

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Optimal taxation in life cycle models

Permalink

<https://escholarship.org/uc/item/4gw0z0qz>

Author

Peterman, William Ben

Publication Date

2011

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Optimal Taxation in Life Cycle Models

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Economics

by

William Ben Peterman

Committee in charge:

Professor Irina A. Telyukova, Chair
Professor Roger Gordon
Professor Isaac W. Martin
Professor Valerie A. Ramey
Professor Krislert Samphantharak

2011

Copyright
William Ben Peterman, 2011
All rights reserved.

The dissertation of William Ben Peterman is approved,
and it is acceptable in quality and form for publication
on microfilm and electronically:

Chair

University of California, San Diego

2011

DEDICATION

To my beloved family and friends, which made this endeavor
attainable.

EPIGRAPH

*The nation should have a tax system that looks like someone designed it on
purpose.*

—William Simon

TABLE OF CONTENTS

Signature Page		iii
Dedication		iv
Epigraph		v
Table of Contents		vi
List of Figures		ix
List of Tables		x
Acknowledgements		xi
Vita and Publications		xii
Abstract of the Dissertation		xiii
Chapter 1	The Effect of Endogenous Human Capital Accumulation on Optimal Taxation	1
	1.1 Introduction	2
	1.2 Literature Review	6
	1.3 Analytical Model	9
	1.3.1 Exogenous Age-Specific Human Capital	10
	1.3.2 Learning-By-Doing	14
	1.3.3 Learning-or-doing	18
	1.4 Computational Model	21
	1.4.1 Demographics	21
	1.4.2 Individual	22
	1.4.3 Firm	23
	1.4.4 Government Policy	23
	1.4.5 Definition of Stationary Competitive Equilibrium	24
	1.5 Calibration and Functional Forms	26
	1.5.1 Demographics	27
	1.5.2 Preferences	27
	1.5.3 Age-Specific Human Capital	28
	1.5.4 Firm	30
	1.5.5 Government Policies and Tax Functions	31
	1.6 Computational Experiment	32
	1.7 Results	33
	1.7.1 Optimal Tax Policies in Exogenous, LBD, and LOD Models	33

	1.7.2	The Effects of Adding Endogenous Age-Specific Human Capital	35
	1.7.3	The Effects of the Optimal Tax Policy in the Exogenous Model	40
	1.7.4	The Effects of Optimal Tax Policy in the LBD Model	42
	1.7.5	The Effects of Optimal Tax Policy in the LOD Model	47
1.8		Sensitivity Analysis	48
	1.8.1	Non-Separable Utility	48
	1.8.2	Sequential Parametrization	51
1.9		Conclusion	53
1.10		Appendix	54
	1.10.1	Analytical Derivations	54
	1.10.2	Competitive Equilibrium	60
	1.10.3	Non-separable Utility	64
	1.10.4	Sequential Parametrization	69
Chapter 2		Determining the Motives for a Positive Optimal Tax on Capital	73
	2.1	Introduction	74
	2.2	Analytical Model	77
	2.2.1	Households' Problem	79
	2.2.2	Primal Approach	80
	2.2.3	Mimicking Age-dependent tax on labor	82
	2.2.4	Government balanced budget leads to non-zero tax on capital	85
	2.2.5	Summary of analytic results	86
	2.3	Computational Model	86
	2.3.1	Demographics	87
	2.3.2	Individual	88
	2.3.3	Firm	88
	2.3.4	Government Policy	89
	2.3.5	Definition of Stationary Competitive Equilibrium	90
	2.4	Calibration	91
	2.4.1	Demographics	92
	2.4.2	Individual	92
	2.4.3	Firm	94
	2.4.4	Government Policy	94
	2.5	Computational Experiment	95
	2.6	Results	96
	2.6.1	Impact of Frisch Elasticity and Accidental Bequests	97
	2.6.2	Determining Impact of Each Assumption	97
	2.6.3	Desire to mimic age-dependent tax	98

