_{it} \]  \hspace{1cm} (1.5)

This is composed of the permanent component discussed above, \( P_{it} \), plus two others components. The purely transitory shock, \( \varepsilon_{it} \), represents a shock that has no persistence. Its effects, except as they persist through differential asset holdings, disappear after a period. Finally there is an economy-wide portion of wages, \( w_t^e \). This has nothing to do with the agent and is determined by the economy-wide marginal product of labor. Since it affects the agent’s decisions, this component of the received wage must be incorporated.

This wage process is more realistic and captures the actual dispersion in wages better than those used in existing models that have looked at taxation in OLG economies. These use a several-state Markov process that, while it captures some important elements of idiosyncratic wage risk, misses some important features. Our wage process allows a higher dispersion of wages which better matches the observed distribution. For taxation issues, the large dispersion matters since those at higher incomes pay a large portion of taxes. Having these higher-income individuals included in the model is key to understanding the effects of tax reform. The greater dispersion means that more individuals have wages that are much higher than the economy average than previous work that considers only the Markov chain process.

So, in a hypothetical case where labor supply did not respond to taxation at all, the

5We do not allow the possibility of different agents having different mean wage paths or growth rates of income. The consensus in the literature since Macurty (1982) has been that there no evidence of heterogeneity in growth rates across individuals exists. This has been challenged recently by Guvenen (2005) who finds evidence for heterogeneity in the growth rates of income across individuals. This heterogeneity can lead to different consumption and savings behavior which would affect the optimal government policies.
wider dispersion of wages would give a larger effect of tax reform as there are more agents making high incomes. For those high incomes a change in tax rates has a larger revenue impact.

All agents retire at 65. Though they can choose to not work before age 65, if they do so they receive no Social Security benefits until they reach 65. The Social Security benefits the agent does receive are based on a progressive non-linear transformation of the final permanent wage similar to the translation between wages and benefits the actual system makes.

The progressive nature and the minimum benefit of the Social Security system play large roles. Both of these features are important to the results and present in the actual system. Social Security features Supplementary Security Income (SSI). This provides a floor level of consumption for very poor individuals. The maximum payments from SSI, which corresponds to the floor level, are about $7000 a year. Social Security also replaces lifetime income progressively. The declining replacement rates with respect to AIME and the cutoffs of AIME at the Social Security taxable maximum both imply a progressive transformation from income to benefits. The modeled Social Security system, with a floor substantially above zero and a progressive transformation of benefits, reflects important features of the actual program.

__6__Allowing an option of retiring early and claiming Social Security benefits could have interesting implications. Labor supply among the credit-constrained old poor agents with low wages would be reduced as they dropped out of the labor force. Higher marginal tax rates could then have an even greater effect in reducing labor supply among the old. This would require higher Social Security tax rates, with their attendant distortions. Though interesting, a detailed investigation of these effects lies beyond the scope of the current paper.

__7__Social security wages are based on final permanent wage rather than some form of AIME (Average Indexed Monthly Earnings). Having them based on an AIME-based formula would add an additional state variable to the agents problem. The increase in realism of such an addition was judged to not be worth the very large increase in computational complexity.

__8__Since we do not allow survival probabilities to vary with income, this ignores the important issue of the actual amount of redistribution in the Social Security system. Liebman (2002)
These features are also important in driving the distributional dynamics of wealth. As shown by Hubbard, Skinner, and Zeldes (1995), Social Security and other government insurance programs drive those with low lifetime earnings to save little. This increases the dispersion of wealth in the economy. In addition, as numerous authors have pointed out, Social Security is a powerful risk-sharing tool.

Given the economic environment outlined above, it is useful to write the household’s problem recursively.

\[
V_t(A_{it}, P_{it}, \varepsilon_{it}) = \max_{c_{it}, l_{it}} u(c_{it}, l_{it}) + \beta \phi_t E \left[ V_{t+1}(A_{i,t+1}, P_{i,t+1}, \varepsilon_{i,t+1}) \right] 
\]

subject to

\[
c_{it} + a_{i,t+1} = a_{it} (1 + R_t) + w_{it} (1 - l_{it}) \\
- \tau_k (R_t a_{it}) - \tau_l (w_{it} (1 - l_{it})) - \tau_{ss} (w_{it} (1 - l_{it})) \\
- \tau_i (R_t a_{it} + w_{it} (1 - l_{it}))
\]

This recursive representation makes it clearer how solving the problem simplifies to solving a sequence of many one-period maximization problems. The state vector includes three variables: two continuous (the level of the permanent component of wages and the agent’s asset holdings) and one discrete (the transitory shock to income).

---

analyzes the actual amount of redistribution in the Social Security system. Differential mortality undoes much of the statutory progressiveness of Social Security.
1.2.2 Production

Since the focus of this paper is the effect of the tax code on individuals, production occurs by a representative firm in perfectly competitive markets. We abstract from all effects of different corporate tax regimes on allocation of capital between firms and focus on the impact of capital taxation on assets held by agents.

A standard Cobb-Douglas production function transforms labor and capital into output:

\[ Y_t = A_t F(K_t, N_t) = A_t K_t^\alpha N_t^{1-\alpha} \]  

(1.8)

Aggregate capital and labor reflect the aggregation of all agent’s decisions. We assume a closed economy and number all agents alive in the economy from 1 to \( I \). Aggregate capital is simply the sum of all individual capital holdings:9

\[ K_t = \sum_{i=1}^{I} a_{it} \]  

(1.9)

Meanwhile aggregate labor supply comes from aggregating labor supply decisions across individuals, weighted by their idiosyncratic individual productivity:

\[ N_t = \sum_{i=1}^{I} P_{it} \varepsilon_{it} (1 - l_{it}) \]  

(1.10)

Agents of different productivity contribute different amounts to aggregate productivity. Each hour of work by a high productivity individual contributes more to aggregate labor supply than the same amount of time spent working by a low productivity individual. Equation 1.10 reflects this by multiplying hours of labor supplied, \((1 - l_{it})\), by the individual productivity, \( P_{it} \varepsilon_{it} \). Economy wide wages are then

9 All of these could be written with integrals instead. For notational clarity and to maintain the link with the computational algorithm we describe in Appendix C.3 we write them as summations.
paid per efficiency unit of labor that an agent supplies.\textsuperscript{10}

Perfect competition implies that factors are paid their marginal products. Depreciation, $\delta$, occurs at the corporate level so agents receive interest rates net of depreciation. Wages and interest rates are then:

\begin{align*}
    w_t^e &= A_t (1 - \alpha) \left( \frac{K_t}{N_T} \right)^\alpha \\
    R_t &= A_t \alpha \left( \frac{N_t}{K_t} \right)^{1-\alpha} - \delta
\end{align*}

\textbf{1.2.3 Government}

The government levies taxes to pay for Social Security and government spending.\textsuperscript{11} The two tax systems are independent. Government spending is fixed at an exogenous percentage of GDP. We take no position on whether government spending belongs in the utility function of the agent. To avoid issues of how to specify utility from government spending we assume, at a minimum, separability of utility from government spending with utility from consumption and leisure.

\textsuperscript{10}One important feature to note is that different types of labor are identical. Changing the production function so that skilled and unskilled labor provide different inputs to the production function could lead to significant changes in the optimal tax schedules. This could occur in the following way. Currently all labor services are the same. A marginal tax rate hike at higher incomes reduces the supply of labor of those with higher incomes. Instead suppose there are two components of labor services. An unskilled component that all workers provide and an additional skilled component on top of that. In this case, increasing marginal rates at the top would push down the supply of skilled but might leave the supply of unskilled labor unchanged. If unskilled labor was a complement to skilled labor this would push down wages paid to unskilled labor hurting the poor.

\textsuperscript{11}The model ignores two potentially very important aspects in describing the government. First the government possesses full commitment. Second the tax change is completely unanticipated. There is obviously a tension between these two assumptions. Further, both conflict with observations of recent history in the United States. The tax code has changed fairly frequently and such changes have been at least partially anticipated. Attempting to include a lack of full commitment by the government, anticipation of the tax changes and potential tax changes would obscure the focus on the heterogeneous agent environment. Using simpler models of household behavior, Auerbach and Hassett (2002) studied how behavior responds to a stochastic tax regime. The seminal work of Kydland and Prescott (1977) deals with the problem of time consistency.
This paper focuses on the risk-sharing properties, distortions, and macroeconomic effects of allowing non-linear capital and non-linear labor taxes: it takes the Social Security system given. This applies both to its payouts and to the taxes that fund it. One could view the Social Security and other taxes as integrated and consider finding the optimal structure of a unified system. Changing the Social Security system at the same time as reforming the tax system, however, presents a different question. Instead we focus on the more limited question of changing the tax code while keeping the Social Security system constant.\footnote{Preliminary investigation indicates that there are large differences in the optimal tax code when Social Security is absent. Removal of this powerful lifetime risk-sharing mechanism increases the desirability of other means of providing insurance. The progressivity of both capital and labor taxes increases without Social Security.} This corresponds to considering the optimal tax reform of the majority of the tax system while leaving Social Security unchanged.

1.2.3.1 Form of Tax Function

The government applies a flexible non-linear tax function to the various quantities it taxes. The most important element of the tax function is its ability to mimic the current system of taxes in the United States while at the same time encompassing many popular reform proposals. Almost as importantly, the tax function does so parsimoniously making optimization over the parameters that define it feasible.

The tax function follows that estimated by Gouveia and Strauss (1994). They approximate the progressive nature of the United States income tax code by:

$$\tau_i(y) = a_0 \left( y - (y^{-a_1} + a_2)^{-\frac{1}{a_1}} \right)$$  \hspace{1cm} (1.13)

where $y$ is total income earned and $\tau_i(y)$ represents taxes paid on that income. First this function approximates the progressive nature of the current system. Then, when...
we separate capital and labor taxes, the functional form above applies to both of them separately. This functional form encompasses a wide range of taxes. Importantly, it embeds several economically and practically interesting cases. We can recover these by appropriate choices of parameters.

For instance, lump sum taxes correspond to \( a_1 = -1 \). A linear tax is represented by \( a_1 = 0 \). Progressive taxes occur when both \( a_1 > 0 \) and \( a_2 > 0 \). When \( a_1 \) is much greater than \( a_2 \) the tax system tends toward an interesting special form of progressive taxation, a flat tax with a large exemption. Figure 1.1 shows these cases and gives the parameter values that produce them.

Social security taxes are much simpler. In keeping with their actual form a linear tax applies to labor income only. We do not model the cap on Social Security earnings so the linear tax applies to all labor earnings.\(^{13}\)

1.2.3.2 Government Budget Constraints

We assume that the government balances its budget each period. Depending on the taxes being investigated, budget balance requires adjustments to different parameters in the tax code. Generally the adjustment is made to the rate of labor income taxation.\(^ {14} \) Like the rest of the budget, Social Security balances period by period. Varying the Social Security tax rate ensures revenues are equal to benefits paid.

\(^{13}\)This has potentially very important implications. Without this cap the marginal rates faced at the top of the income distribution are over ten percent lower than in the case with the cap. Imposing a cap on wages subject to Social Security taxation would greatly decrease the taxes paid by those with high wages. Since, as we see below, the planner desires to redistribute, capping Social Security contributions could substantially raise the optimal marginal tax rates at the high end.

\(^{14}\)By this we mean changing the parameter of our tax function that most closely corresponds to the rate of labor taxation. Generally this will influence the maximum rate of labor income taxation. In cases where there is no labor income taxation we next move to adjusting the overall rate of income taxation.
Figure 1.1: Various Tax Functions

This figure shows how the tax function can reproduce lump sum, linear, flat with an exemption and progressive taxes. Panel (a) shows a lump sum tax with $a_0 = 1$, $a_1 = -1$ and $a_2 = -1$. Panel (b) is a linear tax with $a_0 = 0.33$, $a_1 = 0.0$ and $a_2 = 1$. Panel (c) is a linear tax with an exemption stemming from $a_0 = .33$, $a_1 = 150$ and $a_2 = 1.0$. Finally, Panel (d) is a progressive tax with $a_0 = 0.33$, $a_1 = 1.35$ and $a_2 = 0.1$. 
1.2.4 Social Planners Objective

To compare two steady states requires assigning an objective function to the social planner. From a variety of plausible social welfare functions, we choose a utilitarian one. We assume that the social planner cares about all agents equally and sums the individual utilities of all agents alive. Not only does she care about the current welfare of agents in the economy, but also cares about the future utility of both existing agents and those yet to be born.

\[ U_{SP} = \sum_{t=0}^{\infty} \beta_{SP}^t \left( \sum_i u(c_i, l_i) \right) \]  

This utilitarian social planner has a strong drive to redistribute. If given a fixed amount of GDP and no behavioral responses, the planner would equalize consumption among agents.

When comparing steady states, the social planner need only look at one period’s summed utilities. In steady state, all future periods yield the same utility. Thus their discounted sum is a monotonic transformation of the single period summed utility. Maximizing the one period sum of agents’ utilities therefore maximizes the infinite sum in equation 1.14.

Note that the utilitarian social planner is not the only possible one to consider. For instance, Conesa and Krueger (2005) consider maximizing the initial value function of the agent. This corresponds to a Rawlsian veil of ignorance social welfare function. Agents care about their expected future lifetime utility before the beginning of their economic lives. Relative to our social welfare function, a Rawlsian one places more weight on the young. This differing weight stems from an agent at the beginning of their life down-weighting the future because of internal discounting from pure time preference and the risk of dying. In contrast, the utilitarian social
planner we consider cares just as much about the old as the young.

\section*{1.2.5 Equilibrium Definition and Computation}

Though numerically challenging, the concept of equilibrium in general equilibrium models with incomplete markets and idiosyncratic risk is well understood. For instance, Imrohoroglu, Imrohoroglu, and Joines (1999) contains a very good description of many aspects of solving this class of models in steady state.

Many fewer authors, however, solve for transition paths in this class of economies. The presence of general equilibrium, many different generations and many agents within each generation makes finding solutions challenging. Only recently have expansions in computing power made extensions of the steady state methods to solving for entire transition paths of overlapping generations models with idiosyncratic risk feasible. First solved in Conesa and Krueger (1999), these techniques have been used more recently in Nishiyama and Smetters (2005) and Conesa and Krueger (2005).

Appendix A.1 contains the solution technique for the agent’s life cycle problem. Appendices A.2 and A.3 contain formal definitions of the equilibrium and descriptions of the solution technique for solving for the steady state and transition path. More details are also contained in the references above.

\section*{1.3 Parametrization}

This section discusses parametrization of the model. Our parametrization draws on a variety of sources. Some parameters, especially those related to the agent’s wage processes and preference parameters, come from microeconomic estimates. Many of these are specifically estimated using life cycle models and simulated method of moments estimation. The model uses production parameters that are standard in the
literature. Finally, for taxes we calibrate off of the existing tax code and estimates related to it.

Consumption and leisure are non-separable. The period utility function is

\[ u(c, l) = \frac{(cl^v)^{1-\gamma}}{1-\gamma} \]  

(1.15)

When \( v > 1 \), consumption and work are complements. At lower levels of leisure the marginal utility of consumption increases. Agents who work a lot also consume a lot. \( \gamma \) is the coefficient of relative risk aversion over the combined utility from consumption and leisure. We take \( \gamma = 2 \) and \( v = 2 \). \( \gamma \) is the coefficient of relative risk aversion. Two is quite commonly used in the literature. A value of 2 for \( v \) has most working agents spending about \( \frac{1}{3} \) of their time allotment on market work. These parameters align closely with those estimated by French (2005) for this utility function.\(^{15}\) His simulated method of moments estimates of the preference parameters find this utility function fits better than an alternative which is additively separable between consumption and leisure.

The process for wages is estimated from PSID data. Essentially we estimate the process specified in section 1.2.1. Appendix A.4 contains more details on the estimation of the wage process. Wages follow a pronounced hump shape over the life cycle. Young agents, from twenty up to around thirty, have a very steep, upwardly sloping, wage profile. Wages then peak at around age fifty. Beyond fifty, wages decline.

Nishiyama and Smetters (2005) estimate the same taxation function used here off

\(^{15}\)French defines the utility function slightly differently. He estimates the utility function \( u(c, l) = \frac{(c^{\alpha l^{1-\alpha}})^{1-\gamma}}{1-\gamma} \). In one specification he estimates that \( 1 - \alpha = .602 \) and \( \gamma = 3.78 \). Transforming these to the terminology of our model implies that \( v = 1.51 \) and \( \gamma = 1.50 \). These are quite close to the numbers above.
of the United States statutory tax code. Following them, we parametrize the income tax function as $a_0 = .30$, $a_1 = .85$, $a_2 = .106$.\footnote{We have to make some adjustments to their original estimates. Their original estimates are $a_0 = .41$, $a_1 = .85$, $a_2 = .015$. Nishiyama and Smetters (2005), however, use different units than us. Converting to our correct units (we treat all values in tens of thousands of dollars for numerical stability) gives $a_0 = .41$, $a_1 = .85$, $a_2 = .106$. We also alter the parameter $a_0$ to match the ratio of personal taxes collected by the government to GDP with that observed in the data. Doing this takes into account that few people actually pay the full statutory rate. The tax code has a myriad of deductions and opportunities for income shifting which lowers the effective rate. This finally gives us the values reported above. Note that our approximately 25% reduction in $a_0$ is in line with Nishiyama and Smetters’s similar calculation.} This approximates the progressive nature of the current U.S. tax code while matching the amount of tax revenue the federal government collects. Figure 1.2 graphically shows the implied average and marginal tax rates of these parameters.

Mortality data come from the Social Security Agency. Bell and Miller (2002) contains survival probabilities for 2000 at all ages. We use the average of male and female survival probabilities at each age. Agents are born into our economic

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Source</th>
<th>Parameter Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.965</td>
<td>Gourinchas and Parker (2002); French (2005)</td>
<td>Rate of time preference</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>2.0</td>
<td>Gourinchas and Parker (2002); French (2005)</td>
<td>Risk aversion</td>
</tr>
<tr>
<td>$\nu$</td>
<td>2.0</td>
<td>French (2005); Nishiyama and Smetters (2005)</td>
<td>Trade-off between leisure and consumption</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>$\frac{1}{3}$</td>
<td>Standard</td>
<td>Capital share in output</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.06</td>
<td>Nishiyama and Smetters (2005)</td>
<td>Depreciation rate</td>
</tr>
<tr>
<td>$\phi_i$</td>
<td>varies</td>
<td>Bell and Miller (2002)</td>
<td>Survival probability</td>
</tr>
<tr>
<td>income process</td>
<td>varies</td>
<td>Author’s estimation</td>
<td>Earnings pattern over life cycle</td>
</tr>
</tbody>
</table>

Table 1.1: Summary of Parameter Values, Sources and Main Economic Interpretation
Figure 1.2: Approximation of U.S. Statutory Tax Code
This figure presents the function used to approximate the existing U.S. tax code. The figure shows the average and marginal rates associated with the tax function $\tau_i(y) = a_0 \left( y - (y^{-a_1} + a_2)^{-\frac{1}{a_1}} \right)$ when $a_0 = .30$, $a_1 = .85$, $a_2 = .106$. Taxes apply to total income which is the sum of income from labor and capital.

Our approximation to Social Security replaces income based on an agent’s final wage. It replaces a minimum of around $7000 and rises above that for agents with higher wages. Figure 1.3 graphically shows the transformation. The benefits increase from a floor of around $7,000 (the amount of SSI) to a maximum of about $16,000 for agents with very high wages. Consistent with the progressive nature of the statutory benefits structure, the replacement rate increases fastest near the bottom of the wage distribution and then levels off as wages increase.

\footnotetext{17}{The number of agents who survive to age 100 is under 1 percent.}
Figure 1.3: Social Security Benefits Function.
The transformation of the final wage into Social Security Benefits. Social Security provides a base level of income comparable to the benefits from SSI. The benefits rise in a non-linear manner comparable to the progressive actual replacement rates based on AIME. The “Final Permanent Component of Wages” corresponds to the realized wage if the agent worked the equivalent of full-time. In the model that corresponds to working for $\frac{1}{3}$ of their time allotment. Young and prime age agents spend about $\frac{1}{3}$ of their time in market work.