2.6.4	Government consumption of accidental bequests	100
2.6.5	Individual Liquidity Constraints	102
2.6.6	Government Savings and Debt	103
2.6.7	Social Security Program	106
2.6.8	Summary of Results	107
2.7	Sensitivity Analysis	108
2.7.1	Effect on Optimal Tax Policies in Benchmark Mod- els	109
2.7.2	Frisch Elasticity effect on Channels' Impact	110
2.8	Conclusion	112
2.9	appendix	113
2.9.1	Analytical Derivations	113
Chapter 3	Measuring the Frisch Labor Supply Elasticity: Evidence from a Pseudo Panel	119
3.1	Introduction	120
3.2	Labor Supply Model	122
3.2.1	Individual's Decisions	123
3.3	Data	124
3.4	Results	130
3.4.1	Aggregate Frisch	130
3.4.2	Frisch Elasticity Profile	139
3.5	Conclusion	141
3.6	Appendix	143
3.6.1	Implications of using participation rate elasticity	143

LIST OF FIGURES

Figure 1.1:	Life Cycle Frisch Labor Supply Elasticity in LBD Model	36
Figure 1.2:	Life Cycle Frisch Labor Supply Elasticity in LOD Model	36
Figure 1.3:	Life Cycle Profiles under Baseline Fitted US Tax Policy	39
Figure 1.4:	Life Cycle Profiles in the Exogenous Model	41
Figure 1.5:	Life Cycle Profiles in the LBD Model	43
Figure 1.6:	Life Cycle Profiles in the LOD Model	46
Figure 1.7:	Life Cycle Profiles under Baseline Fitted US Tax Policy (Non-separable Utility)	65
Figure 1.8:	Life Cycle Profiles in the Exogenous Model (Non-separable Utility)	66
Figure 1.9:	Life Cycle Profiles in the LBD Model (Non-separable Utility)	67
Figure 1.10:	Life Cycle Profiles in the LOD Model (Non-separable Utility)	68
Figure 1.11:	Life Cycle Profiles under Baseline Fitted US Tax Policy (Seq. Parametrization)	70
Figure 1.12:	Life Cycle Profiles in the LBD Model (Seq. Parametrization)	71
Figure 1.13:	Life Cycle Profiles in the LoD Model (Seq. Parametrization)	72
Figure 2.1:	Life Cycle Profiles in Model A0 and A1	99
Figure 2.2:	Life Cycle Profiles in Model A0 and A2	101
Figure 2.3:	Life Cycle Profiles in Model A0 and A3	103
Figure 2.4:	Life Cycle Profiles in Model A0 and A4	105
Figure 2.5:	Life Cycle Profiles in Model A0 and A5	106
Figure 2.6:	Life Cycle Profiles in Model A0 and A6	107
Figure 2.7:	Life Cycle Profiles in Benchmark Models with Different Elasticities	110
Figure 3.1:	Hours (All)	126
Figure 3.2:	Hours (Workers)	127
Figure 3.3:	Wages (Workers)	127
Figure 3.4:	All Consumption	128
Figure 3.5:	Food Consumption	128
Figure 3.6:	Participation Rate	129
Figure 3.7:	Average Wage by age for All Individuals	134
Figure 3.8:	Micro Frisch Elasticity Profile	141
Figure 3.9:	Macro Frisch Elasticity Profile	143

LIST OF TABLES

Table 1.1:	Demographics Parameters	27
Table 1.2:	Preference Parameters	29
Table 1.3:	Firm Parameters	30
Table 1.4:	Government Parameters	32
Table 1.5:	Optimal Tax Policies in Benchmark Models	34
Table 1.6:	Aggregate Economic Variables	37
Table 1.7:	Preference Parameters	49
Table 1.8:	Optimal Tax Policies in Non-Separable Models	51
Table 1.9:	Optimal Tax Policies in Sequential Parametrization	52
Table 1.10:	Aggregate Economic Variables (Non-separable Utility)	64
Table 1.11:	Aggregate Economic Variables (Sequential Parametrization)	69
Table 2.1:	Calibration Parameters	92
Table 2.2:	Models	95
Table 2.3:	Optimal Tax Policies	97
Table 2.4:	Optimal Tax Policies	97
Table 2.5:	Aggregate Economic Variables	98
Table 2.6:	Percent Changes Induced by New Optimal Tax Policy in Model	98
Table 2.7:	Calibration Parameters	108
Table 2.8:	Optimal Tax Policies in Benchmark Models with Different Frisch Elasticities	109
Table 2.9:	Optimal Tax Policy in Sensitivity Analysis Under Different Calibrations	111
Table 2.10:	Percentage Point Changes Induced by Change in Model	111
Table 3.1:	Aggregate Micro Frisch Elasticity (Replications)	132
Table 3.2:	Aggregate Micro Frisch Elasticity	136
Table 3.3:	Aggregate Macro Frisch Elasticity	138
Table 3.4:	Micro Frisch Elasticity Profile	140
Table 3.5:	Macro Frisch Elasticity Profiles	142

ACKNOWLEDGEMENTS

I would like to thank Professor Irina Telyukova for her support as the chair of my committee. Through many presentations, multiple drafts, and numerous meetings, her guidance and advice has proved to be invaluable.

I am very grateful for the help and advice from my other committee members, Professor Roger Gordon, Professor Valerie Ramey, Professor Isaac Martin, and Professor Krislert Samphantharak.

I would also like to acknowledge Scott Borger, Kevin Novan, and Vasia Panousi, who provided particularly valuable feedback.

This work benefited from the comments and insights of participants at research seminars at the Federal Reserve Board of Governors, research seminars at the University of California at San Diego, the Annual Eastern Economic Association 2010 Conference, the Annual Missouri Economics Conference 2010, the 2010 Midwest Macroeconomics Meetings, and the International Conference on Computing in Economics and Finance 2010.

VITA

2003	B.A. in Economics and Mathematics, Emory University, Atlanta GA
2005-2011	Teaching Assistant, University of California, San Diego
2011	Ph.D. in Economics, University of California, San Diego

PUBLICATIONS

Bauer, Andy, Nicholas Haltom, and William B. Peterman, “Decomposing Inflation”, *Federal Reserve Bank of Atlanta Economic Review*, 2004 89 (1).

ABSTRACT OF THE DISSERTATION

Optimal Taxation in Life Cycle Models

by

William Ben Peterman

Doctor of Philosophy in Economics

University of California, San Diego, 2011

Professor Irina A. Telyukova, Chair

Whether to tax capital is a central question in both macroeconomics and public finance. Previous research demonstrates that in a life cycle model the optimal tax on capital is typically non-zero for a variety of reasons. My research analytically and quantitatively measures the strength of the different motives for a non-zero tax on capital in a life cycle model. The first chapter considers the impact on the optimal tax policy of including human capital accumulation endogenously. The second chapter measures the relative strength of each motive generally understood to produce a large optimal tax on capital in a standard life cycle model with exogenous age-specific human capital accumulation. The first two chapters demonstrate that the level of the Frisch labor supply elasticity, as well as the pro-

file over the lifetime, have a dramatic impact on the optimal tax policy. The third chapter uses a pseudo panel to estimate the Frisch labor supply elasticity.

Chapter 1

The Effect of Endogenous Human Capital Accumulation on Optimal Taxation

Abstract

This paper considers the impact of endogenous human capital accumulation on optimal tax policy in a life cycle model. Analytically, it demonstrates that including endogenous human capital accumulation, through either learning-*by*-doing or learning-*or*-doing, creates a motive for the government to use age-dependent labor income taxes. If the government cannot condition taxes on age, then it is optimal to use a tax on capital in order to mimic age-dependent taxes on labor income. Quantitatively, this work finds that introducing learning-*by*-doing or learning-*or*-doing increases the optimal tax on capital by eighty or twenty percent, respectively. Including learning-*by*-doing leads younger agents to supply labor relatively less elastically than older agents. Given that taxing capital implicitly taxes younger labor income at a higher rate, the optimal tax on capital is larger in this framework. In the case of learning-*or*-doing, the government increases the tax on capital to encourage individuals to save in the form of human capital as opposed to physical capital.

1.1 Introduction

In their seminal works, Chamley (1986) and Judd (1985) determine that it is not optimal to tax capital in an infinitely-lived agent model. In such a model, taxing capital income is equivalent to an ever increasing tax on future consumption, thus implying an exponentially increasing distortion between the marginal rate of substitution and the marginal rate of transformation. In contrast, in a life cycle model agents live for a finite number of periods so the distortion imposed by a capital tax is bounded and may not necessarily be bigger than the distortions caused by other taxes. Atkeson et al. (1999), Erosa and Gervais (2002), and Garriga (2001) demonstrate in simplified life cycle models that if the government cannot condition labor income taxes on age, the government will generally tax capital in order to mimic an age-dependent tax.¹ The government wants to condition taxes on age since agents vary their consumption and labor over the life cycle. Quantitative exercises, such as Conesa et al. (2009) and chapter 2, demonstrate that in a calibrated life cycle model the inability to condition taxes on age can be a strong motive for a positive tax on capital.