The production parameters for a Cobb-Douglas aggregate production function are standard in the macroeconomics and business cycle literature. Capital depreciates by 6% per year (Cooley and Prescott, 1995). The capital share in the production function, $\alpha$, is $\frac{1}{3}$.

1.4 Model Properties

We start by looking at some of the properties of the model’s steady state properties. This examination serves two purposes: (1) to show the model matches important features of the data, and (2) to build intuition about the economic behavior that
Table 1.2: Steady State Aggregate Quantities of the Model Parametrized to Current U.S. Data

The sources for the statistics in the data column vary. $\bar{K}$ is the average of the ratio of fixed assets to GDP from the BEA’s NIPA tables from 1995-2004. Income at 50 is from French (2005) converted to 2000 dollars using the GDP deflator. The ratio of government spending to GDP is the average from 1995-2004 federal government tax collections, excluding social insurance, divided by GDP. Again, data are from the NIPA tables. The value of $\tau_{SS}$ comes from the statutory rates and includes employers’ contributions. All model statistics are author’s computation.

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{K}$</td>
<td>2.77</td>
<td>2.46</td>
</tr>
<tr>
<td>Mean Income at 50</td>
<td>35,700</td>
<td>36,527</td>
</tr>
<tr>
<td>Mean Assets at 60</td>
<td>195,350</td>
<td>142,800</td>
</tr>
<tr>
<td>$\bar{Y}$</td>
<td>0.116</td>
<td>0.116</td>
</tr>
<tr>
<td>$\tau_{SS}$</td>
<td>12.4</td>
<td>12.6</td>
</tr>
</tbody>
</table>

underlies taxation choices.

1.4.1 Aggregate Statistics

Table 1.2 compares the model with the data along some key dimensions. Since parameters estimated from micro data are used to set many of the parameters, some aggregate statistics do not match exactly as they do in many calibrations. For instance, the model predicts a capital output ratio slightly below the 2.77 observed in the United States. Most large scale OLG models match this by increasing the discount factor $\beta$ closer to unity than our 0.965. Doing so, however, does not match well with the microeconomic evidence on the discount rate agents actually apply. As a consequence of fixing $\beta$ based on microeconomic estimation of life cycle consumption and saving decisions, the model does not perfectly match this ratio.

The low $\beta$ also leads our model to under-predict the assets held at 60. Increasing the discount rate closer to unity would also raise the assets accumulated at 60. This is completely endogenous to the model given the pattern of wages.

The Social Security tax rate also matches up quite well. The model specifies a
benefit function that reflects the stylized facts about Social Security. It then closely matches the tax rate needed with a rate of 12.6 percent as opposed to 12.4 percent. $G$ matches perfectly as we adjust the rate of taxation to ensure that government revenue and spending exactly equals 11.6 percent of GDP.

1.4.2 Life Cycle Patterns of Consumption and Wealth

Figure 1.4 shows the average consumption and wealth of an individual over their life cycle. Agents save little when they are young. What savings they do have, they accumulate for precautionary motives. The combination of the long time until retirement, borrowing constraints and a rapidly rising income profile combine to cause the young agent to save little. Even with an interest rate above their rate of time preference, saving is not optimal for the young. In fact, young agents would like to borrow against their future, higher in expectation, income. Only as they start to approach middle age, and both earn more on average and get closer to retirement, do the agents begin to accumulate substantial assets.

The drop in consumption at retirement is a function of the non-separability of consumption and leisure. The agents both anticipate and accept this drop. As they retire, the amount of leisure they take rises substantially. Given our parameter values, this increase in leisure reduces the marginal utility of consumption. With a lower marginal utility of consumption, the agent consumes less.

Krueger and Fernandez-Villaverde (2005) extensively estimate consumption over

\footnote{Borrowing constraints effectively shorten the time horizon of the agent. An agent who fears the borrowing constraint might be binding in the near future has a much shorter effective time horizon.}

\footnote{One way to view this is that instead of the standard arguments about the agent having a smooth marginal utility of consumption, the agent chooses a smooth marginal utility of the aggregate of consumption and leisure. Numerical results show that this smoothing does take place. Average total utility at retirement shows no drop.}
Figure 1.4: Average Consumption and Wealth over the Life Cycle
the life cycle. Their mean consumption profiles rise essentially smoothly until around age 50, peaks there, and smoothly declines from that age on. At higher ages, mean consumption falls below the consumption of the young. Gourinchas and Parker (2002) document a very similar pattern except their measure of consumption peaks slightly earlier, in the agent’s late 40s. Their estimates stop at age 65, but fall from their peak to that point. Figure 1.4 shows that we match the smooth rise and peak in consumption in the 50s. Instead of a smooth decline in consumption, we have a sharp drop when agents retire at age 65. The sharpness of the drop stems from our simplifying assumption that all agents retire at age 65, but consumption had begun to turn down even before then. More importantly our general equilibrium model with a labor-leisure trade-off matches three stylized life cycle facts about consumption: (1) the rise up until the early 50s (2) the peak around then and (3) the decline thereafter.

1.4.3 Variances of Consumption and Wealth

More recent work addresses how well OLG models, similar to the one studied here, match not just the means of consumption, income and wealth over an individual’s life cycle but also the variances. Storesletten, Telmer, and Yaron (2004), building on their previous work, show that households face idiosyncratic income shocks over the course of their life cycle. Further, they solve a general equilibrium model similar to the current one, though without a choice between leisure and labor. The endogenous consumption variance in their model matches their estimated variance profile of

\footnote{If agents chose their retirement date and had the option of drawing reduced Social Security benefits at an early retirement age at least some of this sharp drop would be smoothed out. Individual agents consumption would still drop on their retirement. The overall average level of consumption as an agent ages, however, would decline more slowly as only some fraction of agents would retire each period. This would generate both the observed fall in consumption around retirement and a smoother pattern of average consumption.}
Figure 1.5: Variances of Consumption and Labor Income Over the Life Cycle.

Cross-sectional variances of consumption and after-tax labor income. Consumption variance is always smaller than labor income variance over the working years. Labor income variance goes to zero for retired agents since all retire at 65 and have identical labor income of zero.

These variances are crucial for studying welfare. Since individual welfare varies with consumption and utility is concave in consumption, broadening the distribution of consumption (holding the mean unchanged) lowers aggregate welfare. Thus, to be able to analyze the impact of alternative tax arrangements on welfare, we need a model that matches the life cycle patterns of consumption, income and wealth variances.

Figure 1.5 shows the performance of the model in matching the variances of labor income and consumption over the life cycle. The profile of labor income variance lines up well with that estimated by Storesletten, Telmer, and Yaron (2004). They estimate income inequality rising over an individual’s life cycle from a fairly low initial
level. It peaks around 55 and slowly declines at higher ages. Consumption inequality also starts out low and rises over the life cycle. Their estimates of consumption inequality do not display the hump shape that income inequality does. Instead it levels off at around age 50. Importantly it always remains lower than income inequality. Storesletten, Telmer, and Yaron, using a model broadly similar to the one here, feed their model their estimated income process. They find their model matches the life cycle profile of consumption variances.

Adding elastic labor supply and income taxation, we still match the main features of the profiles of both consumption and income inequality. As Figure 1.5 shows, in the model, income inequality rises much faster than consumption inequality. Consumption inequality rises until the mid 50s and then levels out. Both of these match the data. We can further match a feature of the data that the model of Storesletten, Telmer, and Yaron do not. Our income variance turns downward as older agents start to reduce their hours of market work. This matches the estimation results of Storesletten, Telmer, and Yaron that the variance of labor incomes turns down around age 60. Since they do not have elastic labor supply, Storesletten, Telmer, and Yaron do not match this downturn.21

1.5 Steady State Optimality Results

We find that the optimal taxes on capital and labor are very dissimilar. The optimal tax on labor is essentially a flat rate of around 12 percent. Meanwhile, most capital

21One important point to note is that Figure 1.5 presents the model with raw variances. Storesletten, Telmer, and Yaron show the log variances. Showing the log variances for our model is not informative. The addition of flexible leisure means that some agents choose to work close to zero hours. Since $ln 0 = -\infty$ the variance of the log blows up. Since Storesletten, Telmer, and Yaron apply selection criteria to their estimation to remove values very close to zero they do not have these issues. Reporting the variance of the raw number is thus more appropriate.
Table 1.3: Comparison of Initial and Optimal Steady States

Table comparing certain key model statistics between the initial tax system and the optimal. This table compares steady state to steady state. Note that $V_0$ and $\sum u(c, l)$ are measures of utility. More is better so the larger value, which is closer to zero, is better. So the -0.98 is better than the -1.0.

<table>
<thead>
<tr>
<th></th>
<th>Initial</th>
<th>Optimal</th>
</tr>
</thead>
<tbody>
<tr>
<td>normalized $\sum u(c, l)$</td>
<td>-1.0</td>
<td>-0.98</td>
</tr>
<tr>
<td>$K/N$</td>
<td>3.83</td>
<td>3.96</td>
</tr>
<tr>
<td>normalized $Y$</td>
<td>1.0</td>
<td>1.051</td>
</tr>
<tr>
<td>normalized $K$</td>
<td>1.0</td>
<td>1.076</td>
</tr>
<tr>
<td>normalized $N$</td>
<td>1.0</td>
<td>1.040</td>
</tr>
<tr>
<td>$w$</td>
<td>1.0</td>
<td>1.011</td>
</tr>
<tr>
<td>Mean Assets at 60</td>
<td>142,800</td>
<td>162,200</td>
</tr>
<tr>
<td>$V_0$</td>
<td>-19.121</td>
<td>-19.127</td>
</tr>
<tr>
<td>$\tau_{SS}$</td>
<td>12.1</td>
<td>12.6</td>
</tr>
</tbody>
</table>

is untaxed. There is a very large exemption on capital income, with income above that being taxed at a higher marginal rate of over 20 percent. Both the rates and the shapes of the tax functions differ significantly between labor and capital. Figure 1.6, showing the marginal and average tax rates for labor and capital, drives home the difference. This difference in both shapes and levels is our main finding.

Table 1.3 showcases some of the major differences between the initial steady state and the optimal. The capital-labor ratio increases significantly from its current value. Both aggregate capital and aggregate labor increase, but the greater increase in aggregate capital than aggregate labor raises the capital-labor ratio above the initial level. GDP increases significantly, rising by over 5 percent from the initial level.

The increased incentives for an individual to save translates into significant increases in asset holdings. The average asset holdings at age 60 increases by about 15 percent. This increase reflects the lower effective marginal rates on capital accumulation for all but the very wealthy. Instead of paying a high marginal rate
Figure 1.6: Optimal Labor and Capital Taxes

The optimal average and marginal rates on capital and labor income. Labor taxes are essentially a flat rate of 11%. Capital taxes feature a large exemption. Assets up to about $100,000 face no taxes. Above that the marginal tax rates rises quickly to over 20 percent.
each additional dollar of asset income, most agents pay zero. This induces significant extra saving by agents.

1.5.1 Differences from Initial System

The optimal system differs significantly from the existing system. The marginal rate on most labor income falls significantly. Before the reform, marginal rates of at least 12 percent applied to any income over $12,000. Above that income level rates could rise to over 30 percent. All labor income over this level now faces lower marginal rates. At the lower end, however, agents face higher marginal and average rates. More of the tax burden falls on poor agents than before. Despite the lowering of marginal rates, agents tax burden can increase. For agents not too far above $12,000 in income, total taxes paid rise because of higher average rates. High wage agents are much better off as both their marginal and average rates fall substantially.

Average and marginal capital tax rates drop substantially as most capital now escapes taxation. The zero marginal rate on most asset holdings greatly reduces the number of agents who pay any taxes on their capital income. Compare this to the situation in the pre-reform economy in which all capital paid a marginal tax rate greater than zero. Most paid a much higher marginal rate since additional capital income was taxed at the marginal rate when both labor and capital income were already added in.

1.5.2 Life Cycle Tax Burden Distribution

Figure 1.7 shows the differential contributions of labor and capital taxes at different points in the life cycle. For the majority of an average agent’s life the tax burden falls entirely on labor. The large capital income exemption means that young agents pay no capital tax on their savings. Since they pay no capital tax it raises their...
rate of return to saving inducing them to save more leading to the higher capital accumulation discussed above. Note also that only agents near retirement age, those who have the most savings, pay any capital taxes. Even for these agents capital taxes make up only a small share of total taxes they pay.

The distribution of the tax burden over the life cycle fits in with the theoretical results of Erosa and Gerivas and Garriga. These recent theoretical papers consider OLG economies with linear taxes and without idiosyncratic risk. In this simpler case, they attain theoretical results about the nature of optimal taxes in life cycle economies. They show that, when the government cannot condition on age, it taxes capital a non-zero rate.\textsuperscript{22}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure1}
\caption{Life Cycle Pattern of Taxes Paid Under Optimal Tax System}
\end{figure}

\textsuperscript{22}Garriga (2003) shows that for the special case of $u(c, l) = \frac{c^{1-\gamma}}{1-\gamma} + \frac{l^{1-\gamma}}{1-\gamma}$ a zero tax on capital is optimal. The result holds only for the case where there is no heterogeneity within generations. Thus, even if we assumed these preferences, it would not hold in our model.
As Erosa and Gerivas (2002) note, the planner would like to have a declining tax on labor income as the agent ages.\textsuperscript{23} Since this is not possible, taxing capital helps to do this. Taxing capital creates intertemporal price distortions making leisure today cheaper than leisure tomorrow. This effectively allows the planner to tax the leisure of the old more and their labor supply less.

Non-linear taxes on capital give the planner a better tool to do this than the linear taxes that Erosa and Gerivas and Garriga (2003) consider. Now, instead of inter-temporally distorting all agents, the planner focuses on only distorting older agents. The ability to levy non-linear taxes allows the planner to tax only the capital of the old instead of having to tax all agents. Since the young agents hold capital below the exemption, they are not taxed. They face a zero marginal tax on capital so their intertemporal margin is not distorted. Older agents, who have built up their capital for life cycle reasons, are distorted. Since this distortion allows partial taxation of leisure of the old, a good that is consumed in relative abundance, it increases social welfare.

1.5.3 Increase in Average Consumption

Moving to the optimal tax system increases average consumption significantly throughout the agent’s life. Two factors drive this rise in consumption: (1) the macroeconomic feedback to higher wages, and (2) the increase in labor supply in response to lower marginal rates. This consumption increase is one of the main drivers of the welfare gain.

The first effect follows familiar logic. The microeconomic impacts of lowering the

\textsuperscript{23}The planner would like to make leisure, the untaxed good, more expensive when the agent consumes a lot of Agents consume the most leisure close to retirement. So the planner would like to lower taxes on labor income, thus making leisure more expensive, for old agents.
tax rate on saving leads to an increase in overall capital. This in turn increases the capital labor ratio and hence the wage. This increase in the wage then allows the agent to consume at a higher level.

The second effect is also familiar. Lowering the marginal tax rate on labor income induces agents to substitute work for leisure raising their earnings. The pattern of consumption also changes. Agents consume more in their peak earnings years both because they have more but also because they are working more. The increase in work reduces leisure and increases the marginal utility of consumption. Both the more pronounced increase during the working years and the overall rise in consumption are evident in Figure 1.8.

Along with the rise in the level of consumption comes a substantial increase in the variance of consumption. Two factors drive this increase. First, the substitution of an essentially linear tax on labor and no tax on most capital increases the resources of high earners relative to low earners. This pushes their consumption further above that of low earners increasing the variance of consumption. Second, through the non-separability outlined above, this pushes up their contemporaneous consumption. The combination of these significantly increase the variance of consumption over its current level.

1.5.4 Labor Tax Balancing

To understand why the optimum features these tax rates it helps to look more deeply into the driving economic forces in the model. Given the strong redistributive motives of the planner, a more progressive tax function seems like the expected outcome making a labor tax so close to linear quite surprising. Instead we get the linear labor tax as the planner balances opposing desires.

The vastly different elasticities over the life cycle sits on one side of the balance
Figure 1.8: Average Consumption Profiles Under Current and Optimal Tax Systems

The figure shows the age profile of consumption under the current tax system and after switching to the optimal. Note that average consumption rises significantly from the baseline. The lower marginal rates on middle aged agents induce a significant increase in labor supply and consumption.
when choosing optimal labor tax rates. How this elasticity varies greatly over the life cycle drives the trade-offs the planner makes. Figure 1.9 shows the evolution of the average labor supply elasticity over the life cycle.\textsuperscript{24} Young agents have a very low elasticity of labor supply. As agents age, the magnitude of their labor supply elasticity starts to slowly increase.\textsuperscript{25} As they approach retirement, their labor supply becomes very elastic, increasing sharply as they pass 60. Then, between 60 and 65, the elasticity doubles.

Precautionary saving explains a good deal of this movement. Young agents start with no stock of assets. Without any assets they are unable to buffer the idiosyncratic income shocks they face. Thus the amount of labor they supply varies little with the shocks they face. As they age, agents start to build up savings. Building up these savings allows them to use a buffer stock of assets to smooth their consumption. Since they can smooth their consumption they can vary their labor supply more in response to shocks.

Also, as they approach retirement, the combination of a large stock of assets and imminent tapping of Social Security increases the labor supply elasticity of old agents even more. Since agents at this point have built up a substantial stock of assets for retirement, they can easily buffer their consumption with these assets and decrease their amount of work when hit by negative shocks. Since Social Security has a substantial floor and a progressive benefits transformation, the wage shocks have only a limited impact on the agents' future expected benefits. So they do not

\textsuperscript{24}The elasticity is in response to a change in the permanent component of an agent's wage. We choose to measure in response to a permanent change as this is closer in spirit to considering a change induced by taxation. Though the thought experiment of considering a change in their permanent component of wages is not quite the same as considering how their labor supply changes in response to a small change in taxation it is quite close and the intuition quite useful in understanding our results.

\textsuperscript{25}French (2005) points out the age dependence of labor supply elasticity and in a broadly similar model finds a similar pattern over the life cycle.
Figure 1.9: Labor Supply Elasticity of Agent

The labor supply elasticity of an agent over the life cycle. Young agents, with no buffer stock of savings, have a very low elasticity. As agents age, build up assets, and get closer to retirement their labor supply elasticity increases greatly becoming very large near retirement.

increase work to offset future lost benefits. This greatly increases the temptation to cut back work in the present and wait for the benefits they will soon be able to claim.

Returning to the underlying intuition of optimal taxation, taxing an inelastically supplied good minimizes the distortions induced by taxation. On this basis, a planner desires to tax the labor supply of the young heavily. Their labor supply responds very little to changes in their wage. The distortions from taxing the young are thus quite low. Meanwhile the labor supply of the old reacts much more to changes in their wages. If the old are taxed heavily they cut back the amount of time they work dramatically. Since they distort the decisions of the old so much, high marginal tax rates on the old have much larger losses stemming from the taxes’ distorting effects.
Offsetting this, however, is a desire to redistribute. Young agents have both low wages and low assets. Returning to Figure 1.4, the average consumption of a 65 year old is approximately 50 percent higher than a 20 year old. This combination leads the young to have a high marginal utility of consumption relative to older agents. The combination of an upwards sloping profile of average wages, borrowing constraints, and incomplete markets means that agents cannot smooth consumption across their lifetime. Figure 1.10 shows how a marginal dollar of assets greatly increases the value function of a young agent. As agents age, they both save a stock of assets and have their wages increase. This combination makes the marginal dollar much less valuable to them. Given this pattern a social planner can increase aggregate welfare by transferring money from older agents to younger agents. The combination of incomplete markets for their idiosyncratic wage risk and borrowing constraints prevent the agents from transferring the consumption across time themselves. Since they cannot transfer the future consumption of their middle aged selves back to their young selves government measures that do so increase their welfare.