Age-specific human capital is responsible for causing an agent to vary his labor supply and hence the non-zero tax on capital result.² Even though age-specific human capital is a driving mechanism for the positive optimal tax on capital, it is typically incorporated in models exogenously through age-specific productivity levels. By including human capital accumulation exogenously, the models ignore any effect that endogenous accumulation may have on the optimal tax policy. This paper assesses, both analytically and quantitatively, the impact of including endogenous age-specific human capital accumulation in a life cycle model on the optimal capital tax.

¹Atkeson et al. (1999) demonstrate a related result. They show conditions under which the optimal tax on capital is zero if age-dependent taxes on labor income are allowed. Gervais (2010) demonstrates that a progressive labor income tax can also mimicking age-dependent taxes on labor income.

²I define age-specific human capital as changes in an individual's labor productivity throughout his work life. In the US, this can be thought of as productivity changes for individuals beginning in their early 20s. Some other model features, including liquidity constraints and retirement, may also cause variations in consumption and labor over the life cycle.

Specifically, this paper explores the effect on optimal tax policy of two forms of endogenous age-specific human capital accumulation: learning-*by*-doing (LBD) and learning-*or*-doing (LOD). In LBD, an agent acquires human capital by working. In LOD, an agent acquires human capital by spending time training in periods in which he is also working.³ With LBD an agent determines his level of age-specific human capital by choosing the hours he works, while with LOD, an agent determines his human capital by choosing the hours he trains. I analyze the effects of both forms since there is empirical evidence that each form is responsible for age-specific human capital accumulation and each is commonly employed in quantitative life cycle models.⁴

Including endogenous human capital accumulation changes the properties of the optimal capital income tax both qualitatively and quantitatively. In a simplified life cycle model with a utility function that is both separable and homothetic with respect to consumption and hours worked, I analytically demonstrate that, including either form of endogenous age-specific human capital accumulation creates an incentive for the government to condition labor income taxes on age. If age-dependent labor income taxes are not in the feasible policy set, then the optimal tax on capital is non-zero in order to mimic the wedge created by conditioning labor income taxes on age. Specifically, a positive (negative) tax on capital imposes the same wedge on the marginal rate of substitution as a relatively larger (smaller) tax on young labor income. The motive to use age-dependent taxes in the endogenous model is in contrast to Garriga (2001) that demonstrates in a similar model with exogenous age-specific human capital the government does not want to condition the labor income tax on age, meaning that the optimal capital tax is zero.

³In LOD, separate time for training and working are both deducted from leisure. LOD is sometimes referred to as Ben-Porath type skill accumulation or on-the-job training.

⁴Examples of life cycle studies that include variations of LBD or LOD include Hansen and Imrohoroglu (2009), Imai and Keane (2004), Chang et al. (2002), Jones et al. (1997), Jones and Manuelli (1999), Guvenen et al. (2009), Kuruscu (2006), Kapicka (2006), and Kapicka (2009). Topel (1990), Cossa et al. (1999), Altuğ and Miller (1998) provide empirical evidence of LBD. The authors show that past hours worked and length of current job tenure impact current wages. With regards to LOD, numerous studies provide evidence that individuals partake in training and that the training is responsible for wage growth. For examples see Mulligan (1995), Frazis and Loewenstein (2006), Kuruscu (2006), and Mincer (1989).