The derivative of the value function with respect to assets, $\frac{\partial V}{\partial a}$, rather than the marginal utility of consumption is the appropriate object to consider. Young agents who receive extra assets may not immediately consume it. Instead they will use at least part of it to augment their buffer stock of savings. Spreading it out and avoiding the possibility of binding liquidity constraints means their marginal value will generally exceed their marginal utility of consumption. So looking at the marginal utility of consumption would understate how much a young agent values those assets.

The increasing social welfare by transferring assets holds only with appropriate caveats about the distortions not being too large.

Hurst and Willen (2005) note this same potential gain with Social Security. Using a life cycle model they show how shifting the Social Security tax burden from the young to the middle aged can lead to welfare improvements. Much of those gains come from paying off higher rate unsecured debts of the young.
1.5.5 Capital Tax Balancing

The capital tax also balances competing concerns of the social planner. Here the social planner is again balancing her desire to redistribute, but this time against the distortions of the capital tax. Taxing capital has the extremely negative effect of retarding capital formation which translates to lower wages. This argues for lower capital taxes. Against that is the fact that in a world with Social Security the poor hold little capital. Since they have low assets, a higher tax on capital income and a lower tax on labor income benefits them greatly. The planner again balances out these two effects.

In a world with Social Security, the majority of capital is held by those who have had a series of positive shocks. Those who have a high level of income have a strong desire to save to prevent a drop in consumption on retirement. For those at low levels of income, Social Security replaces a large fraction of their income. Since so much of their income is replaced they need to save much less to smooth their consumption and so hold fewer assets. For those at higher levels of income Social Security, benefits are insufficient to replace the lost income. To maintain an appropriate level of consumption requires much higher savings rate.

The capital tax also presents a way to get around some of the balancing that takes place in the labor tax. By taxing the capital of those close to retirement, the capital tax provides a way to tax these old agents who have a very high labor supply elasticity. Since taxing their labor supply at a high marginal rate causes them to cut back their labor supply sharply, high marginal rates on them is not an effective

\footnote{The tax on labor income could be lower because of the additional revenue raised by the higher capital tax.}

\footnote{Note that we say appropriate rather than equal level of consumption. Our assumptions on preferences mean that even for rich agents retirement involves a drop in consumption. This follows from the non-separability of consumption and leisure.}
strategy for achieving redistributive goals. Their capital holdings, however, exist mainly to smooth consumption when their income drops on retirement. They are less sensitive to changes in the marginal tax rate on capital. So, taxing capital of agents with large supplies of it taxes agents with a high labor supply elasticity in a less distortionary way.

1.5.6 Discussion of Optimum

Despite increasing aggregate welfare in steady state, the move to the optimal tax system does not maximize every reasonable welfare criteria. In fact, using a Rawlsian measure of the expected utility of a newborn agent, welfare actually decreases.
Table 1.3 on page 30 shows this with $V_0$, the initial utility, being higher under the current system than the optimal. A newborn agent would, despite the reasons for improvement outlined above, prefer a tax system like the current one to the optimal.

Why do we get this result? It stems from where in the life cycle the objective function places weight. With the Rawlsian emphasis is placed on a young agent. The Rawlsian objective discounts the far future heavily because of both normal discounting and mortality risk. In contrast, a utilitarian welfare function places as much weight on old agents who have survived as the young ones. Recall from Figure 1.10 that any transfers of wealth to a young agent have a very high marginal value for the young agent. The optimal system lowers capital taxes making older agents better off on average. That revenue must be made up somewhere. It comes from taxing younger agents more at higher average, though lower marginal, rates. This reduces the expected lifetime welfare of a young agent. In addition, losses from less risk-sharing come from moving from a more redistributive system to a less redistributive one.

This helps to explain why our results on optimal taxation look quite different than those of Conesa and Krueger (2005). They examine a Rawlsian social welfare function. Doing so, they find the optimal income tax features a large exemption and then a flat rate. Since their planner cares more about the young than ours, she redistributes more to the young be exempting essentially all of their income from taxation.

Solving for the optimal steady state tax system is not the same as solving for the optimal tax system for the government to impose now. As discussed below, it

\[31\text{Conesa and Krueger (2005) examine only combined income taxation. Some of the difference might also stem from the inability to separate labor and capital and the quite different effects of taxes on each.}\]
fails to take account of the costs of transition and different generations bearing the costs from those that reap the gains. So why look at it? Solving for the optimal transition path of taxes, even with our parsimonious tax function, is computationally infeasible.32 We view solving for the optimal and then considering the transition as an approximation. Looking at the optimal steady state and also including the transition to it gives an idea about whether or not such a tax reform leads to welfare gains or losses. It shows the direction that such an optimal reform would probably move in as well as giving an idea of how far our current system is from optimal.

1.6 Adding a Lump Sum Levy

Capital and labor taxes are not the only taxes that government have access to and use in practice. Besides these, governments either now tax or have in the past total income, consumption and lump sum taxes. Economic theory suggests advantages for all of these. Lump sum taxes have the attractive property that they do not distort any marginal decisions. Income taxes use information on total income which might better proxy for an agent’s total resources. Linear consumption taxes do not distort the labor supply margin. These are favorable properties which indicate that in more complex models, such as the current one, there might be gains from these various other sorts of taxes

As mentioned before, Conesa and Krueger (2005) study the optimal taxation of income in a similar environment. Similarly, Smyth (2006b) examines how consum-

32Solving for the optimal transition involves solving the non-linear optimization problem where the number of parameters being optimized over is very large. The set of parameters is now the same ones we have now except at every period along the transition. Solving over this large of a state space is not feasible. Finally, even parameterizing the transition to be a smooth class of functions, proves computationally intractable. Solving for the transition path takes two orders of magnitude more computing time than calculating a steady state.
tion taxes should be formulated to optimally balance the issues discussed here. This section considers a perturbation of the admissible set of taxes. This perturbation allows the government access to a lump sum levy, either positive or negative, in addition to the taxes on capital and labor.

1.6.1 Adding a Lump Sum Levy

Purely lump sum taxes would not be optimal in this economic environment. The presence of idiosyncratic risk and the distribution of agents it cause mean that a planner who weights agents equally will desire to reduce the variance of utility across agents. Since marginal utility declines with increases in consumption, the planner will desire to use taxes to move some consumption from agents with high consumption to agents with low consumption. Lump sum taxes do not allow this since all agents pay the same tax. The planner is willing to pay the price of distortionary taxation in order to redistribute. Also, a lump sum tax can hit agents with very low consumption hard. These agents then end up with extremely low consumption. Since the marginal utility of consumption becomes infinite as it approaches zero, given some agents very low consumption has extremely negative effects.

Adding a lump sum extends the set of taxes the government has access to allowing two possible results. First, the government could increase redistribution by choosing a negative lump sum tax (so a positive lump sum transfer). This would allow non-distortionary, except for income effects, transfers to the poor. However, there would be a distortionary increase in other taxes to fund this. In contrast, the government could raise money in a non-distortionary manner with the lump sum tax and use the revenue from that to lower other taxes in the economy. Which raises welfare more

\[33\text{And in fact, are not. Replacing the income tax with a lump sum tax yields a lower steady state aggregate utility.}\]
is a quantitative matter which the model allows us to investigate.

Formally, adding the lump sum tax is a simple modification of the agents budget constraint. Instead of a budget constraint given by Equation 3.3, the agent faces a new one with lump sum taxes added. Considering the case with only capital and labor taxes, this becomes

\[
c_{it} + a_{i,t+1} = a_{it} (1 + R_t) + w_{it} (1 - l_{it}) - \tau_{LS} - \tau_k (R_t a_{it}) - \tau_l (w_{it} (1 - l_{it})) - \tau_{ss} (w_{it} (1 - l_{it}))
\]

(1.16) (1.17)

where \( \tau_{LS} \) is the lump sum component of taxation. Since the lump sum taxation is constant across agents, it does not depend on any of the agents characteristics.

1.6.2 Changes in Tax Functions Induced by Lump Sum Component

The optimal lump sum tax turns out to positive. The planner levies a tax of just over $2000 on all agents in the economy. Along with the lump sum tax occur changes in the shape and level of the optimal capital and labor taxes. The capital tax maintains the same shape, a big exemption and a fairly steep phase in to a maximum rate. Since some of the governments revenue now comes from the lump sum taxation, the government is able to slightly lower the tax on capital, from around 22% to 18%. Also the government shifts the point at which the capital tax is above zero up significantly. Previously capital above about $100,000 was taxed. The addition of lump sum taxes shifts this up to over $150,000. This shift removes almost all taxation from capital. As shown below in Figure 1.12, now only a trivial amount of capital is taxed.
Figure 1.11: Optimal Taxation when the Government has Access to Capital, Labor and Lump Sum Taxes
The tax on labor, in contrast, shifts a lot. Instead of being nearly flat, the labor tax now also takes the form of an exemption, a moderately gradual phase in, and a constant rate above that. This new labor tax is more progressive than the flat labor tax without the lump sum component. As Figure 1.11 Panel (a) shows, the lump sum tax causes average taxes to start out much higher and decline with income up to a point. At about $30,000 in labor income average rates reach their minimum and turn up again beyond that. Note from Panel (c), which compares the average taxes under the optimal systems with and without access to lump sum taxation, that there is an area where average taxes are lower than under the optimal system without lump sum taxation. This region, from about $20,000 to about $45,000 happens to be where the majority of agents are. The ability of the government to use a lump sum tax enables lower average tax rates on what is essentially the lower middle class.

1.6.3 Impacts Over Life Cycle

The addition of the lump sum tax allows more progressivity than before. Since taxation is now more efficient because of the lump sum portion, the labor portion of the tax code becomes more progressive, though it must be to offset the regressivity of the lump sum portion. The more efficient taxation, operating through the lower marginal rates, can be seen in Figure 1.12 Panel (a). Consumption is higher for the access to lump sum sum taxes case than under the current system, though at most ages not as high as for the optimal capital and labor taxes. Panel (b) makes clear the increase in redistribution that raises social welfare, despite the mostly lower consumption. Under the system with lump sum taxes, the variance of consumption

\[34\] In this discussion I assign all the average tax from the lump sum to income taxes rather than to capital taxes. Since the lump sum tax is not associated with marginal decisions for either activity, there is no compelling reason to assign it to either. The assignment made is primarily for purposes of elucidation.
(a) Mean Consumption
(b) Variance of Consumption
(c) Total Taxes Paid under Different Systems
(d) Burden of Different Forms of Taxation

Figure 1.12: Optimal Taxation with Capital, Labor and Lump Sum Taxes
drops relative to either the current system or the optimal without lump sum taxes. The variance of consumption is much lower than under the optimal capital tax system. The increased efficiency of taxation allows for a corresponding increase in the amount of compression of incomes the planner can cause.

The planner chooses not to use capital taxes to try shift some burden from the middle aged toward those near retirement as it did before. Panel (c) shows the total taxes paid under the current system, the optimal capital and labor taxes and the optimal capital and labor with lump sum. Adding the lump sum taxation takes some of the tax burden off the agents between 50 and 65 relative to the other two systems. It taxes retirees much more heavily. Looking at (c) and (d) the higher burden placed on retirees by the lump sum portion of the tax is clear. Also clear from (d) is how small of a role capital income taxes now play. Even before, only a minority of agents paid them. This has dropped yet further and almost all capital is exempt from taxation.

The substitution of lump sum taxes for some of the capital and labor taxation has positive macroeconomic impacts. Table 1.4 shows how these change. The capital-labor ration jumps 10% from 3.91 to 4.21 inducing an increase in both wages and output. A slight decrease in labor supply and increase in capital compared to the taxes without access to lump sum taxation drives this. The higher taxes on the high productivity individuals pushes down their labor supply relative to the no lump sum optimal case. Welfare, measured either by the sum of all agents utilities or the value function of a newborn agent also increases. The increase for an entering agent is equivalent to giving them a cash payment of $4,300.
Table 1.4: Comparison of Optimal Steady States Under Different Tax Regimes
Table comparing certain key model statistics in steady state between the various optimal tax systems.

<table>
<thead>
<tr>
<th></th>
<th>Initial</th>
<th>Optimal $\tau_k, \tau_l$</th>
<th>Optimal $\tau_k, \tau_l, \tau_{LS}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>normalized $\sum_u(c, l)$</td>
<td>-1.0</td>
<td>-.990</td>
<td>-0.979</td>
</tr>
<tr>
<td>$\frac{K}{N}$</td>
<td>3.83</td>
<td>3.91</td>
<td>4.21</td>
</tr>
<tr>
<td>normalized $Y$</td>
<td>1.0</td>
<td>1.048</td>
<td>1.050</td>
</tr>
<tr>
<td>normalized $K$</td>
<td>1.0</td>
<td>1.064</td>
<td>1.121</td>
</tr>
<tr>
<td>normalized $N$</td>
<td>1.0</td>
<td>1.040</td>
<td>1.017</td>
</tr>
<tr>
<td>$w$</td>
<td>1.0</td>
<td>1.010</td>
<td>1.032</td>
</tr>
<tr>
<td>$V_0$</td>
<td>-19.194</td>
<td>-19.182</td>
<td>-18.94</td>
</tr>
</tbody>
</table>

1.7 Transition

To fully assess the gains or losses from switching to a different tax system simply comparing steady states is insufficient. The paths of all variables along the transition will vary. Since agents have an existing distribution of asset holdings it will take time for this distribution to converge to the new steady state one. Along the transition there can be shifts in aggregate prices that either help or hurt agents. The direction the change in welfare goes depends on the direction of the price changes and the initial state of the agent. This implies that a tax reform can make agents worse off once the transition is accounted for despite gains that accrue in the steady state. Thus the only way to consider welfare is to explicitly compute the transition path and compare welfare from making the switch to no change.

In addition, the presence of multiple generations makes a large difference. The agents who already exist and who bear any burdens along the transition path are different than those who eventually reap the gains from the improved steady state. For instance, shifting to a higher tax on capital disadvantages existing old agents. They saved under the assumption of lower rates on capital income. Suddenly increas-
ing their rates of capital income taxation, even if it increases steady state aggregate welfare, hurts them. These agents bear the cost of the higher taxes on their savings and then die without receiving any of the benefits that accrue.\footnote{For instance, raising the tax on capital might allow a lowering of the tax on labor income. This in turn would effectively transfer money from old agents to young agents. These initial old agents would pay the cost in terms of having to give some of their consumption to young agents. They would never have received the benefits of the increased consumption when young.}

In considering the transition we take the following scenario: the government makes a surprise transition to the new tax system that no agents anticipate. Up until the announcement, agents expect the old tax system to continue indefinitely. Upon this sudden transition, agents labor supply, savings and consumption decisions all change. To reach a new steady state takes a transition of many years as the agent’s asset holdings and the economy-wide distribution of asset holdings adjusts only slowly. In addition, to maintain a balanced budget, tax rates over the transition vary. Appendix ?? contains details on solving for the transition path.

\subsection{Evolution of Aggregates}

Figure 1.13 shows the evolution of output, the capital-labor ratio, aggregate consumption, wages and interest rates. Over the course of the transition the capital-labor ratio rises by over 3 percent. Though both capital and labor increase, capital increases more than labor. Despite this, there is an initial fall in the capital-labor ratio. The initial fall comes from labor supplied to the market jumping up on imposition of the new tax scheme. Capital adjusts upward only slowly leading to the initial fall in the ratio. Building up to the new steady-state capital stock takes a while.\footnote{This is the reason the results of Chamley (1986) and Judd (1985) need not hold in OLG economies. Their very high short run tax rates on capital act as a lump sum tax on existing capital. This raises welfare as it is later made up for by the removal of the intertemporal distortions. Crucial is that the same agents who are being hurt in the present are helped in the future. In OLG economies this is no longer the case. Those agents who are hurt now die before they reap the gains.}
Figure 1.13: Evolution of Aggregate Variables Along Transition Path

(a) GDP

(b) Capital Labor Ratio

(c) Consumption

(d) Interest Rate

(e) Wage
Agents have to forgo consumption to save and increase the capital stock. Saving is, however, the last thing they want to do since they work more which increases their demand for consumption. So labor supply actually overshoots its long run level and slowly comes back down.

Interest rates and wages reflect the movements in the capital-labor ratio. In the short run, interest rates rise and wages fall. Note that, despite the fall in wages, total labor income increases from the expansion in labor supply. From an individual agent’s perspective, post-tax labor income increases significantly. The rise in the interest rate helps induce agents to save more, pulling up the aggregate capital stock. Beyond this short run fall, the after-tax interest rate still increases for agents because of the drop in marginal rates on capital income.

The increase, however, in these aggregates does not necessarily indicate an increase in welfare. The losses from risk-sharing can outweigh the gains in overall production. All the benefits might be in the future so that the short-run pain of the transition is greater than the benefits since the future benefits are discounted. Also, if all the gains accrue to the rich agents, aggregate welfare can decline. So we next turn to examining the welfare of agents alive at the transition.

1.7.2 Welfare gains

Accounting for the transition, changing to the new tax system generates only small aggregate welfare gains. We take as a measure of welfare gains the amount of extra assets an agent would have to have to be as well off under no reform as if the reform took place. This is the amount of money an agent would pay to have the reform implemented. If the amount is negative this is the amount the agent pay to not have
Figure 1.14: Percent of Initially Alive with Welfare Gains

Percent of each initially alive cohort who experience expected gains in welfare from the transition to the optimal tax system taking account of all the changes along the transition path. Welfare gains correspond to an increase in the value function, $V(a_i, w_i)$, when the new tax system is announced given an agent's existing assets and wage.
the reform implemented.\textsuperscript{37} This ignores any effects on aggregate quantities of these transfers and the funding of these transfers.\textsuperscript{38}

The aggregate welfare gains are only about 10 percent of initial GDP. When we sum all these individual welfare gains, accounting for the transition, there is a positive though not large increase in our measure of welfare.

The gains from the reform of the tax system do not accrue uniformly. Some agents do much better than others. Agents with higher wealth levels and higher incomes benefit more from the reduction in marginal rates at the top. Young agents benefit from the lower marginal tax rates as well as the higher expected wages in the future.\textsuperscript{39} In addition, all agents lose to varying degrees from the decline in risk sharing due to the less progressive nature of the new tax system.\textsuperscript{40} The tax system no longer provides as much insurance against the idiosyncratic wage shocks as before.

Figure 1.14 shows how the welfare gains accrue across the age range of the population. Only 45 percent of the population would vote to support a move to the optimal. Despite the aggregate gains in steady state there are many losers from the tax reform.

\textsuperscript{37}Formally consider the value function without reform as $V_{NR}(a_i, P_i, \varepsilon_i)$ and the value function with a reform as $V_{R}(a_i, P_i, \varepsilon_i)$. Our measure of the gains (or losses if negative) from the reform is $g$. When $V_{R} > V_{NR}$ we define $g$ as solving the equation $V_{R}(a_i, P_i, \varepsilon_i) = V_{NR}(a_i + g, P_i, \varepsilon_i)$. If $V_{R} < V_{NR}$ we instead define $g$ as solving $V_{R}(a_i - (-g), P_i, \varepsilon_i) = V_{NR}(a_i, P_i, \varepsilon_i)$ If $g > 0$ the agent would need to receive a transfer of $g$ to be as well off as if the reform took place. If $g < 0$ the reform hurts the agent. The addition or subtraction of $g$ is broken into two parts to avoid situations where adding or subtracting $g$ moves assets outside the range over which the value function is defined.

\textsuperscript{38}In considering the efficiency of consumption taxes, Nishiyama and Smetters (2005) account for these transfers fully. They take into account the distributional effects of doing so on agents savings and leisure time decisions as well as the macroeconomic impacts. To do so they construct a Lump Sum Redistribution Authority to make all agents at least as well off as they were before the reform. They find that shifting to a consumption tax then entails welfare losses.

\textsuperscript{39}The young agents will live and be working for the higher future wages. Older agents will not. Thus the beneficial effects further out in the future benefit them more.

\textsuperscript{40}Though this loss is outweighed by the gains for rich and young agents.
The old all dislike the reform. Their dislike stems from the increase in taxes on their Social Security benefits. Social Security benefits are taxed as labor income within the model. For old agents, benefits are low enough that the average taxes paid on on their Social Security benefits rises. This more than offsets the lower taxes they now pay on capital income.\footnote{There are also interactions with the annuity market. Since Social security is the only source of annuities agents value it greatly. Capital holdings can partially proxy for annuities, but the agents then face longevity when choosing how to consume down their assets. So, increasing the tax they face on the Social Security benefits causes large welfare losses for older agents since it affects their annuitized wealth, which is particularly valuable. If agents could annuitize wealth, this effect would decline.} The older agents have not built up enough assets to offset this drop in their effective Social Security benefits. In the new steady state, mean asset holdings before retirement are higher. These old agents, however, have not been able to do the additional saving over their life cycle and so are worse off. Reducing the models implied taxes on Social Security would increase the support among the old.

Younger agents benefit from the reform. Their marginal tax rates when they are working are reduced. The combination of lower marginal tax rates and increased hours of work increases their post-tax income. They also benefit from the future higher wages as they are still working when wages are higher. This is enough to offset their losses in reduced risk sharing. Finally, and importantly, they can build up a higher capital stock to offset the reduced Social Security benefits.

1.8 Conclusion

Using a calibrated general equilibrium model with uninsurable idiosyncratic risk we showed that the optimal taxes on capital and labor are extremely different. The government taxes labor at a flat 12 percent rate. In contrast the vast majority
of capital is untaxed. Capital income from assets under $100,000 is exempt from taxation while income above that threshold is taxed at a marginal rate of 20 percent.

Despite the complexity of the model, conventional economic factors drive these results. For the labor tax, the social planner weighs the different labor supply elasticities over the agents’ life against the desire to redistribute from the rich old to the young poor. Likewise, the capital tax involves balancing between the distortionary effects on capital formation and the same desire to redistribute.

Adding a lump sum to the arsenal of taxation does not cause the planner to redistribute through it. Instead, it removes even more capital from being subject to taxation. Along with this, average taxes are cut for the middle class and the progressivity of the labor part of the tax system increased. More efficient taxation allows for an increase in redistribution.

This class of models generally has difficulty matching the very top of the income distribution. The extremely wealthy achieve that status through means not modelled in this paper. Entrepreneurship and dynastic connections across generations could be more important for these individuals. As their behavior matters greatly for both tax revenues and the overall level of capital in the economy, integrating these concerns into our models of taxation will help to provide insights about optimal taxation.

Another limitation of the model is in describing the way old agents draw down their assets. Actual old agents draw down their assets much more slowly than those in the model, with many old agents maintaining substantial stocks of assets. This suggests roles for bequests or health shocks hitting the old as either of these would induce old agents to hold more assets.

Finally, the actual tax code has much greater complexity than we modeled. We abstracted from corporate income taxes, capital gains taxation, deductions for housing and income from personal businesses. All of these affect the distribution of
income and the distribution of the tax burden across the population. Incorporating these into this class of model can help to understand the actual incidence and distributional consequences of these taxes.

The results also suggest several directions for further research. Most capital being free of taxation while capital above some exemption being taxed heavily is reminiscent of IRAs and 401(k)s. Adding tax-free savings accounts to an income tax system places a zero marginal tax on most capital. Capital outside the accounts faces a high marginal tax rate. Such a system could well approximate the optimal non-linear capital tax we found and so capture some of the gains. This would not, however, have the beneficial impacts on labor supply.

Another possible direction for further research is the interaction of more complex models with age-specific taxes. At least some of the gains come from the capital tax differentially taxing the old and the young. If tax functions at different ages could be shifted directly this might remove the need for any capital taxation.

Finally, one of the main economic drivers was the differing elasticities of labor supply across the life cycle. The elasticity of the old is much larger than that of the young. Estimation of this elasticity over the life cycle could provide valuable information about whether this key theoretical effect exists in practice. This could also help elucidate some of the puzzles of differing estimates of the elasticity of income with respect to changes in the tax code.
Chapter 2

Value of a Statistical Life Over the Life Cycle

with Joe Aldy
2.1 Introduction

Extensive theoretical and empirical research has focused on how individuals' willingness to pay for mortality risk reduction varies over their life cycle. Early theoretical work showed that the value of life declines with age assuming that markets exist for individuals to smooth their consumption over their life cycle. Subsequent studies illustrated how the value of life could take an inverted-U shape with respect to age if individuals at young ages face borrowing constraints. Revealed preference and stated preference studies have estimated a variety of age patterns for the value of life. Some have found that the value of life declines with age; others found an inverted-U age relationship for the value of life; a few papers found no age variation in the value of life; and one recent revealed preference paper estimates that the value of life may increase with age. Recent empirical research has also identified differences in the value of life by income, gender, and ethnicity.

The lack of consensus in the literature may explain the differences in how measures of age-specific value of life are used by government agencies. The Food and Drug Administration has employed a constant value of a statistical life-year to monetize benefits of risk-reducing regulations.\(^1\) A constant value of a statistical life-year is consistent with perfect life cycle consumption smoothing and the presence of perfect annuity markets, although such a presumption is inconsistent with empirical observation. The Environmental Protection Agency and the Consumer Product Safety Commission typically monetize risk reduction benefits with an age-invariant value of a statistical life.\(^2\) We are not familiar with any theoretical work in the literature that

\(^1\)Refer to the FDA's regulation limiting tobacco use among adolescents (61 Federal Register 44396) as an example.

\(^2\)Refer to the EPA's regulation requiring emissions controls on cars and light trucks to limit sulfur dioxide emissions (65 Federal Register 6698) and the CPSC's regulation establishing label requirements for charcoal (60 Federal Register 40785) as examples.
supports the notion that individuals’ willingness to pay for mortality risk is constant over their entire life cycle. The Environmental Protection Agency, in its analysis of the Clear Skies Initiative, and the Canadian health ministry, in an assessment of a tobacco regulation, used a value of a statistical life step function: constant until age 70 (age 65 for the Canadian analysis), then a discrete drop to a lower value for all older individuals (Hara Associates, 2000; Environmental Protection Agency, 2002). Again, the literature is silent on a theoretical basis for such an approach.

These different standards for assessing the benefits of mortality reducing policies can have important implications for public policy. For instance, in the Clear Skies Initiative referenced above, 70% of the mortality reduction occurs among those over 65. If the statistical lives saved among the elderly were valued less than under the current analysis, this could change the optimal level of pollution allowed. Numerous other regulations, which by executive order must be subjected to benefit-costs analysis, feature this differential mortality by age.

At the same time, it is important to note the political realities that are intrinsically part of valuing life, especially when done in a democratic process. Imposing differential values on lives of different agents, whether they differ by age, income, wealth, race or sex, is politically contentious. The above methodologies might be an imperfect compromise that is necessary to get any benefit-cost analysis done. Imperfectly accounting for differences in the value of life, but still trading off the benefits and costs of a regulation, improves policy over cases where the issues are not considered at all. Knowing where and why these commonly used methods fail, however, can provide additional useful insight into considering these problems.

To investigate the question of how the value of life varies over the life cycle and attempt to explain some of the variations in the empirical literature and policy
practice, we develop a life cycle model of the consumption and labor supply choice problem facing an individual. Numerically solving this problem under uncertainty, we assess how much wealth an individual would forego for a marginal reduction in mortality risk. Calibrated to labor market, wealth, and longevity data for the United States, we can address how the willingness to pay to reduce mortality risk varies with age, income, wealth, and other individual characteristics correlated with longevity expectations.

This analysis makes several contributions to the literature. First, our numerical approach advances the theory and simulation literature focused on valuing mortality risk reduction. We have modeled the individual’s problem in a more realistic fashion that allows for the individual to choose optimal consumption and leisure over the life cycle. Unlike earlier papers that only specified consumption as a choice variable in the life cycle problem (Shepard and Zeckhauser, 1984; Johansson, 1996), this allows us to better match our model to observed consumption and leisure patterns and to specify willingness to pay for risk reduction as a trade-off among consumption, leisure, and risk. Leisure has value to agents so neglecting it can understate the benefits to mortality risk reduction. In addition, ours is the first paper in this literature to explicitly account for uncertainty in the income process. Modeling uninsurable idiosyncratic labor income risk yields consumption and saving behavior consistent with the empirical evidence on precautionary saving among the young and wealth accumulation typically starting in middle age. This has important results for how willingness to pay varies both over the life cycle and across the cross-section of the population.

Second, our numerical simulations explain some of the observed variations in the value of a statistical life (VSL) estimates in recent empirical work.\(^3\) Aldy and

---

\(^3\)The value of a statistical life is simply the aggregation of a population’s willingness to pay for

In the next section, we outline the individual’s life cycle model. The third section discusses the calibration of the model and the numerical implementation. In the fourth section, we employ the model to assess our base case and provide results from these simulations. In the following section, we examine a variety of special cases involving variation in individual characteristics, such as income, gender, and race. The final section concludes the paper while an Appendix provides more information on the computational methods.

2.2 Model

Our model captures several important features that play a role in determining an individual’s VSL. First, the model allows idiosyncratic, age-specific mortality that a given risk reduction such that it equals the total amount that population would pay to reduce one statistical mortality. For example, if each member of a population of 10,000 is willing to pay $500 to reduce a 1 in 10,000 risk, then the value of a statistical life for this population is $5 million.
can differ across demographic groups. Second, agents face uninsurable wage shocks whose magnitudes can differ across different types of agents. Third, agents value both consumption and leisure. Fourth, agents choose optimally their savings and allocation of time to market work accounting for both earnings and mortality risk.

This type of life cycle model with idiosyncratic risk is common in the macro literature, for recent examples see French (2005), Nishiyama and Smetters (2005) or Smyth (2006a). These innovations to the typical life cycle model used in the mortality risk valuation literature allow us to conduct simulations that closely represent actual means and distributions for key economic measures over the life cycle, including labor income, consumption, wealth accumulation, labor participation, and mortality.

2.2.1 Household

Households choose their consumption and leisure to maximize the discounted sum of future utility:

$$\max_{c_t, l_t} \mathbb{E} \sum_{t=0}^{T} \beta^t \phi_t u(c_t, l_t)$$

(2.1)

Agents discount the future at a constant, time-invariant rate $\beta$. Their mortality differs by age with $\phi_t$ representing the age-specific probability of survival to the next period. $\phi_t$ can differ across agents of different type, for instance, across agents of different races. Especially for older agents, this age-dependent survival probability, which declines as agents get older, shortens their effective time horizon and affects decisions on consumption and leisure as well as their values for mortality risk reduction. Older agents behave as if they have substantially higher discount factors because of the risk of not surviving to the next period.

Agents face incomplete markets: no private markets provide insurance against future wage shocks. Agents can, however, self-insure through savings. All saving
is risk-less and pays the market rate of return. This ability to partially self insure, as well as conduct life cycle savings, provides a very important means of smoothing consumption in the face of idiosyncratic shocks. There are two uses of an agent’s resources: consumption and saving for future consumption. The agent’s resources consist of prior savings, \( a_{it} \), interest on those savings, \( R_{it}a_{it} \) and labor market compensation, \( w_{it} (1 - l_{it}) \). The agent allocates one unit of time between leisure, \( l_{it} \), and market work. Working in the market generates additional resources of \( w_{it} (1 - l_{it}) \) for consumption and savings, but with the opportunity cost of foregoing some leisure. This yields a standard budget constraint:

\[
c_{it} + a_{i,t+1} = a_{it} (1 + R_{it}) + w_{it} (1 - l_{it}) - \tau SS w_{it} (1 - l_{it})
\]

The base case is a borrowing constraint of zero. A borrowing constraint of zero ensures that no agents die in debt:

\[
a_{t+1} > a_{\min}
\]

\[
a_{\min} = 0
\]

We also consider cases where agents can borrow by relaxing the condition in equation (2.4). In these cases the age-dependent borrowing constraint starts fairly loose and rises to zero as the agent ages. The tightening borrowing constraint prevents agents from dying in debt with certainty.\(^6\)

\(^4\) Periods are one year.

\(^5\) If agents can borrow, some agents who die young will owe money. Setting a borrowing constraint of zero ensures that this case never happens and is an easy solution. A borrowing constraint of zero also helps in mimicking the observed distributions of consumption, income and wealth. In cases where we relax the borrowing constraint we ignore any agents who die in debt as their number is fairly small and offset by those who die holding assets.

\(^6\) If agents were allowed to borrow at older ages they would have incentives to borrow, consume
2.2.2 Wage Process

Mean wages, the life cycle pattern of wages and the idiosyncratic dispersion of wages over the agent’s lifetime all play key roles in generating the heterogeneous values for common changes in mortality risk. To capture the actual wage risks agents face, we use a wage process common in the macro-consumption literature that has both permanent and transitory shocks. Many studies estimate a similar process from microeconomic data (Sanzick and Carroll, 1997; Gourinchas and Parker, 2002; Storesletten, Telmer, and Yaron, 2004). In these studies, this process has been successful in matching several important empirical features of the U.S. economy, including wages, asset holdings, and consumption.

This wage process tracks the permanent component of wages and a transitory component both of which are subject to i.i.d. shocks. Together, these two determine the agent’s realized wage. The permanent component, $P_{it}$, is characterized by:

$$P_{it} = P_{i,t-1} \left( \frac{M_t}{M_{t-1}} \right) \eta_{it}$$ (2.5)

$M_t$ is the mean wage in period $t$ while $M_{t-1}$ was the mean wage the previous period. Their ratio, $\frac{M_t}{M_{t-1}}$, gives the deterministic element of wages over the life cycle. The permanent shock, $\eta_{it}$, persists into future periods through $P_{it}$ depending on the previous permanent component of wages, $P_{it-1}$.

The actual wage received by an agent is:

$$w_{it} = P_{it} \varepsilon_{it}$$ (2.6)

what they borrowed, and die in debt with certainty. Making agents pay back all debts before retirement ensures that this does not occur.
This is composed of the permanent component, \( P_{it} \), and a purely transitory shock, \( \varepsilon_{it} \), representing a shock without persistence. Its effects, except as they persist through differential asset holdings, disappear after a period.

All agents retire at 65. Although they can choose to retire (by working no hours) before age 65, they cannot receive Social Security benefits until they reach 65. The Social Security benefits the agent receives are based on a progressive non-linear transformation of the final permanent wage similar to the translation between wages and benefits in the actual system.\(^7\)

These features of Social Security are also important in driving the distributional dynamics of wealth. As shown by Hubbard, Skinner, and Zeldes (1995), Social Security and other government insurance programs drive those with low lifetime earnings to save little. This increases the dispersion of wealth in the economy. In addition, as numerous authors have pointed out, Social Security is a powerful risk-sharing tool.

### 2.2.3 Recursive Representation

Given the economic environment outlined above, we can write the individual’s problem recursively:

\[
V_t(A_{it}, P_{it}, \varepsilon_{it}) = \max_{c, l} u(c_{it}, l_{it}) + \beta \phi_t \mathbb{E}[V_{t+1}(A_{i,t+1}, P_{i,t+1}, \varepsilon_{i,t+1})]
\]

\(^7\)Social security wages are based on final permanent wage in our model rather than some form of AIME (Average Indexed Monthly Earnings). Basing benefits on an AIME-type formula would add an additional state variable to the agent’s problem. The increase in realism of such an addition was judged to not be worth the large increase in computational complexity.
subject to

\[ c_{it} + a_{i,t+1} = a_{it} (1 + R_t) + w_{it} (1 - l_{it}) - \tau w_{it} (1 - l_{it}) \]  

(2.8)

This recursive representation illustrates how solving the problem simplifies to solving a sequence of many one-period maximization problems. The state vector includes three variables: two continuous (the level of the permanent component of wages and the agent’s asset holdings) and one discrete (the transitory shock to income). Appendix B.1 presents details on our numerical methods for solving the individual’s problem.

2.2.4 Defining the Value of a Statistical Life

Using the above representation of an agent’s problem, we define a discrete time analogue to the more commonly used continuous time definition of mortality risk reduction valuation. Consider, at period \( t \), a small increase in the probability of surviving to the next period \( \partial \phi_t \). The probability of survival is now \( \phi_t + \partial \phi_t \) instead of \( \phi_t \). All future survival probabilities remain unchanged so the function \( \mathbb{E} [V_{t+1} (A_{i,t+1}, P_{i,t+1}, \varepsilon_{i,t+1})] \) also remains unchanged. Since agents change their savings behavior in response to the change in survival probabilities in that one period, the assets carried forward will be different and the evaluated expectation will also differ. This allows us to compare two distinct value functions representing agents’ utility maximization problem over consumption and leisure. \( V_t (A_{it}, P_{it}, \varepsilon_{it}) \) is the original or baseline value function and \( V_t^* (A_{it}, P_{it}, \varepsilon_{it}) \) reflects the case of the 1-period increase in agents’ survival probabilities. The change in the mortality risk the agents face represents the only difference in parameters.

To evaluate the value of the mortality risk reduction we consider how much an
agents would pay for this increase in their survival probability. Consider the change in wealth, $\partial A_{it}$, such that

$$V_t(A_{it} + \partial A_{it}, P_{it}, \varepsilon_{it}) = V_t^*(A_{it}, P_{it}, \varepsilon_{it})$$  \hspace{1cm} (2.9)$$

This increment $\partial A_{it}$ is the amount of additional wealth that makes the person indifferent between living longer in expectation and living the same length of time as before with that additional wealth. The value of a statistical life typically refers to the aggregate willingness to pay of a sub-population to reduce the mortality risk borne by this sub-population equivalent to one statistical fatality. We construct a VSL from this wealth-risk trade-off by assuming a set of individuals identical to an agent in a given model run and aggregating the $\partial A_{it}$ over this sub-population. This yields a VSL of:

$$VSL = \frac{\partial A_{it}}{\partial \phi_t}$$  \hspace{1cm} (2.10)$$

Note that since our population is heterogeneous in all of $A_{it}$, $P_{it}$ and $\varepsilon_{it}$ the VSLs are also heterogeneous across the population. We report VSLs throughout this paper to facilitate comparison with the literature by assuming that each agent represents a sub-population of individuals of size $\frac{1}{\partial \phi_t}$.

This corresponds to the definition of the VSL in simple one-period models and most revealed preference studies (Viscusi and Aldy, 2003; Johannesson, Johansson, and Lofgren, 1997). For example, hedonic wage models in which the wage is regressed on occupational job mortality risk and other controls yield VSL estimates by differentiating the estimated wage equation with respect to annual job mortality risk. The VSL is the change in the wage with respect to the change in risk and is given by the estimated coefficient for job mortality risk (for linear specifications and the coefficient scaled by the wage for log-linear specifications) aggregated over a
population. This definition corresponds to our one-period change in wealth divided
by the one-period increment in the survival probability expressed above.

2.3 Calibration

The parametrization of our model draws on a variety of sources. We employ some
parameters from the literature estimated through simulated method of moments
and other life cycle based microeconometric procedures. We estimate the parameters
characterizing the wage process with data from the Panel Study of Income Dynamics
(PSID) and we employ survival probability estimates from the 2002 U.S. Life Tables.

2.3.1 Agent’s Utility

Based on previous research by, among others, French (2005), Murphy and Topel
(2005) and Nishiyama and Smetters (2005) we assume that consumption and leisure
are non-separable. The period utility function is:

\[ u(c, l) = k + \frac{(cl^v)^{1-\gamma}}{1-\gamma} \] (2.11)

When \( v > 1 \), consumption and work are complements. At lower levels of leisure
the marginal utility of consumption increases. Agents who work a lot also consume
a lot. \( \gamma \) is the coefficient of relative risk aversion over the combined utility from
consumption and leisure. We assume \( \gamma = 2 \) and \( v = 2 \). A value of 2 for \( \gamma \) is
quite commonly used in the literature. A value of 2 for \( v \) implies that most working
agents spend about \( \frac{1}{3} \) of their time allotment on market work. These parameters
align closely with those estimated by French (2005) for this utility function.\(^8\) His

\(^8\)French defines the utility function slightly differently. He estimates the utility function
\[ u(c, l) = \left( (cl^v)^{1-\gamma} \right) \frac{1-\alpha}{1-\gamma}. \] In one specification he estimates that \( 1-\alpha = .602 \) and \( \gamma = 3.78. \)
simulated method of moments estimates of these preference parameters and this utility function better fits the empirical data than an alternative which is additively separable between consumption and leisure.