Adding LBD to the model causes an agent to supply labor relatively less elastically early in his life, which alters the optimal tax policy. In a model with exogenous skill accumulation, an agent's only incentive to work is his wage. In a model with LBD, the benefits from working are current wages as well as an increase in future age-specific human capital. I refer to these benefits as the "wage benefit" and the "human capital benefit," respectively. The importance of the human capital benefit decreases as an agent approaches retirement. Adding LBD thus causes the agent to supply labor relatively less elastically early in his life compared to later in his life. Relying more heavily on a capital tax reduces the distortions that this tax policy imposes on the economy, since it implicitly taxes this less elastically supplied labor income from younger agents at a higher rate than older agents.⁵

Adding LOD to the model also causes a non-zero tax on capital to be optimal if age-dependent taxes are unavailable. There are two channels through which LOD affects the optimal tax policy: the elasticity channel and the savings channel. First, adding LOD changes an agent's elasticity profile. Training is an imperfect substitute for labor as both involve forfeiting leisure in exchange for higher lifetime income. The substitutability of training decreases as an agent ages since he has less time to take advantage of the accumulated skills. Therefore, introducing LOD causes a young agent to supply labor relatively more elastically. The elasticity channel lowers the optimal tax on capital since the tax policy implicitly taxes labor income from younger agents at a lower rate. The second channel, the savings channel, arises because training is an alternative method of saving, as opposed to accumulating physical capital. Therefore, the government can increase an agent's incentives to train by taxing capital (or taxing young labor income at relatively higher rates) since it makes training a relatively more desirable way to save. Since these two channels have counteracting effects, one cannot analytically determine the cumulative direction of their impact on the optimal tax policy.

Next, I quantitatively assess the effect of adding each form of endogenous

⁵A standard result in public finance is that, if it is necessary to use distortionary taxes, it is optimal to tax inelastically supplied factors at relatively higher rates since this policy will minimize the distortions to the economy.

age-specific human capital accumulation on optimal tax policy in a calibrated life cycle model using the specific utility function from Garriga (2001). The optimal tax rates in the model with exogenous age-specific human capital accumulation (exogenous model) are 11.8% on capital and 24.7% on labor. I find that adding either form of endogenous human capital increases the optimal tax on capital. In the model with LBD the optimal tax rates are 21.1% on capital and 23.3% on labor. The optimal tax rates in the model with the LOD framework are 14.3% on capital and 24.3% on labor. Adding endogenous age-specific human capital accumulation raises the optimal tax on capital over twenty percent in the LOD framework and over eighty percent in the LBD framework.

I test the sensitivity of these results with respect to both the methodology used to calibrate the parameter values and the utility function. I find that using an alternative procedure to determine the parameter values in which all the values are the same across the models, does not change the effect of endogenous human capital accumulation on optimal tax policy. I find that using an alternative utility function that is neither separable nor homothetic with respect to consumption, and hours worked implies that the optimal tax on capital is much larger in the exogenous model. I still find that including either form of endogenous human capital accumulation with this utility function causes the optimal tax to increase; although, the large optimal tax on capital in the exogenous model crowds out some of their impact.

This paper is organized as follows: Section 2 discusses the relevant literature. Section 3 examines an analytically tractable version of the model in order to demonstrate that including endogenous human capital accumulation creates a motive for the government to condition labor income taxes on age. Section 4 describes the full model and the competitive equilibrium used in the quantitative exercises. The calibration and functional forms are discussed in section 5. Section 6 describes the computational experiment, and section 7 presents the results. Section 8 tests the sensitivity of the results with respect to calibration parameters and utility specifications. Section 9 concludes.

1.2 Literature Review

Following the zero tax on capital result in Chamley (1986) and Judd (1985), subsequent works identify at least three sufficient conditions in which it is optimal to tax capital in a life cycle model: (i) when individuals face uninsurable idiosyncratic risk; (ii) when an individual's earnings increase over his lifetime and he faces borrowing constraints; and (iii) when the government cannot tax all factors of production or sources of income at distinct rates.

Chamley (2001) demonstrates that when individuals face uninsurable idiosyncratic risk, the optimal tax on capital is non-zero in order to provide partial insurance. In a stochastic economy, the direction of the tax depends on the correlation between consumption and savings. It is optimal to have a positive (negative) tax on capital when consumption is positively (negatively) correlated with savings. Similarly, Aiyagari (1995) demonstrates that when agents face uninsurable idiosyncratic earnings risk, the optimal tax on capital is positive. Panousi (2009) finds that a tax on capital is welfare improving in a model that includes idiosyncratic investment risk.