We include a constant $k$ in our utility function. Decisions regarding savings, consumption, and leisure are all invariant to including a constant as it effects no decisions and simply shifts the period utility up or down. Excluding a constant, however, yields negative values of utility for all positive values of consumption and leisure with a coefficient of risk aversion in excess of 1. This implies that the standard normalization that dead agents receive zero utility would result in agents preferring death to being alive.

Adding the constant to the utility function makes an agent prefer living to dying and affects how much each year of living is worth. Murphy and Topel (2005) employ a constant in their model as well. They define their constant as a ratio to a minimum level of consumption above which the agent prefers death, however this definition can be rewritten to match ours.

Transforming these to the terminology of our model implies that $v = 1.51$ and $\gamma = 1.50$.

Consider consumption of $20,000$ (denoted in tens of thousands of dollars) and the parameter values presented above. Assume that agents devote $\frac{2}{3}$ of their time to leisure. This yields a period utility, without a constant, of less than zero:

$$u_{\text{alive}} = \left( \frac{2}{\frac{2}{3}} \right)^{1-2} = \frac{9}{8}$$

Assuming that zero consumption is equivalent to death, this expression implies that the utility of being dead is:

$$u_{\text{dead}} = 0$$

This implies that agents prefer being dead to being alive. Assigning a sufficiently large negative number to the utility of death would achieve the same result as adding a positive constant to the utility function.

This is implicit in any model that has expected utility and mortality but does not explicitly model a value on death. The utility of death is typically normalized to zero, which is then averaged with all the other states. Since this does not affect agents’ decisions, the implicit assumption normally makes no substantive difference.

Murphy and Topel (2005) define $z = u(c, l)$ which is aggregate of consumption and leisure.
We choose $k$ to match the peak VSL estimated in empirical studies. Aldy and Viscusi (2006) have recently estimated workers’ life cycle VSL patterns based on labor market hedonic wage analyses over 1993-2000. They estimate a peak VSL of approximately $8 million and we calibrate $k$ to match that peak. Note that this calibration only implies that the peak estimated in our simulations will be about $8 million on average and has no bearing on the life cycle VSL pattern. The results that follow are robust to different values of $k$. Changing $k$ acts as a shifter to VSL, moving it up or down, but has little effect on either distributions or patterns of VSL.\(^{12}\)

They further define $z_0$ as the minimum level of utility for which the agent prefers life to death. Utility is then $u(z) = \frac{z_0^{1-\gamma} - z^{1-\gamma}}{1-\gamma}$ where pulling out the $z_0$ gives $u(z) = \frac{z_0^{1-\gamma} - z^{1-\gamma}}{1-\gamma}$. The second term is positive when $\gamma > 0$, as we assume and is commonly assumed. So our $k$ corresponds to Murphy and Topel’s $-\frac{1}{1-\gamma}$.

\(^{12}\)Life cycle VSL figures analogous to those presented in sections 4 and 5 for alternative measures of $k$ are available from the authors upon request.
2.3.2 Life Cycle Pattern and Variance of Wages

We estimate the life cycle pattern of labor market compensation from PSID data over 1967-1992. We limit the sample to heads of household between 20 and 65 years of age working at least 1,000 hours per year. We also exclude heads of household with top-coded income. This results in a total sample of 107,381 observations for 11,040 households. Labor market compensation measures are converted to constant year 2000 dollars using the CPI-Urban deflator. We then convert labor market incomes to hourly rates and take the natural logarithm of these values.

We estimate the deterministic life cycle component of labor market compensation by specifying the log of wages as a quintic function of age. This quintic version yields very similar life cycle labor market compensation profiles as more flexible specifications, such as those with age-specific dummy variables, but this more parsimonious parametric form allows for more efficient numerical simulation. Wages increase steeply for young agents between 20 and 30. Wages peak at around age fifty and decline thereafter. We have also estimated such age quintic specifications by gender and by race to complement the results for the total population and facilitate our analysis of the heterogeneity of the value of life across demographic characteristics.

To characterize the variance in labor market compensation as the effect of both permanent and transitory shocks, we employ the variance decomposition method developed by Samwick and Carroll (1997). For our sample, we first remove the predictable component of labor market compensation. We regress the log of the hourly wage on demographic characteristics, age, education, industry, occupation, interactions of these terms, and year dummy variables. We normalize a worker’s wage by constructing the ratio of the actual wage to the predicted wage from this regression. For each worker in a given year we difference labor market incomes by
Table 2.2: Permanent and Transitory Variances of Wage Shocks

<table>
<thead>
<tr>
<th></th>
<th>Permanent ($\sigma_\eta^2$)</th>
<th>Transitory ($\sigma_\varepsilon^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Sample</td>
<td>.0069722</td>
<td>.0263392</td>
</tr>
<tr>
<td>Males</td>
<td>.006681</td>
<td>.0282521</td>
</tr>
<tr>
<td>Females</td>
<td>.0070766</td>
<td>.0207519</td>
</tr>
<tr>
<td>Whites</td>
<td>.0067114</td>
<td>.0272783</td>
</tr>
<tr>
<td>Blacks</td>
<td>.0061862</td>
<td>.0260105</td>
</tr>
</tbody>
</table>

subtracting previous years’ incomes from the current year’s income:

$$\Delta y_{td} = y_{t+d} - y_t \quad \forall d = 1, \ldots, D \quad (2.12)$$

Having removed the predictable component of labor market income, the square of these differences should represent the variance in labor market income over time. Samwick and Carroll (1997) show how this can be decomposed into permanent and transitory shocks to income by OLS regression on a worker-by-worker basis of these differences on the time length associated with the difference, $d$, and the constant 2:

$$\Delta y_{td}^2 = d\sigma_\eta^2 + 2\sigma_\varepsilon^2 \quad (2.13)$$

The coefficient estimates yield the permanent and transitory income shocks. Following Samwick and Carroll, we exclude observations with $d$ of 1 or 2 to ensure that our results do not suffer from serial correlation. Table 2.2 contains the estimates of the variances for the total sample and various subgroups of interest.
2.3.3 Survival Probabilities

We use in our model survival probabilities from the 2002 U.S. Life Tables constructed by the National Center for Health Statistics (Arias, 2004). The Life Tables provide us with the age-conditional probability of surviving to the next year for the whole population, by male and female populations, and by black and white populations. Agents are born into our economic environment at age 20 and can live to be at most 100 at which age they die with certainty. Since less than 1% of agents survive to 100, truncating life at that age has no significant effects on the results. Also note that the U.S. Life Tables do not report age-specific survival probabilities for those older than 100.

2.3.4 Social Security Benefits Function

Our approximation to Social Security replaces income based on an agent’s final wage. It replaces a minimum of approximately $7,000 and increases for agents with higher wages until reaching the maximum of about $16,000. Consistent with the progressive nature of the statutory benefits structure, the replacement rate increases fastest near the bottom of the wage distribution and then levels off as wages increase.

2.4 Base Case

For all cases, we simulate the model with 10,000 agents. Agents in a given population start with the same initial labor market compensation at age 20, experience identical deterministic wage profiles, face the same survival probability profile over the life cycle, and have shocks drawn from the same distribution. For example, with the model calibrated to the characteristics of the total population (our base case), the 10,000 agents have the identical starting wage at age 20. We introduce the change
in the probability of survival for that one period and assess the change in wealth for the alternative value function with the initial probability of survival for each agent necessary to make agents indifferent between the two profiles. This generates the age 20 VSLs for the agents. We then move the agents forward to age 21, recognizing that only 99.9% of them survive to age 21, introduce labor market shocks, and conduct the risk-wealth change comparison again. This yields the age 21 VSLs for these agents. This proceeds on until age 100 when all agents die. These simulations yield age-specific VSLs for the agents in our model runs.
2.4.1 Entire Population

The agents in the base case with the total population calibrated model only differ over the life cycle as a result of the shocks they face in the labor market. Figure 2.1 shows the mean age-specific VSL profile for the agents in this base case. The mean VSL for these agents starts at just over $3 million at age 20 and increases to over double that, peaking at over $8 million in the mid-forties. After peaking, the VSL pattern steadily declines over the rest of the life cycle. This pattern is consistent with previous results in the literature. The inverted-U shape of the VSL reflects the findings in Shepard and Zeckhauser (1984) for their “Robinson Crusoe” case in which an agent cannot borrow against future income. The inability to insure against idiosyncratic income risk and the non-negative borrowing constraint in our model constrains agents in much the same way as the borrowing constraint in their Robinson Crusoe case. Our result is driven, however, much more by the inability to insure against income shocks than it is by the borrowing constraint – we have also undertaken simulations with looser borrowing constraints and we still find this inverted-U shape because risk-averse agents are not inclined to borrow substantially from future income under the possibility of future adverse income shocks. Our VSL shape also mimics the life cycle pattern of consumption, an analytic result of Johansson’s (2002) model and evident in empirical results in Kniesner, Viscusi, and Ziliak (2006).

The results also very closely mirror recent estimates of the VSL over the life cycle. On the same graph we have also plotted estimates of the VSL from Aldy and Viscusi (2006). Controlling for both age and cohort effects in panel data, they use a minimum distance estimator to flexibly allow for age variation in VSL. The only model parameter calibrated off of their estimates is the additive constant, \( k \), in the
utility function. All other model inputs come from different sources. So while we expect to match the peak VSL, there is nothing in our calibration that forces our model to match their estimates.

The model achieves results strikingly close to their estimates. The estimates rise and fall together extremely closely with similarities in both the shape of the VSL profile and level at the different ages not calibrated to match. The two curves track almost exactly from ages 25 to 55. At low ages our model underpredicts VSL by about a half million dollars. Since we ignore transfers from parents and bequests this slight underprediction is not especially worrisome. At older ages, our model slightly overpredicts relative to the empirical results. Though the reasons for this are less obvious, early retirement might play a role here as in reality old agents with low wages can drop out of the labor market. Since the Aldy and Viscusi estimates are only of the working population, this could pull their estimates down relative to the model.

In contrast to the deterministic simulation models used previously in this literature, our model with uncertainty in the income process allows us to investigate the distribution in VSLs. Figure 2.2 shows the life cycle VSL profiles for the mean, median, and 5th and 95th percentiles of the 10,000 sample of agents calibrated to the total population.\(^\text{13}\) This figure depicts two stark findings. First, the inverted-U shape is common to most agents across the distribution. The inverted-U takes an even more pronounced shape for those at the upper end of the distribution. For agents who experience a series of very negative shocks, however, the pattern of their VSL can exhibit a quite different pattern as it falls over their life cycle, evident by those at the 5th percentile of the distribution. Second, those agents who experi-

\(^{13}\) Percentiles are over cross-sections at each age. Using the cross-section gives a better picture of the distribution of VSL across the population at different ages.
ence a sequence of positive permanent labor market shocks have much, much higher VSLs. The mean VSL is nearly 33% higher at its peak than the median for this set of agents, reflecting the effect of the high VSLs on the mean. The VSLs for those at the 95th percentile are an order of magnitude larger than the VSLs for those at the 5th percentile. As the first analysis to account for uncertainty in the labor income process, we show substantial heterogeneity in the value individuals place on mortality risk reduction simply as a function of the labor compensation outcomes over their life cycle.

Figure 2.3 explicitly illustrates this heterogeneity in the VSL at one point in the life cycle, age 40. The distribution of agents’ values is very skewed. Although most agents’ mortality risk reduction valuation yields a VSL ranging between $2 million and $6 million, the long right tail shows the existence of much higher values.
This was the same effect that resulted in the divergence between the mean and the median: the relatively small number of high VSL agents pulls the mean above the median. The mean VSL of about $8 million at age 40 masks the variation over the population. This also has interesting implications, which we discuss below, for estimates based on compensating differential for mortality risk in revealed preference studies and mortality risk reduction valuations in stated preference studies.

Figure 2.4 presents the variance in the VSL estimates for the total population in our model. These age-specific variance estimates also follow an inverted-U with respect to age. The variance increases until around age 50, and then declines steadily until age 65. This initial decline in the variance reflects the decline in the mean wage.
Figure 2.4: Variance of VSL Over Life Cycle

profile at this stage of the life cycle. At retirement (mandated at 65 for all agents in our model), the variance declines precipitously. This reflects the progressive nature of social security – those agents who experienced bad labor market shocks enjoy relatively high social security benefits in comparison with their labor market income but do so now with full leisure time. At the top of the income distribution, these agents now move from making high incomes from the labor market to relatively modest social security benefits. Agents also start to consume their accumulated wealth holdings. Since this is also when mortality rates begin rising rapidly, the effective discount factor of agents also rises and they run down their asset holdings. Since only the top of the income distribution has substantial amounts of wealth to run down, VSLs decline much more for those at the top of the income distribution
than for those at the bottom, resulting in the smaller variance starting at 65.\textsuperscript{14}

### 2.4.2 Without Income Shocks

Figure 2.5 shows the effects of labor market shocks on the life cycle VSL pattern by presenting the mean values for the deterministic pattern (no shocks), the VSL shape with only permanent labor market shocks (notrans), the shape with only transitory shocks (noperm), and the full model presented previously in Figure 2.1. The permanent shocks to labor market compensation drive the results for the full model, as the mean values for the full model and the no transitory shocks version track each other very closely. The deterministic model and the no permanent shocks version also track each other very closely. Since transitory shocks have only a very small effect on lifetime income and can be almost entirely self insured against by saving, their small effect on VSL is not surprising. In contrast, permanent shocks lead to large differences in lifetime income and can influence consumption and leisure choices. Their inclusion thus induces large changes in VSLs.

In contrast, the median VSL is actually higher when we turn shocks off. As before, both the no shocks and the only transitory shocks cases move almost exactly together while the full shocks and the only permanent cases do the same. The cases with more shocks now exhibit a uniformly lower VSL than before. They also peak slightly earlier in the life cycle and the post-peak decline as the agent ages is less steep. The size of the discrepancy between the cases with permanent shocks and cases without also shrinks when considering median VSL rather than mean VSL.

The variance in income, consumption, and leisure induced by the idiosyncratic wage shocks makes most agents less well off. The lower median VSLs for agents in

\textsuperscript{14}This effect is due in large part to the presence of Social Security. Poor agents have little incentive to save since Social Security benefits replace a large fraction of their consumption.
Figure 2.5: Impact of Income Shocks on Mean and Median VSL
Table 2.3: Lifetime Value and Value of Statistical Life
Numbers in the table above are the value function and VSL evaluated at the mean assets and wages across the 10,000 simulations. Agents are aged 40 years and consequently are near the peaks of their VSL. The VSL row is in millions while the $E(V)$ row has no units.

<table>
<thead>
<tr>
<th></th>
<th>Base</th>
<th>No Shocks</th>
<th>No Permanent Shocks</th>
<th>No Transitory Shocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E(V)$</td>
<td>136.57</td>
<td>135.19</td>
<td>135.61</td>
<td>136.27</td>
</tr>
<tr>
<td>VSL</td>
<td>6.60</td>
<td>5.56</td>
<td>5.69</td>
<td>6.25</td>
</tr>
</tbody>
</table>

the presence of permanent shocks reflects this. The median is not influenced by the skewed distribution of VSLs whereas the mean is largely driven by that.

Table 2.3 presents the lifetime remaining value for agents with the same assets and permanent wages with and without shocks to labor compensation. Shocks uniformly reduce both the lifetime expected value and the VSL conditional on a given level of assets and wages. The increase in the mean VSL instead stems from the dispersion in assets and wages. Agents who have higher assets and wages are willing to pay more to ameliorate risks.

### 2.5 Individual Characteristics

#### 2.5.1 Income

Individuals should be willing to pay more for a given mortality risk reduction as their incomes increase. Viscusi and Aldy (2003) estimate an income elasticity for the value of statistical life on the order of 0.5-0.6. In contrast, Kaplow (2005) argues that the income elasticity should exceed 1.0 based on theoretical grounds and empirical estimates of the coefficient of relative risk aversion. Consistent with this, Costa and Kahn (2004) estimate an income elasticity of about 1.5 for a series of U.S labor market cross-sections over 1950-1990. With our model we can explicitly address the
Table 2.4: Elasticity of VSL

Regressions are run on simulated data. There are 408,002 observations as regressions are run for all working age (under 65) agents. The coefficients in the first two rows of the table report the elasticities of VSL with respect to realized income and the permanent component of income respectively. \(a\) is the asset holdings of an agent. Some cases control for age using a cubic. All coefficients are significant in these cases. All regressions also contain a constant which is not reported.

<table>
<thead>
<tr>
<th>(\bar{w}(1-l))</th>
<th>(A)</th>
<th>(B)</th>
<th>(C)</th>
<th>(D)</th>
<th>(E)</th>
<th>(F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\bar{w})</td>
<td>.554</td>
<td>.523</td>
<td>.552</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.0006)</td>
<td>(.0006)</td>
<td>(.0006)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\bar{a})</td>
<td>1.709</td>
<td>1.656</td>
<td>1.672</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.0007)</td>
<td>(.0007)</td>
<td>(.0006)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>other controls</td>
<td>.106</td>
<td>.041</td>
<td>.1708</td>
<td>.100</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.0004)</td>
<td>(.0002)</td>
<td>(.0008)</td>
<td>(.0003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.674</td>
<td>0.935</td>
<td>0.7142</td>
<td>0.941</td>
<td>0.756</td>
<td>0.9619</td>
</tr>
</tbody>
</table>

question of how the value of life varies with income.

We have estimated income elasticities in separate regressions of the natural logarithm of VSL on the natural logarithm of the permanent component of wages and the natural logarithm of realized income. The two measures differ because agents will choose to adjust their time allotment to leisure and labor as labor compensation changes: a percentage increase in the permanent component of wage will typically exceed the percentage increase in realized income because workers will increase their leisure. Since accumulated wealth also plays a role in determining the willingness to pay for mortality reductions, we also control for it in some regressions. The simulated data from our model runs generate more than 400,000 observations for the all working agents sample and at least 9,200 observations for the age-specific working agents samples.

Table 2.4 shows our calculations of the elasticity of VSL with respect to both
the permanent component of wages and the actual income of agents. All estimates generate statistically significant positive elasticities for VSL with respect to either realized income or permanent wages. Running a simple regression without controls yields economically and statistically significant different elasticities for VSL with respect to realized income and the idiosyncratic productivity. When using income, the elasticity of VSL is 0.554 but when using productivity this jumps to 1.709. The productivity measure also does a much better job predicting the value an agent places on these mortality risk reductions with an $R^2$ of 0.935 as opposed to 0.674. The regressions with productivity are thus doing a better job of approximating the true function that relates our state variables to VSL.\(^{15}\)

Columns (C), (D), (E) and (F) add a variety of controls. Since assets and age should also affect the observed VSL we add controls for these. For assets we simply add the level of assets while for age, given the patterns we observed in the previous section, we allow non-linearities with a cubic in age. Adding assets pulls down the estimate of elasticity with respect to either measure slightly. Adding a cubic control for age further brings the estimates back up slightly to 0.552 and 1.672 for income and productivity respectively. These controls, however, leave intact the large difference between the two measures. Using our quasi-statistical procedure, these estimates are different from each other at far above the 1% level. Further, the differences are large in terms of their economic meaning since the VSL with respect to productivity is three times as large.

\(^{15}\)Theoretically, VSL is a four dimensional function of age, assets, the permanent level of productivity and the transitory shock, \(\text{VSL}_{it} = \text{VSL}_{iit}(\text{age}, a_{it}, P_{it}, \varepsilon_{it})\). This function can generally be non-linear. In running the regression we are estimating it as a linear function of these state variables. Errors in the linear regression are thus not random error but deviations of the true underlying function from the linear approximation. Adding more controls thus allows us to better approximate the function. The age specific regressions we perform later are then separately estimating the series of functions \(\text{VSL}_{i} = \text{VSL}_{ii}(a_{i}, P_{i}, \varepsilon_{i})\) where since each is estimated for a different age that is no longer a state variables.
Table 2.5: Elasticity of VSL at Different Ages

Regressions are run on simulated data. Each coefficient reported in this table represents the results for a separate regression. Income and productivity are not included in the same regressions. For the cases with all agents there are 408,006 observations. For other cases the number of observations ranges from 9,200 observations to 9,997 observations. Standard errors are in parentheses below the reported elasticities. The $R^2$ for the “All Working” are presented in Table 2.4 while those for the permanent component of wages range from $0.993 - 0.998$ and those for income from $0.775 - .967$.

<table>
<thead>
<tr>
<th></th>
<th>All Working</th>
<th>Age 30</th>
<th>Age 45</th>
<th>Age 60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elasticity of VSL w/r Permanent Component of Wages</td>
<td>1.672 (.0003)</td>
<td>2.011 (0.001)</td>
<td>1.562 (.002)</td>
<td>.814 (.003)</td>
</tr>
<tr>
<td>Elasticity of VSL w/r Realized Income</td>
<td>0.552 (.0006)</td>
<td>.819 (0.005)</td>
<td>.555 (.004)</td>
<td>.147 (.002)</td>
</tr>
<tr>
<td>Controls</td>
<td>Wealth, Age Cubic</td>
<td>Wealth</td>
<td>Wealth</td>
<td>Wealth</td>
</tr>
</tbody>
</table>

Further, as shown in Table 2.5, the elasticities exhibit substantial age variation. Young workers have a much higher income elasticities for both measures than do old workers. For example, an age-30 worker has an estimated VSL-income elasticity for the permanent component of wages of 2.01, and this declines to 1.56 for an age-45 worker and still further to 0.81 for an age-60 worker. Likewise, an age-30 worker has a VSL-income elasticity for realized income of 0.82, nearly twice the age-45 estimate of 0.56 and nearly five times the age-60 estimate of 0.15. For young agents the VSL varies much more with changes in either permanent wages or income than for old agents. Since they have many more years of possibly higher consumption (and/or leisure) from their higher income, they should be willing to pay more to maintain this. Young agents also have the ability to stretch each additional dollar out over more years. Since marginal utility of both consumption and leisure declines with more of either, the same amount of extra consumption spread out over a longer time period raises expected utility more. Thus lifetime expected utility rises by more for a given change in the probability of dying. Agents value the mortality risk reduction
more and thus pay more for it. These economic drives explain why an extra dollar of either permanent wages or income has a larger effect on the VSL for young agents than old agents.\footnote{Assets have the opposite pattern. A extra dollar a of assets has a large impact on the VSL for old agents than young ones. Since older agents don’t have future labor income which is affected by changes in their permanent component of wages, their main difference is their asset holdings. Higher asset holdings support a higher future level of consumption making a given mortality change that increase the chances of living to enjoy that consumption more valuable.}

The differences between the elasticities with respect to income and permanent wages is also very interesting. The elasticity with respect to income is closer in spirit to the empirical work that looks at the elasticity of VSL with respect to income in cross-sectional data. This elasticity takes into account the fact that agents with extremely high wages or assets can cut back on hours of market work and take more leisure. Since these cross-sectional studies generally do not take into account leisure in the analysis, they can understate how much a larger wage makes agents willing to pay for extra life years. It is important to recognize that a model that does not account for both leisure choice and income shocks (that can generate VSL heterogeneity) would not be able to illustrate this effect.

In contrast, the estimates based on the permanent component of wages are more closely akin to estimates that vary across time. As agents become richer in terms of their lifetime earning potential they value their life more. To a degree, the permanent component of wages captures that agents have some stock of future earnings potential that they can take in either consumption or leisure. Since agents value both, an extra dollar of earnings potential is very valuable to the agents.

These differences in the VSL elasticity measures for the realized income and the permanent component of wages provides some numerical and theoretical context for the existing literature. Costa and Kahn (2004) found an income elasticity with
a series of U.S. cross-sections over 40 years indistinguishable to our estimate for the permanent component of wages VSL elasticity. Likewise, Kaplow (2005) and Murphy and Topel (2005) suggest that the income elasticity should exceed unity. Across cohorts, we should expect such an elastic response of VSL to income changes. A variety of meta-analyses of labor market VSL estimates – summarized, replicated, and extended upon by Viscusi and Aldy (2003) – tend to find much lower VSL income elasticities. Most of the studies included in these meta-analyses are based on labor market survey data over a relatively short period of time from the mid-1970s through late-1980s. The variation in income in these studies’ samples likely reflects differences within cohorts – such as from sample composition effects by focusing on specific occupations, specific countries, etc. – and less differences across cohorts. The smaller estimated income elasticities by integrating these VSL results in a meta-analysis would appear to square with the findings for our assessment of how VSL changes with realized income. In fact, our VSL elasticity for this measure is indistinguishable from the estimated VSL income elasticity in Viscusi and Aldy (2003).

To show how the elasticity with respect to permanent wages corresponds more closely to the VSL rising with wages rising over time we consider the following experiment. Increasing the initial wage of agents entering the economy is analogous to the overall rise in wages that occurs over time. An increase in the entering wage persists permanently and increases the permanent component of wages for every agent in the economy. Given this secular increase in permanent incomes, VSLs across the population should also rise and we can consider the elasticity of this with respect to the increase in permanent income.

This approach yields an estimate of the VSL-income elasticity of 1.56.\textsuperscript{17} This is

\textsuperscript{17}Formally the procedure is very similar to finding the cross-sectional elasticities above. We consider increasing the initial wage by 10\%, 20\%, 50\%, 75\%, 125\% and 200\%. We then take the log of the population average of VSL. We then run the regression of this on the log of the initial...
effectively the same as our estimate of the elasticity with respect to the permanent component of wages presented in Table 2.5. This elasticity of VSL in response to a secular increase in wages for our model economy supports our interpretation above of a difference between measuring an elasticity with respect to wages and income. Since the elasticity with respect to income fails to take into account the utility from leisure, it understates how much more agents are willing to pay for reductions in mortality risk as their incomes increase.

2.5.2 Race

Recent research by Viscusi (2003) based on hedonic wage models shows that black workers have a value of life equal to about one-half that of white workers. The component of permanent wages. The above estimate is the coefficient on initial permanent wages.
discrepancy is even larger when focused on black and white male workers. We assess
the extent to which the difference between black and white willingness to pay for
risk reduction reflects variation in income profiles and expected longevity.

Figure 2.6 depicts the mean VSLs for each of three sets of 10,000 agents for
the model calibrated to the total population, the white population, and the black
population. The white and black populations of agents have VSL patterns that
follow similar inverted-U shapes over the life cycle. The modest difference in the
initial (age 20) VSLs between blacks and whites reflects the longer longevity and
higher expected future incomes for whites. White VSLs peak at over $9 million in
their late forties, nearly double the peak for blacks at a similar age, comparable to
Viscusi’s findings.

To decompose the effects of differences in income profiles and longevity, we
present four mean life cycle VSL patterns in Figure 2.7. Two of the patterns are

the black and white life cycle VSL profiles presented in Figure 2.6. The other two

reflect a modification of a black population characteristic to its white population

representation. In one profile, we have modified the black population calibrated

model so that it is has the white population’s deterministic wage profile while main-

taining the black labor market income variances and survival probabilities. In the

other profile, we have modified the black population calibrated model so that it has

the white population’s deterministic wage profile and labor market income variances

while maintaining the black survival probabilities.

Figure 2.7 clearly shows that the vast majority of the difference in VSL between

blacks and whites results from their differences in wages. Improving the survival

rates of black agents to the levels of white agents only yields a small increase in the

mean VSL, closing less than ten percent of the gap between peaks. The very modest

effect of changing survival probabilities at early stages of life reflects the effect of
discoun ting. A 20-year old worker expecting to live to 73 (black life expectancy

conditional on reaching 20) or 78 (white life expectancy) would not value these

differences much since they would occur more than 50 years in the future. The

figure does show that moving from black survival probabilities to white survival

probabilities does increase in absolute and percentage terms the VSL for those at

older ages because they have fewer years over which to discount before enjoying

the longer life expectancy. Improving the black wage profile to match the white

wage profile closes over ninety percent of the gap between the VSLs. The difference

in wages between blacks and whites drives the differences in VSL that have been

observed in the empirical literature.
2.5.3 Gender

Men and women differ in both their labor market compensation (mean and variances) and their life expectancy. Based on the 2002 Life Tables, women should expect to live 5.5 years longer than men. This should increase the VSL of women relative to men. Countering this effect, however, is the higher labor market compensation men experience. We assess the life cycle VSLs for men and women to discern which of these two effects dominate.

Figure 2.8 depicts the mean VSLs for each of three sets of 10,000 agents for the model calibrated to the total population, the male population, and the female population. Both male and female populations have VSL patterns that follow similar inverted-U shapes over the life cycle. Female VSLs peak in their late forties, about five years later than the male VSL peak, reflecting the effects of differences in life
Figure 2.9: Breakdown of Sources of Differences in VSL between Males and Females expectancy. At their peaks, men have VSLs of over $ million, more than 33% greater than women’s VSLs of $5.2 million at their peak.

To decompose the effects of differences in income profiles and longevity, we present four mean life cycle VSL patterns in Figure 2.9. Two of the patterns are the male and female life cycle VSL profiles presented in Figure 2.9. The other two reflect a modification of a female population characteristic to its male population representation. In one profile, we have modified the female population calibrated model so that it is has the male population’s deterministic wage profile while maintaining the female labor market income variances and survival probabilities. In the other profile, we have modified the female population calibrated model so that it has the male population’s deterministic wage profile and labor market income variances while maintaining the female survival probabilities.
The figure illustrates that the vast majority of the differences in VSL between males and females, as in the black-white comparison, derives from differing wage processes. Giving females the shorter lives of males results in only a modest reduction in their VSL. Since wages for females increase at a slower rate than those of males, their VSL also rises less steeply and has a longer plateau than the male VSL. Giving females the wages of males, however, makes a substantial difference. With their longer life expectancies, their VSL now surpasses that of males at all ages. The mortality difference makes little difference in the shape with VSLs as males’ and females’ VSLs have almost the exact same shape once they have the same pattern of wages.

2.6 Conclusion

We have developed a numerical life cycle model that is novel to the mortality risk reduction literature by accounting for choice over consumption and leisure and stochastic wage and mortality processes. Our calibrated model matches several important regularities in the U.S. economy, including life cycle consumption patterns, precautionary saving among young adults, wealth accumulation starting in middle age, and labor market compensation and participation outcomes. The realism of the model and the accounting for uncertainty allows us to explore important questions on the value of mortality risk reductions that existing models cannot address.

We find that the value agents’ place on reductions in mortality risk (or their population-equivalent VSL) varies substantially with income and assets. Accounting for empirically estimated life cycle wage patterns and idiosyncratic labor market compensation shocks yields a wide range of VSL estimates both over the life cycle and across the population. The mean, median, 95th percentile, and variances of the VSL follow an inverted-U shape over the life cycle, peaking in agents’ mid-40s.
at about the same point in the life cycle as the peak in consumption. Agents who experience very negative labor market shocks have a VSL that declines over their life cycle, as evident by the 5th percentile of the distribution. The mean VSL substantially exceeds the median as the VSL distribution is non-normal with an extremely fat, long upper end tail. The peak VSLs for agents at the 95th percentile are an order of magnitude larger than the peak VSLs for those at the 5th percentile of the distribution. The inverted-U shape is driven by the idiosyncratic labor market shocks that we assume, consistent with real-world practice, are uninsurable (except through savings-based self insurance). Accounting for these labor market shocks, extending beyond the current VSL literature, allows us to characterize the entire age-specific VSL distribution and not just a single, deterministic series of values.

The elasticity of the VSL with respect to income differs substantially between the wage and realized income. Estimating the elasticity with respect to realized income ignores the substantial value that agents place on leisure. We find that the VSL-income elasticity for all working agents based on the permanent component of the wage is on the order of 1.5, three times greater than the estimate for realized income. We believe that this difference, only evident in such a model with leisure choice and a stochastic labor market compensation process over the life cycle, explains the difference in VSL-income elasticities in the revealed preference literature: the permanent wage results are consistent with cohort-based income elasticities and the realized income results are consistent with cross-sectional results. We also show that the VSL-income elasticity declines significantly with age for both income measures.

We have also assessed how the VSL varies over the life cycle by race and gender. We find that differences in the life cycle wage process and mortality profile drive the substantial differences in VSLs between blacks and whites and males and females. Assuming no preference heterogeneity across demographic groups, we show
that virtually all of the black-white gap in recent revealed preference research can be explained by labor market compensation. Although the longer female life expectancy should increase female VSLs relative to males' VSLs, the higher labor market compensation for men more than offsets this effect.

Our results may have several implications for future research. First, the dispersed and skewed distribution of agents' VSLs indicates that empirical estimates based on assumptions of normality might be biased. Second, these results suggest the possibility of explaining some other observed risk anomalies based on the agents' position in wage and asset distributions. Third, our estimates of an elasticity in response to changes in wages substantially in excess of one implies that as an economy becomes richer more resources might be devoted to health care and other measures to reduce mortality risk. Fourth, some have suggested that age-specific discount rates and risk attitudes could explain some of the empirical literature's findings of an age-invariant VSL (or in the case of Smith, Evans, Kim, and Taylor (2004), a VSL that increases with age). A model similar to ours could be extended to explore the necessary evolution over the life cycle of time and risk preferences to yield such different life cycle VSL patterns.

The substantial heterogeneity in the VSL over the life cycle and across the population could have several implications for the application of VSL estimates to mortality risk reduction policy proposals. First, the life cycle variation in the value of statistical life in this study does not support current EPA practice of a constant VSL irrespective of age or the past FDA practice of a constant value of a statistical life-year irrespective of age. The inverted-U life cycle pattern we find implies that EPA practice overestimates the benefits of mortality risk reduction for the elderly and underestimates the benefits of risk reduction for prime-aged adults and FDA practice overestimates the benefits of mortality risk reduction for young adults. Sec-
ond, the VSL-income elasticity results can inform assessments of long-term mortality risk reduction policies, such as reducing UV-B exposure through the phase-out of chlorofluorocarbons. Such policies would reduce mortality risk for future generations that would be expected to have higher lifetime incomes and would be willing to pay more for risk reduction. Third, the heterogeneity across the population suggests that a one-size fits all VSL would not appropriately reflect how much individuals would be willing to pay for risk reduction if they had such an opportunity to do so in a competitive market. For proposed policies that may affect the mortality risk profile for a specific demographic group, population-specific VSLs may be appropriate. Given the negative response to the so-called “senior discount” used by EPA in its 2002 evaluation of the Clear Skies Initiative, this may raise political and ethical concerns. Such concerns may be entirely legitimate for those contexts in which at least some of the differences in lifetime incomes and wage profiles reflect gender- or ethnic-specific discrimination in labor markets.
Chapter 3

Optimal Consumption Taxes
3.1 Introduction

European countries tend to have a tax system that raises much more revenue from consumption taxation than the United States. Countries such as the UK, France, Germany, Canada and many others raise a substantial amount of revenue from consumption, rather than income, taxation. In contrast, the United States does not use consumption taxes on a national level, though states often fund themselves through consumption taxes.\(^1\) Instead, the United States uses an income tax to raise the majority of government revenue.

The debate on consumption taxes has tended to focus on linear versions because they are more commonly implemented and are easier to implement. There is no reason, however, why consumption taxes must be constrained to be linear. The fundamental feature of a consumption tax is the base that the tax is applied to, not the linearity. Progressive taxes can be applied to consumption as well as income.

European nations impart some progressivity to consumption taxation by taxing different goods at different rates. Necessities, such as food, are often taxed at lower rates than normal good. Luxuries, say yachts, often are taxed at higher rates. A positive correlation between the percent of consumption spent on necessities and income combined with a negative correlation between percent of consumption spent luxuries then makes these different rates of taxation partially proxy for a progressive consumption tax.

A more direct way of implementing a progressive tax on consumption would involve tracking income from all sources and contributions to to any form of saving. Subtracting these two then gives a direct measure of consumption. Progressive taxes could then be levied on this measure. This would, however, raise questions of what

\(^1\)State consumption taxes are normally referred to as sales taxes.
counts as savings - particularly pertinently with housing.

Independent of implementability, enough theoretical interest has focused on the efficiency aspects of linear consumption taxes that studying their effects, taking into account risk sharing and redistributive aspects of taxation, is interesting. Dating back to Auerbach and Kotlikoff (1987) and beyond, a robust finding has been that consumption taxes have positive efficiency effects - increasing the capital stock, output and consumption. Does a progressive consumption tax maintain these beneficial effects of consumption taxation? Does a progressive undo some of the negative distributional consequences of a flat consumption tax, and if so does that overturn the increases in output and consumption.

This paper examines these questions. In an OLG model with uninsurable idiosyncratic risk, I calculate the optimal non-linear tax on consumption. This optimal tax is very progressive. Though it raises the capital stock, it actually lowers aggregate output compared to the current income tax system. The lower output stems from the sharp drop in labor supply of the high productivity individuals induced by the high taxes. Despite this drop in earnings, all individuals actually save more as capital income is untaxed.

Recent work by Nishiyama and Smetters (2005) examined the costs and benefits of transition to a linear consumption tax system. Other recent work by Smyth (2006a) and Conesa and Krueger (2005) has studied the optimality of non-linear taxation in OLG models. Finally, a recent and growing strand of the literature started by Golosov, Kocherlakota, and Tsyvinski (2003) studies the problems involved in private information on the part of agents, problems that are implicit in these other works.

The rest of the paper is structured as follows: Section 3.2 lays out the OLG model used to evaluate the possible tax functions; Section 3.3 presents the optimum and
discusses its impacts both over the life cycle of an agent and upon macroeconomic aggregates; Section 3.4 compares the optimal tax system to a linear consumption tax along these same lines while Section 3.5 concludes.

3.2 Model and Calibration

This section presents the OLG model in which the optimality of various consumption taxes is assessed. Several major features define the model: (1) agents face idiosyncratic wage risk, (2) agents face a potentially non-linear tax on income or consumption, (3) general equilibrium linkages of individual decisions to aggregate quantities, (4) a deterministic hump shaped component of wages and (5) a Social Security program. Smyth (2006a) contains a much more detailed description of the model.

3.2.1 Household

Rational, forward looking households choose their consumption and leisure to maximize their expected discounted sum of future utility. Agents discount the future at the time invariant rate $\beta$ and have mortality probabilities that vary with age of $\phi_t$. This implies I can write their maximization problem as:

$$\max_{c,l} E \sum_{t=0}^{T} \beta^t \phi_t u(c, l) \quad (3.1)$$

As I pursue numerical solutions to the agents problem, I have to specialize and choose a specific utility function. Utility is non-separable in consumption and leisure based on evidence presented in French (2005):
\[ u(c, l) = \left( c^\gamma \right)^{1-\gamma} \frac{1}{1-\gamma} \] (3.2)

Agents accumulate wealth from their prior savings which pay the risk-less, market rate of return and through their work in the labor market. From those resources they pay taxes, which the government levies on income and consumption. There are two uses for this after tax accumulation, consumption and saving for next period. This yields a budget constraint of:

\[ c_t + a_{t+1} = a_t (1 + R_t) + w_t (1 - l_t) - \tau_c (c_t) - \tau_i (R_t a_t + w_t (1 - l_t)) - \tau_{ss} (w_t (1 - l_t)) \] (3.3)

where \( \tau_{SS} \) is a linear tax on wage income and \( \tau_i (\cdot) \) and \( \tau_c (\cdot) \) are flexible non-linear tax functions applied to income and consumption. We further assume incomplete markets in that agents cannot insure against idiosyncratic wage shocks and that agents are unable to borrow against their risky future income. This adds the constraints:

\[ a_{t+1} > a_{\text{min}} \] (3.4)
\[ a_{\text{min}} = 0 \]

Finally, wages are stochastic. The agents faces two different shocks, a permanent one, \( \eta_t \), and a transitory one, \( \varepsilon_t \). The permanent component of wages, \( P_t \), persists through time and carries forward the effects of \( \eta_t \), while growing at a deterministic rate of \( \left( \frac{M_t}{M_{t-1}} \right) \) where \( M_t \) is the mean wage of all age \( t \) agents. \( M_t \) follows a hump shape rising up til the agents mid-fifties and slightly declining thereafter.
\[ P_{it} = P_{i,t-1} \left( \frac{M_t}{M_{t-1}} \right) \eta_{it} \]  

(3.5)

\[ w_{it} = w_{it}^e P_{it} \varepsilon_{it} \]  

(3.6)

Appendix C.1 contains details on estimation of this wage process from Panel Study of Income Dynamics (PSID) data.