Hubbard and Judd (1986) quantitatively demonstrate that a positive tax on capital income can lead to welfare gains by shifting some of the tax burden from lower income to higher income years. Agents generally prefer to smooth their consumption. Therefore, if an agent's earnings increase over his lifetime, he would prefer to smooth his consumption by borrowing against earnings from later years in order to increase consumption in earlier years. Borrowing constraints hinder an agent's ability to make such transfers, creating a role for tax policy to facilitate such a transfer. Since an individual typically accumulates more assets later in life, increasing the tax on capital income and decreasing the tax on labor income will allocate more of the lifetime tax burden to an individual's later years, helping to smooth consumption. Furthermore, İmrohoroğlu (1998) confirms in a richer quantitative exercise that it is optimal to tax capital in an overlapping generations model (OLG). The author concludes that the presence of borrowing constraints, combined with the timing of the burden of taxation over the life cycle, drives the positive tax on capital income. In his model, he assumes labor is determined ex-

ogenously which means that the author may over emphasize the effects of liquidity constraints since agents cannot ease binding liquidity constraints by increasing their labor supply.

Correia (1996), Armenter and Albanesi (2009), and Jones et al. (1997) demonstrate that if the government cannot tax all factors of production at separate rates then a non-zero tax on capital may be optimal in order to mimic such taxes. Erosa and Gervais (2002) and Garriga (2001) demonstrate a specific example of this result; it is optimal to tax capital in order to mimic an age-dependent tax on labor income when it is not in the feasible policy set. The shape of an agent's lifetime labor supply elasticity profile affects how the benevolent planner would want to condition age-dependent taxes on labor income. For example, if an agent supplies labor more elastically as he ages, then it is typically optimal to impose labor income taxes that decrease over the agent's lifetime. This tax policy minimizes the distortions it introduces into the economy by taxing factors that are supplied less elastically at a higher rate. A positive capital income tax can replicate this type of age-dependent labor income tax because it implicitly taxes labor income for younger individuals at a higher rate.

Conesa et al. (2009), henceforth CKK, solve a calibrated life cycle model with endogenously determined labor in order to determine which, if any, of these motives are quantitatively significant. Their model differs from İmrohorođlu (1998) in that an agent's labor supply is determined endogenously. The model that CKK solve numerically is similar to the models that Atkeson et al. (1999), Erosa and Gervais (2002), and Garriga (2001) solve analytically, but include liquidity constraints along with additional life cycle features such as accidental bequests, retirement and social security. They determine that the optimal tax policy is a flat 34% tax on capital and a flat 14% tax on labor income.⁶ They state that a primary motive for imposing a high tax on capital income is to mimic a relatively larger labor income tax on younger agents when they supply labor relatively less elastically. An agent

⁶This is model M4 in Conesa et al. (2009). I refer to CKK's model that abstracts from idiosyncratic earnings risk and within-cohort heterogeneity because they find that these features do not affect the level of the optimal tax on capital. Therefore, I also abstract from these features in my analysis.

supplies labor more elastically as he ages because his labor supply is decreasing, and the authors use a utility specification in which the agent's Frisch labor supply elasticity is a negative function of hours worked. Chapter 2 confirms that this is an economically significant motive for the positive tax on capital, but concludes that the government's inability to save and to consume accidental bequests also contribute to the positive tax on capital. Fuster et al. (2008) demonstrate that the welfare gains from increasing the relative tax on capital decrease as the strength of altruism increases. Nakajima (2010) demonstrates that the optimal tax on capital falls if preferential tax treatment for owner-occupied housing is included in a life cycle model. The tax on ordinary capital falls in order to narrow the tax wedge between housing and non-housing capital. This paper extends these studies of optimal tax policy by determining the effects of endogenous age-specific human capital accumulation on the optimal tax policy in a standard life cycle model.