This implies the agent’s problem can be written in the common recursive form of

\[ V_t (A_{it}, P_{it}, \varepsilon_{it}) = \max_{c,l} u(c_{it}, l_{it}) + \beta \phi_t E \left[ V_{t+1} (A_{i,t+1}, P_{i,t+1}, \varepsilon_{i,t+1}) \right] \]  

(3.7)

Writing the problem in this form makes solving the agent’s life cycle problem easier as it reduces to a series of one period maximization problems. Appendix C.2 contains details on the numerical solution procedure used to solve the problem for the agent.

### 3.2.2 Firm

A representative firm produces output from capital and labor. A standard Cobb-Douglas production function transforms capital and labor into output.

\[ Y_t = A_t F(K_t, N_t) = A_t K_t^\alpha N_{1-t} \]  

(3.8)

Labor and capital are the aggregation of all individual capital holdings and labor inputs. Labor varies with the productivity of the individual and the amount of time they choose to devote to market work. Aggregating these across all individuals in the economy gives:
\[ K_t = \sum_{i=1}^{I} a_{it} \quad (3.9) \]

\[ N_t = \sum_{i=1}^{I} P_{it} \varepsilon_{it} (1 - l_{it}) \quad (3.10) \]

With perfect competition, the wages and interest rates in the economy are given by by the marginal products of capital and labor respectively:

\[ w_t^e = A_t (1 - \alpha) \left( \frac{K_t}{N_T} \right)^\alpha \quad (3.11) \]
\[ R_t = A_t \alpha \left( \frac{N_t}{K_t} \right)^{1-\alpha} - \delta \quad (3.12) \]

where \( \delta \) is the depreciation rate in the economy. The various parameters governing production are standard from the literature and summarized, along with sources, in Table 3.1.

### 3.2.3 Government

The government runs a balanced budget and pays for exogenously given government expenditures. I assume that government spending is separable in utility from agents own consumption to alleviate the need to specify the utility of government consumption.

The governments imposes non-linear taxes on either consumption or income to pay for these expenditures. Tax functions take the form:

\[ \tau_i(y) = a_0 \left( y - (y^{-a_1} + a_2)^{-\frac{1}{a_1}} \right) \quad (3.13) \]
Gouveia and Strauss (1994) introduced this approximation of the tax function. They, Smyth (2006a), Nishiyama and Smetters (2005) and Conesa and Krueger (2005) all discuss its properties more extensively. Given appropriate choices of \( a_0 \), \( a_1 \) and \( a_2 \) the function approximates progressive taxes, exemption based taxes and linear taxes. Government spending, financed either by the income or consumption tax, is calibrated from the average of federal government tax collections, excluding social insurance, from 1995-2004 divided by GDP.

### 3.2.4 Steady State of Model

The model is closed when individual decisions are consistent with aggregate quantities. So, given wages, interest rates and tax functions the savings and labor supply decisions of the household imply aggregate amounts of capital, labor and tax revenue that generate these wages and prices. Appendix C.3 contains a formal definition of

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Source</th>
<th>Parameter Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta )</td>
<td>0.965</td>
<td>Gourinchas and Parker (2002); French (2005)</td>
<td>Rate of time preference</td>
</tr>
<tr>
<td>( \gamma )</td>
<td>2.0</td>
<td>Gourinchas and Parker (2002); French (2005)</td>
<td>Risk aversion</td>
</tr>
<tr>
<td>( \nu )</td>
<td>2.0</td>
<td>French (2005); Nishiyama and Smetters (2005)</td>
<td>Trade-off between leisure and consumption</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>( \frac{1}{3} )</td>
<td>Standard</td>
<td>Capital share in output</td>
</tr>
<tr>
<td>( \delta )</td>
<td>0.06</td>
<td>Nishiyama and Smetters (2005)</td>
<td>Depreciation rate</td>
</tr>
<tr>
<td>( \phi_i )</td>
<td>varies</td>
<td>Bell and Miller (2002)</td>
<td>Survival probability</td>
</tr>
<tr>
<td>income process</td>
<td>varies</td>
<td>Author’s estimation</td>
<td>Earnings pattern over life cycle</td>
</tr>
</tbody>
</table>
equilibrium for the model and details on the numerical solution methods.

3.2.5 Social Welfare

To consider optimal taxation requires defining the optimum. The model maximizes the utility of a benevolent social planner who cares about all agents equally.

\[ U_{SP} = \sum_{t=0}^{\infty} \beta_t^{SP} \left( \sum_i u(c_i, l_i) \right) \]  

(3.14)

Since this paper compares between steady states, the discounting of the social planner is irrelevant and we can consider maximizing one cross-section of agents utilities. Smyth (2006a) contains a fuller discussion of the social planner as well as discussion of an alternative Rawlsian social welfare criterion.

3.3 Optimal Consumption Taxes

This section presents the optimal non-linear tax on consumption. First it shows the form of the optimal consumption tax and how and how taxes paid varies with the amount consumed. The later subsections discuss the macroeconomic impacts of a system of consumption tax relative to an income tax system like the United States currently has and how a consumption tax shifts various quantities of interest to the agent across the life cycle.

The model compares the steady state under the optimal consumption tax to an approximation of the current United States income tax system. Using a progressive tax function estimated by Nishiyama and Smetters (2005), we calculate the steady state of the model. Smyth (2006a) discusses how the model matches the data.

I then optimize over the set of possible consumption taxes spanned by the tax function. For consumption taxation, this involves optimizing over two parameters
Figure 3.1: Optimal Tax on Consumption
The optimal tax on consumption. Note that even though marginal rates exceed 100% an agent might still wish to consume above that level. Since the budget constraint is $a_{t+1} = w(1-l) - c - \tau_c(c)$. A tax of greater than 100% on consumption just means an agent has to earn two dollars to finance one of consumption.

free parameters of the Gouveia and Strauss (1994) tax function given in Equation 3.13. The function has three parameters but the third is tied down by matching the total amount of government spending.

3.3.1 Form of Optimal Consumption Tax

Figure 3.1 shows the optimal consumption tax. Taxation on consumption up to around $20,000 is quite light. The marginal rates start at zero and gradually increase over to slightly over 10%. Above $20,000 marginal rates start to rise rapidly and continue rising up rapidly up to around $100,000 where they are nearly 200%. Above this they level off and slowly approach their maximum of 212%.

It is important to be clear what these marginal rates mean, since the intuition
behind them is quite different than the normal intuition regarding marginal tax rates from income or labor taxation. Recall our definition of the agents budget constraint from Equation 3.3:

\[ c_t + a_{t+1} = a_t (1 + R_t) + w_t (1 - l_t) - \tau_c (c_t) - \tau_i (R_t a_t + w_t (1 - l_t)) - \tau_{ss} (w_t (1 - l_t)) \]  

(3.15)

Now consider consuming an extra dollar. For example, take an agent paying the 100% marginal tax rate on consumption. Consuming the extra dollar raises the individuals tax liability by a dollar. So, to increase consumption by a dollar takes two dollars of resources, one to pay for the consumption and one to pay for the taxes on it. The 100% marginal rate does not have the intuition that it prevents any consumption from occurring above that level. Similarly, the top marginal rate of just above 200% means agents need a little more than three dollars to fund an additional dollar of consumption. These rates are high, but do not rule out any consumption occurring in the area where marginal rates are over 100%.

The effect of these rising consumption taxes is to encourage consumption smoothing. For agents who expect their consumption to be higher in the future, the progressive consumption tax encourages consuming now. The price of current consumption is lower than the price of future consumption so these agents consume now instead of saving. For agents who expect a decline in consumption, the consumption tax encourages savings. Since consuming in the future will be cheaper than consuming now they reduce their current consumption and save to increase their future consumption. For agents who expect their consumption to rise and then fall, as young agents generally do, the net effect is ambiguous.

Overall, the optimal tax on consumption is very progressive. Marginal tax rates start low and continue to rise throughout the income distribution only topping out
Figure 3.2: Mean Consumption

This figure depicts mean consumption across the life cycle for agents under two tax regimes. The one labeled “Current” refers to a progressive income tax estimated from the current U.S. system. The line “Optimal Cons Tax” is if consumption was taxed using the optimal rates presented in Figure 3.1.

at a quite high level, over $200,000 in consumption. Average rates at higher incomes are also quite high. At the same $200,000 the average rate is slightly over 100%. So consuming $200,000 costs the agent a bit over $400,000 in resources.

3.3.2 Life Cycle Effects

Compared to the current income tax system, changing to the optimal consumption tax raises consumption of the young and lowers that of the middle aged. Figure 3.2 graphically depicts how the mean consumption of agents compares to consumption under an income tax system. Young agents can consume slightly more than before as taxes on them are slightly reduced. Middle aged and older working agents bear the brunt of the higher consumption taxation. Their consumption is pushed down
Figure 3.3: Variance of Consumption

This figure depicts the cross-sectional variance of consumption across the life cycle for agents under two tax regimes. The one labeled “Current” refers to a progressive income tax estimated from the current U.S. system. The line “Optimal Cons Tax” is if consumption was taxed using the optimal rates presented in Figure 3.1.

significantly. At its peak, consumption under the progressive consumption tax is about 10% lower than under the income tax.

Given that total aggregate consumption is now lower than before, where do the aggregate gains in utility come from. Figure 3.3, which depicts the cross-sectional variance of consumption across the life cycle, provides insight into this. The variance of consumption is much lower under the progressive consumption tax. Since the social planner values all agents equally and since marginal utility declines in consumption, the social planner dislikes variance. The drastic reduction in variance helps raise aggregate welfare above what it was before and offset the slightly lower mean level of consumption.
Figure 3.4: Life Cycle Pattern of Mean Taxes Paid

This figure depicts average taxes paid to the government across the life cycle for agents under two tax regimes. The one labeled “Current” refers to a progressive income tax estimated from the current U.S. system. The line “Optimal Cons Tax” is if consumption was taxed using the optimal rates presented in Figure 3.1.
In addition, the young agents who get to consume slightly more, as discussed above, are agents who have low consumption and thus a very high marginal utility of consumption. Figure 3.4 shows the average taxes that an agent pays under the current income tax system and the optimal consumption tax. Young agents pay lower taxes under the optimal consumption tax allowing them to increase their consumption above what they consume under the income tax system. The biggest differences, however, occur later in life. In prime working age the average taxes paid decreases significantly. Agents aged around 40 years pay significantly less under consumption taxes. The average tax take drops by 20%. However, in contrast to income taxes, the consumption taxes paid by the agent continues to rise - with taxes paid peaking in the late fifties as opposed to around forty for the income taxes.

Taxes collected from from agents who are about to retire are much higher than under the income tax system. Instead of tax revenue from these agents falling sharply, it continues to rise peaking around fifty, rather than in the agent’s mid-forties. Theoretical results of Erosa and Gerivas (2002) and Garriga (2003) indicate that with age specific taxes the planner would like to be able to tax these agents near retirement more heavily, but is unable to do so since they have a high elasticity of labor supply. The numerical results of Smyth (2006a) mirror this as the planner uses non-linear capital taxes to increase the total tax burden on these agents near retirement. Under the income tax system, they cut back on the amount of work they do and lower their income taxes. Under the consumption tax, however, they pay more since they are unwilling to cut back too much on consumption. Since they still consume at a fairly high level the planner can raises tax revenue from them using the consumption tax. Under an income tax, cutting back labor supplied to the market and living of savings until they claim Social Security at 65 allows them to substantially lower their taxes. Under a consumption tax doing so no longer lowers
their tax burden.

Retired agents also pay significantly higher taxes under consumption taxes. For newly retired agents, their average tax burden is almost double what they pay under the current income tax system. Since they are still consuming at a fairly high level immediately following retirement, their tax burden does not drop as much as under the labor income tax where their taxes paid decreases dramatically since they no longer have labor income. The consumption of older retired agents declines as they run down their assets, so their tax bill also declines.

Overall the optimal consumption tax smooths the distribution of the tax burden across agents of different ages much more. The peak of the taxation burden is lower and there are many fewer years where there is a very low burden. Also, those years where there is a low burden no coincides better with years when there is relatively low utility, and hence a high marginal utility. Likewise, the heaviest burden lies on years when utility is high and marginal utility is low. Perhaps most pronounced, and leading to some of the positive efficiency aspects, is the ability of the consumption tax gives to compel agents right before and after retirement to continue paying substantial taxes. These are agents who have some of the highest average utility so taxing them and lower taxes elsewhere raises social welfare.

3.3.3 Macroeconomic Impacts

These life cycle effects lead to significant macroeconomic effects. Table 3.2 shows how some major macroeconomic aggregates differ between the income tax system and the optimal consumption tax system. These are all steady state comparisons so they do not take into account all the costs associated with changing between tax systems.

Importantly, since it was the goal of optimization over the consumption tax
Table 3.2: Comparison of Initial and Optimal Steady States
Table comparing certain key model statistics between the initial tax system and the optimal consumption tax. This tables compares steady state to steady state. Note that $V_0$ and $\sum u(c, l)$ are measures of utility. More is better so the larger value, which is closer to zero, is better.

<table>
<thead>
<tr>
<th></th>
<th>Initial Income Tax</th>
<th>Optimal Consumption Tax</th>
<th>Linear Consumption Tax</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{\Delta}{N}$</td>
<td>4.12</td>
<td>4.77</td>
<td>4.72</td>
</tr>
<tr>
<td>normalized $Y$</td>
<td>1.0</td>
<td>.99</td>
<td>1.11</td>
</tr>
<tr>
<td>normalized $K$</td>
<td>1.0</td>
<td>1.09</td>
<td>1.22</td>
</tr>
<tr>
<td>normalized $N$</td>
<td>1.0</td>
<td>0.94</td>
<td>1.06</td>
</tr>
<tr>
<td>$w$</td>
<td>1.0</td>
<td>1.05</td>
<td>1.04</td>
</tr>
<tr>
<td>$R$</td>
<td>6.7</td>
<td>5.5</td>
<td>5.7</td>
</tr>
</tbody>
</table>

function, the aggregate utility in the economy increases. The normalized sum of utilities increases by 2% so social welfare is significantly higher.\(^2\) Measured by the Rawlsian criteria of the welfare of a newborn agent, utility also improves. The increase is even larger than the increase in overall utility. Much of this has to do with a large part of the tax burden being shifted onto older agents. Since the young agents discounts this both because of pure time preference and the risk of not surviving to see it, they are much better off. In fact, to make a newborn as well off under the income tax system as they would be under the consumption tax system would require a transfer of $15,000.

Despite these positive welfare effects, aggregate output decreases slightly. Steady...

\(^2\)Since this is a measure in terms of utility, we cannot directly make comparisons to consumption or other quantities of interest. This is further complicated the sum contains many agents at different levels of utility for who extra consumption or leisure are worth different amounts. The number still gives an idea of the general direction and magnitude. The declining marginal utility of consumption implies this will be greater than a 2% in consumption equivalents.
state output is 1% below what it is under the income tax system. The main driver of the fall in output is a large drop in the aggregate labor supply, which falls by over 5%. The fall comes from high productivity agents cutting back their supply of labor. The high marginal tax rates on consumption for these high productivity individuals mean that they face high marginal tax rates no matter when they consume. These agents current and future consumption is going to be high enough that they cannot smooth consumption enough to get low rates. If they try to consume high amounts now, they face high marginal tax rates. If they try to save for the future and consume then, they will still face the high marginal rates. Leisure, on the other hand, faces no such high marginal tax rates. Thus they cut back on their hours of work and take more leisure pushing down the aggregate labor supply.

Concurrently, the capital stock increases by a large amount with associated positive effects. The aggregate amount of capital increases by 9% from its level under the income tax. Combined with the decrease in labor supply, this pushes up the capital-labor ratio by over 16%. The new, higher, capital-labor ratio pushes down the interest rate and pulls up the wage level. Both of these changes help the young and low productivity agents at the expense of the old and high productivity agents. The lower interest rate does not hurt the lower end of the productivity distribution since their optimal response to the presence of Social Security is to save little (Hubbard, Skinner, and Zeldes, 1995). Meanwhile, they get the majority of their income through wage income, so rising wages helps them a lot. High marginal rates on high levels of consumption affects these agents less. They never reach the levels of consumption where the high marginal rates of the optimal consumption taxes reduce their consumption or significantly alter their decisions on labor supply because of anticipated high future marginal tax rates on consumption. Thus they benefit greatly and have little downside.
These general equilibrium effects operate through what I call a “labor shallowing” effect - akin to the more commonly noted capital deepening. By pushing down the labor supply of high productivity workers, the progressive taxes help raise the capital-labor ratio, $K/N$. Instead of operating through raising $K$, as conventional capital deepening does, “labor shallowing” works by decreasing the denominator $N$. The effects are the same: raising wages and pulling down the interest rate. Both of these changes help those at the bottom of the income distribution.

3.4 Efficiency and Redistribution Compared to Linear Consumption Taxes

To assess the efficiency and redistribution of the optimal consumption taxes compared to other consumption taxes we compare the optimal taxes on consumption to another more common and more easily implementable tax on consumption - a linear consumption tax.

The linear consumption tax has important benefits from simplicity. A linear consumption tax would be extremely easy to collect. Since it’s a flat levy on consumption, all that has to be done is collect it from all transactions. Most states in the United States already have some form of linear sales tax so the mechanism is already in place. In addition, there are few opportunities for avoidance of a straight linear consumption tax since all agents are paying at the same rate and shifting around measured consumption between agents would not change taxes paid.

3.4.1 Linear Consumption Tax

A linear consumption tax yields steady state social welfare that is slightly lower than under the optimal consumption tax, but higher than under the current income tax.
Figure 3.5: Life Cycle Patterns - Linear Consumption Tax Compared to Linear Consumption Tax

This figure depicts various cross sectional quantities across the life cycle for agents under two tax regimes. “Linear Cons Tax” refers to all consumption being taxed at a constant linear rate. The line “Optimal Cons Tax” is if consumption was taxed using the optimal rates presented in Figure 3.1.
Figure 3.5 shows differences between the linear consumption tax and the optimal non-linear one found earlier. Panel (a) shows that average consumption is much higher under a linear tax than under the optimal tax. Reducing the marginal rates on consumption taxation for many individuals, and, as importantly, never having it reach very high levels leads to greatly increased work incentives. High productivity agents are much more willing to work and save knowing that they will be able to consume high levels without paying extremely high levels of taxes. This increases labor supply - as shown in Table 3.2 - and hence also average earnings and consumption.

Panel (b) shows that this increase in average consumption comes with a cost. The variance of consumption is much higher under the flat consumption tax system than under the optimal. Since high productivity agents work and consume more, their levels of consumption are now much higher than the poorer agents. This also has the effect, the reverse of the “labor shallowing” noted above, of pushing up the denominator of the capital-labor ratio which then brings down wages. Despite lower levels of both capital and labor, again shown in Table 3.2, the capital labor ratio is higher under the optimal tax system. The results of this on an individual level can be seen in Panel (d) where average wealth holdings are higher for agents when they face linear consumption taxes. Along with these effects on capital and labor, aggregate output is substantially - 12% - higher under the linear consumption tax system.

Under a linear tax system average taxes, presented in Panel (c), follow a much more pronounced hump. The optimal consumption tax system places a greater share of the lifetime tax burden on agents around middle age. The young and the old, who previously faced low average tax rates, now face higher ones. The young and old agents generally have a higher marginal utility of consumption - due to their lower
consumption - so taxing them less heavily and allowing them to raise consumption raises aggregate welfare.

The optimal versus the linear consumption tax systems display how the trade-off for equity vs. efficiency plays out. The linear consumption taxes cause substantially increases in output. In the end, either system gives a substantial gain over the current income tax system and close to the same benefit to a benevolent to a social planner. Using another criteria for social welfare, the Rawlsian veil of ignorance where we consider the lifetime expected future utility of a newborn agent, there are substantial benefits to being born into the more progressive economy. Using that measure, the optimal consumption tax system significantly outperforms the linear one.

These results stands in partial contrast to recent results by Nishiyama and Smetters (2005). They find, accounting for the transition, that welfare is lower under a flat consumption tax than under the current income tax system. This difference comes from two main sources. First, their model compensates all current agents for losses along the transition. Having to compensate the existing agents exacts a resource cost which is borne into the future. Paying this cost causes some losses for future agents. Since this paper does not account for the costs of the transition, these costs do not show up. Second, they measure the utility of agents entering the economy. This is closer to the Rawlsian criteria discussed above. When we examine a Rawlsian criteria, the gains from a consumption tax, compared to the current income tax system, are much lower.
3.5 Conclusion

Optimal consumption taxes are very progressive. Imposing high marginal rates, they greatly reduce the variance of consumption. Despite a large increase in the capital stock, aggregate output drops slightly compared to a benchmark progressive income tax system. Labor supplied drops enough to cause this output drop. This combination pushes the capital-labor ratio much higher raising wages and lowering interest rates. The optimal consumption tax also does a good job of taxing agents who have high utility, those who will soon retire. Taxing them more heavily and taxing other agents less raises social welfare.

The linear consumption tax provides social welfare very close to what the planner can achieve by the optimal. The channels through which the linear consumption tax increases social welfare are completely different. Instead of being quite redistributive and reducing variance, the linear tax increases aggregate output and works through increasing the size of the pie rather than dividing it more equally. Though the linear consumption tax gives good results for a social planner who cares about all equally, it does less well from a Rawlsian perspective.

Given these results, the planner faces a choice. Either of two consumption taxes gives quite good results in terms of social welfare. Presented, however, is a clear efficiency-equity trade-off. The linear consumption tax substantially raises aggregate output through its positive effects on the efficiency of taxation. It comes at the cost, however, of reducing the amount of risk sharing in the tax code. The optimal tax system provides a lot or risk sharing, and good matching of the life cycle timing of taxation to high consumption periods, at the cost of distortions that lower efficiency.

An important take away from this paper and Smyth (2006a) is that the optimal response of the planner, when the set of taxes become more efficient, is to increase
the progressivity in the rest of the tax code. The improved efficiency of taxation makes the planner able to indulge its desire for redistribution without hurting theoretical results in simpler models, where the channels through which the mechanisms described above operate are clearer, would help to illuminate and provide context to this numerical finding.
Bibliography


Appendix A

Appendices to Chapter 1

A.1 Solving the Life Cycle Problem

Solving the individuals life cycle problem requires numerical methods as analytical solutions do not exist. Recall that there are three state variables (1) the amount of assets (2) the permanent component of wages and (3) the transitory shock to wages. Since the government tax functions are dependent on the levels of these state variables we cannot make some of the useful simplifications that are often made\(^1\). Lacking these simplifications, we solve the more complicated problem with many state variables.

Recall that we recursively represented the agent’s problem as:

\[
V_t = \max_{c,t} u(c,t) + \beta \phi_t E[V_{t+1}(A_{t+1}, w_{t+1}, \varepsilon_{t+1})]
\]  
(A.1)

\(^1\)An example is the common normalization of using the homothecy of the value function to reduce the state variable to \(\frac{A_t}{P_t}\), the ratio of assets plus permanent income to permanent income. Since the level of both assets and income affects tax rates and hence individual decisions we lose this convenient normalization.
subject to:

\[ c_t + a_{t+1} = a_t (1 + R_t) + w_t (1 - l_t) \]
\[ \quad - \tau_k (R_t a_t) - \tau_l (w_t (1 - l_t)) - \tau_{ss} (w_t (1 - l_t)) \]
\[ \quad - \tau_c (c_t) - \tau_i (R_t a_t + w_t (1 - l_t)) \]

For each value of the discrete state, the transitory wage shock, there is a grid over assets and wages. Solutions generally use 3 transitory shocks. The grid over assets and wages has 12 points for both assets and wages. Adding more grid points does not significantly change any of the decisions or aggregate quantities of the model. Since the decision rules feature high levels of curvature of low levels of assets and wages and much lower curvature at higher levels we use a non-equally spaced grid. The density of the grid increases at lower levels of both wages and assets to better account for the rapid changes in the value function.

Shape preserving splines approximate the value function between grid points. The value function inherits monotonicity the underlying utility function. Agents always prefer more assets and higher wages. This implies that approximation methods which preserve this are preferred. Two dimensional Akima shape preserving splines are relatively fast and efficient (Schneider and Eberly, 2003). Importantly these splines provide information not only about the function but also its derivative. This is necessary to use the much faster Newton’s method based optimization algorithms described below.

To solve the individuals problem at each of these grid points we use the OPT++ library of Mesa (1994). This package solves general non-linear programming problems with arbitrary constraints.

Starting in the last period the solution is simple, do not work and consume
everything. We then know the the value function $V_{80}$ at our grid points. The splines described above then approximate the value function over the entire area. Given $V_{80}$, we then solve for $V_{79}$ taking $V_{80}$ as given and construct splines again. From this the algorithm iterates back to the beginning of an agent's economic life at age 0.

In cases where the interest rates, wages and tax functions change over the life of the agent the solution is exactly the same except for facing different prices in different years.

When solving for the optimum in the value function we also get decision rules for consumption and leisure. Since these are again on the grid the same spline procedure approximates them.

### A.2 Steady State Solution

#### A.2.1 Steady State Equilibrium Definition

When given a sequence of government spending decisions, $\{G_t\}_{t=0}^{\infty}$ and set of tax policy variables, $\{a_0^l, a_1^l, a_2^l, a_0^k, a_1^k, a_2^k, \text{ss}(P), \tau_{SS}\}_{t=0}^{\infty}$ a competitive equilibrium is a set of value functions and decision rules for the agent, $\{V_t, c_t, l_t, a_{t+1}\}_{t=1}^{T}$, stationary measure of agents $\{\Gamma\}$, and prices $\{R_t, w_t^e\}_{t=0}^{\infty}$ such that:

1. Households maximize utility. Agents solve their recursive problem, equation C.1 subject to equations 3.3 and 3.4 given the sequence of wages and interest rates $\{w_t^e, R_t\}_{t=0}^{\infty}$.

2. Markets clear so that:

$$K_t = \sum_{i=1}^{I} a_{it} \tag{A.3}$$
and
\[ N_t = \sum_{i=1}^{I} P_{it} \xi_{it} (1 - l_{it}) \]  \hspace{1cm} (A.4)

3. Perfect competition gives factor prices.

\[ w_t^f = A_t (1 - \alpha) \left( \frac{K_t}{N_T} \right)^\alpha \]  \hspace{1cm} (A.5)

\[ R_t = A_t \alpha \left( \frac{N_t}{K_t} \right)^{1-\alpha} - \delta \]  \hspace{1cm} (A.6)

4. Aggregate feasibility so:

\[ C_T + (1 - \delta) K_t + G_t \leq F(K_t, L_t) \]

5. The government balances both its Social Security and general budget

\[ G_t = \sum_{i=1}^{I} l_{it} \]  \hspace{1cm} (A.7)

and

\[ \sum_{i=1}^{I} s_{it}^{bens} = \sum_{i=1}^{I} \tau_{ss} (P_{it} \xi_{it} (1 - l_{it})) \]

6. The measure of agents of is stationary. Let \( \Omega \) be the transition function between measures of agents. Then:

\[ \Gamma_{t+1} = \Omega_t (\Gamma_t) = \Gamma_t \]

A.2.2 Steady State Numerical Solution

Models with overlapping generations, incomplete markets and uninsurable idiosyncratic risk are generally not easy to solve. Finding the steady state requires both
numerical solution of the Bellman Equation and iteration over the endogenously determined prices and government policies

The algorithm can be summarized as follows:

1. Pick an initial guess for the capital labor ratio and the government policy functions.

2. Recursively solve the agent’s problem.

3. Simulate a large number of agent histories and calculate aggregate statistics.

4. Update our estimates of $\frac{K}{N}$ and government policies. We move $\frac{K}{N}$ toward that which comes from the above simulations. We move elements of the tax functions toward Social Security taxes balancing with benefits paid and the total revenue raised by the government being the target percentage of GDP.

5. Check if the algorithm converged. If yes, then stop. If no, then go to 2.

More details on the computational procedure are available from the author on request.

A.3 Transition Path Solution

A.3.1 Transition Path Equilibrium Definition

A similar definition of equilibrium holds along the transition path. Instead of repeating all equations, the definition we simply summarize: the definition in C.3.1 now holds period by period along the transition path.
A.3.2 Transition Path Numerical Solution

Our algorithm is similar to Nishiyama and Smetters (2005) and Conesa and Krueger (1999, 2005). These papers describe the solution in more detail. We here just present a sketch of the algorithm.

The solution for the transition path is similar to the steady state solution except that every generation alive must be solved for and differing prices along the transition path allowed. The algorithm can be summarized as follows:

1. Pick an initial guess for the sequence of capital labor ratios and government policy functions.

2. Recursively solve the agent’s problem for each generation of agents that will be alive at any point along the transition.

3. Simulate a large number of agent histories and calculate aggregate statistics.

4. Update our estimates of \( \frac{K}{N} \) and government policies for all periods. We move \( \frac{K}{N} \) toward that which comes from the above simulations. We move elements of the tax functions toward Social Security taxes balancing with benefits paid and the total revenue raised by the government being the target percentage of GDP.

5. Check if the algorithm converged. If yes, then stop. If no, then go to 2.

More details on the computational procedure are available from the author on request.
A.4 Estimation of Wage Process

The wage process is estimated from PSID data. We use household labor income where top coded values are dropped. Nominal values are converted to real using the CPI-Urban deflator. Incomes are divided by hours worked and then logged.

The sample is restricted to be from 1973 to 1992. Heads of household with missing observations are dropped leaving us with 931 heads of household to estimate over.

We estimate the deterministic component of the wages with respect to age with a quintic in age. In this regression we also control for race, gender, region and occupation.

The variances of the permanent and transitory shock are calculated using the methodology of Samwick and Carroll (1997).

Aldy and Smyth (2006) provides more information on all the estimation procedures.
Appendix B

Appendices to Chapter 2

B.1 Solving the Life-Cycle Problem

Solving the individual’s life-cycle problem requires numerical methods as analytical solutions do not exist. Recall that there are three state variables (1) the amount of assets (2) the permanent component of wages and (3) the transitory shock to wages.

The particular form of our problem makes it impractical to make some common dimensionality reducing state reductions\(^1\). Lacking these simplifications, we solve the more complicated problem with many state variables.

Recall that we recursively represented the agent’s problem as:

\[
V_t = \max_{c, l} u(c, l) + \beta \phi_t E[V_{t+1}(A_{t+1}, w_{t+1}, \varepsilon_{t+1})]
\]  

(B.1)

\(^1\)An example is the common normalization of using the homothecity of the value function to reduce the state variable to \(\frac{A_t}{w_t}\), the ratio of assets plus permanent income to permanent income. Since we have both a permanent and transitory shock combined with an endogenous labor supply decision we lose this convenient normalization.
subject to:

\[ c_t + a_{t+1} = a_t (1 + R_t) + w_t (1 - l_t) - \tau_{SS} w_{it} (1 - l_{it}) \]  

(B.2)

For each value of the discrete state, the transitory wage shock, there is a grid over assets and wages. Solutions generally use 3 transitory shocks. The grid over assets and wages has 20 points for both assets and wages. Adding more grid points does not significantly change any of the decisions or aggregate quantities of the model. Since the decision rules feature high levels of curvature at low levels of assets and wages and much lower curvature at higher levels we use a non-equally spaced grid. The density of the grid increases at lower levels of both wages and assets to better account for the rapid changes in the value function.

Shape preserving splines approximate the value function between grid points. The value function inherits the monotonicity underlying utility function. Agents always prefer more assets and higher wages. This implies that approximation methods which preserve this are preferred. Two dimensional Akima shape preserving splines are relatively fast and efficient (Schneider and Eberly, 2003). Importantly these splines provide information not only about the function but also its derivative. This is necessary to use the much faster Newton’s method based optimization algorithms described below.

To solve the individuals problem at each of these grid points we use the OPT++ library of Mesa (1994). This package solves general non-linear programming problems with arbitrary constraints.

Starting in the last period the solution is simple, do not work and consume everything. We then know the the value function \( V_{80} \) (corresponding to age-100 agents) at our grid points. The splines described above approximate the value function over
the entire area. Given $V_{80}$, we then solve for $V_{79}$ taking $V_{80}$ as given and construct splines again. From this the algorithm iterates back to the beginning of an agents economic life at age 0 (age-20 agents).

When solving for the optimum in the value function we also get decision rules for consumption and leisure. Since these are again on the grid the same spline procedure approximates them.
Appendices to Chapter 3

C.1 Estimation of Wage Process

The wage process is estimated from PSID data. We use household labor income where top coded values are dropped. Nominal values are converted to real using the CPI-Urban deflator. Incomes are divided by hours worked and then logged.

The sample is restricted to be from 1973 to 1992. Heads of household with missing observations are dropped leaving us with 931 heads of household to estimate over.

We estimate the deterministic component of the wages with respect to age with a quintic in age. In this regression we also control for race, gender, region and occupation.

The variances of the permanent and transitory shock are calculated using the methodology of Samwick and Carroll (1997).

Aldy and Smyth (2006) provides more information on all the estimation procedures.
C.2 Solving the Life Cycle Problem

Solving the individuals life cycle problem requires numerical methods as analytical solutions do not exist. Recall that there are three state variables (1) the amount of assets (2) the permanent component of wages and (3) the transitory shock to wages. Since the government tax functions are dependent on the levels of these state variables we can not make some of the useful simplifications that are often made\footnote{An example is the common normalization of using the homothecity of the value function to reduce the state variable to $\frac{A + \Pi}{P}$, the ratio of assets plus permanent income to permanent income. Since the level of both assets and income effects tax rates and hence individual decisions we lose this convenient normalization.}. Lacking these simplifications, we solve the more complicated problem with many state variables.

Recall that we recursively represented the agent’s problem as:

$$V_t = \max_{c, l} u(c, l) + \beta \phi_t E[V_{t+1}(A_{t+1}, \omega_{t+1}, \varepsilon_{t+1})]$$  \hspace{1cm} (C.1)$$

subject to:

$$c_t + a_{t+1} = a_t (1 + R_t) + \omega_t (1 - l_t)$$  \hspace{1cm} (C.2)

$$-\tau_k (R_t a_t) - \tau_l (\omega_t (1 - l_t)) - \tau_{ss} (\omega_t (1 - l_t))$$

$$-\tau_c (c_t) - \tau_l (R_t a_t + \omega_t (1 - l_t))$$

For each value of the discrete state, the transitory wage shock, there is a grid over assets and wages. Solutions generally use 3 transitory shocks. The grid over assets and wages has 12 points for both assets and wages. Adding more grid points does not significantly change any of the decisions or aggregate quantities of the model. Since the decision rules feature high levels of curvature of low levels of assets and
wages and much lower curvature at higher levels we use a non-equally spaced grid. The density of the grid increases at lower levels of both wages and assets to better account for the rapid changes in the value function.

Shape preserving splines approximate the value function between grid points. The value function inherits monotonicity the underlying utility function. Agents always prefer more assets and higher wages. This implies that approximation methods which preserve this are preferred. Two dimensional Akima shape preserving splines are relatively fast and efficient (Schneider and Eberly, 2003). Importantly these splines provide information not only about the function but also its derivative. This is necessary to use the much faster Newton’s method based optimization algorithms described below.

To solve the individuals problem at each of these grid points we use the OPT++ library of Mesa (1994). This package solves general non-linear programming problems with arbitrary constraints.

Starting in the last period the solution is simple, do not work and consume everything. We then know the the value function $V_{80}$ at our grid points. The splines described above then approximate the value function over the entire area. Given $V_{80}$, we then solve for $V_{79}$ taking $V_{80}$ as given and construct splines again. From this the algorithm iterates back to the beginning of an agent’s economic life at age 0.

In cases where the interest rates, wages and tax functions change over the life of the agent the solution is exactly the same except for facing different prices in different years.

When solving for the optimum in the value function we also get decision rules for consumption and leisure. Since these are again on the grid the same spline procedure approximates them.
C.3 Steady State Solution

C.3.1 Steady State Equilibrium Definition

When given a sequence of government spending decisions, $\{G_t\}_{t=0}^{\infty}$ and set of tax policy variables, $\{a^l_0, a^l_1, a^l_2, a^k_0, a^k_1, ss(P), \tau_{SS}\}_{t=0}^{\infty}$ a competitive equilibrium is a set of value functions and decision rules for the agent, $\{V_t, c_t, l_t, a_{t+1}\}_{t=1}^{T}$, stationary measure of agents $\{\Gamma\}$, and prices $\{R_t, w_t\}_{t=0}^{\infty}$ such that:

1. Households maximize utility. Agents solve their recursive problem, equation C.1 subject to equations 3.3 and 3.4 given the sequence of wages and interest rates $\{w_t, R_t\}_{t=0}^{\infty}$.

2. Markets clear so that:

$$K_t = \sum_{i=1}^{I} a_{it} \quad \text{(C.3)}$$

and

$$N_t = \sum_{i=1}^{I} P_{it} \xi_{it} (1 - l_{it}) \quad \text{(C.4)}$$

3. Perfect competition gives factor prices.

$$w_t^e = A_t (1 - \alpha) \left( \frac{K_t}{N_T} \right)^{\alpha} \quad \text{(C.5)}$$

$$R_t = A_t \alpha \left( \frac{N_t}{K_t} \right)^{1-\alpha} - \delta \quad \text{(C.6)}$$

4. Aggregate feasibility so:

$$C_T + (1 - \delta) K_t + G_t \leq F(K_t, L_t)$$

143
5. The government balances both its Social Security and general budget

\[ G_t = \sum_{i=1}^{I} l_{it} \]  
\[ \sum_{i=1}^{I} s_{ss}^{bens} = \sum_{i=1}^{I} \tau_{ss} (P_{it} \xi_{it} (1 - l_{it})) \]

6. The measure of agents of is stationary. Let \( \Omega \) be the transition function between measures of agents. Then:

\[ \Gamma_{t+1} = \Omega_t (\Gamma_t) = \Gamma_t \]

### C.3.2 Steady State Numerical Solution

Models with overlapping generations, incomplete markets and uninsurable idiosyncratic risk are generally not easy to solve. Finding the steady state requires both numerical solution of the Bellman Equation and iteration over the endogenously determined prices and government policies.

The algorithm can be summarized as follows:

1. Pick an initial guess for the capital labor ratio and the government policy functions.
2. Recursively solve the agent’s problem.
3. Simulate a large number of agent histories and calculate aggregate statistics.
4. Update our estimates of \( \frac{K}{N} \) and government policies. We move \( \frac{K}{N} \) toward that which comes from the above simulations. We move elements of the tax
functions toward Social Security taxes balancing with benefits paid and the total revenue raised by the government being the target percentage of GDP.

5. Check if the algorithm converged. If yes, then stop. If no, then go to 2.

More details on the computational procedure are available from the author on request.