Past work that analyzes the trade off between labor and capital taxes in a model that includes endogenous age-specific human capital accumulation, typically does not use a life cycle model.⁷ For example, both Jones et al. (1997) and Judd (1999) examine optimal capital tax in an infinitely lived agent model in which agents are required to use market goods to acquire human capital similar to ordinary capital. They find that if the government can distinguish between pure consumption and human capital investment, then it is not optimal to distort either human or physical capital accumulation in the long run. Reis (2007) shows in a similar model that if the government cannot distinguish between consumption and human capital investment, then the optimal tax on capital is still zero as long as

⁷Jacobs and Bovenberg (2009) analyze the trade off between a labor and capital tax in a life cycle model with pre-work education. The authors find that in a two period model where agents acquire education in the first period and work in the second period the optimal tax on capital is generally positive if educational investment is not verifiable. The tax on capital reduces the tax on labor income which in turn reduces the distortions on the benefit to education. Jacobs and Bovenberg (2009) is related to the current work, however they focus on human capital accumulation prior to working while the current study examines the impact of endogenous human capital accumulation once an agent begins to work. Additionally, a related line of literature examines the effect of taxes on human capital accumulation in models with endogenous human capital accumulation. However, these studies tend to focus on the interaction of endogenous human capital accumulation and the progressivity of an overall income tax and not the effect on the optimal ratio of separate tax instruments. For example, see Jacobs (2005), Kapicka (2006), Kapicka (2009), Caucutt et al. (2000), and Bénabou (2002).

the level of capital does not influence the relative productivity of human capital. Chen et al. (2010) find in an infinitely lived agent model with labor search, that including endogenous human capital accumulation through both LBD and LOD causes the optimal tax on capital to increase because a higher tax on capital unravels the labor market frictions. In this model, a tax on capital causes the wage discount to increase, thus causing firms to post more vacancies which in turn causes an increase in worker participation. The current paper is related to these other works however it differs in that it analyzes optimal tax policy in an OLG model as opposed to an infinitely lived agent model . Therefore, these other studies do not account for the effects of endogenous human capital accumulation through life cycle channels. It is especially important to include the life cycle channel since Conesa et al. (2009) and chapter 2 demonstrate that this channel is quantitatively important for motivating a positive tax on capital.

1.3 Analytical Model

In this section I reexamine the result from Garriga (2001) that for a specific utility function the government has no incentive to condition labor income taxes on age. I find that adding endogenous human capital accumulation to this model causes the optimal tax policy to include age-dependent taxes and if age-dependent taxes are unavailable then a non-zero tax on capital is optimal. I derive analytical results in a tractable two-period version of the computational model. For tractability purposes, the features I abstract from include: retirement, population growth, progressive tax policy, and conditional survivability. Additionally, for the purposes of the analytical model I assume that the technology is such that the marginal products of capital and labor are constant. This assumption permits me to focus on the life cycle elements of the model, in that changes to the tax system do not affect the pre-tax wage or rate of return. Since there is no variation in the factor prices, I suppress the time subscripts on the factor prices in this section.

I begin by replicating the result in Garriga (2001) that it is not optimal to condition labor income taxes with a benchmark utility function that is homothetic

with respect to consumption and hours worked,

$$U(c, h) = \frac{c^{1-\sigma_1}}{1-\sigma_1} - \chi \frac{(h)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}}.$$

I set up the household problem and demonstrate that a positive (negative) tax on capital induces a wedge on the marginal rate of substitution that is similar to a relatively larger tax on young (old) labor income. Using the primal approach I then solve for the optimal tax policy in the exogenous model, which confirms the Garriga (2001) result that since the government has no incentive to condition labor income taxes on age the optimal tax on capital is zero. I then add endogenous human capital to the exogenous model and demonstrate that it creates a motive for the government to condition labor income taxes on age. I also demonstrate the channels by which the forms on endogenous human capital accumulation affect the optimal tax policy.

1.3.1 Exogenous Age-Specific Human Capital

General Set-up

In the analytically tractable model, agents live with certainty for two periods and with preferences over consumption and labor represented by

$$U(c_{1,t}, h_{1,t}) + \beta U(c_{2,t+1}, h_{2,t+1}) \tag{1.1}$$

where β is the discount rate, $c_{j,t}$ is the consumption of an age j agent at time t , and $h_{j,t}$ is the percent of the time endowment the agent works.⁸ Age-specific human capital is normalized to unity when the agent is young. At age 2, age-specific human capital is ϵ_2 .⁹ The agent maximizes equation 1.1 with respect to

⁸Time working is measured as a percentage of endowment and not in hours. However for expositional convenience, I also refer to $h_{j,t}$ as hours.

⁹Since age-specific human capital is exogenous and predetermined in this model, it could also be considered age-specific productivity.

consumption and hours subject to the following constraints

$$c_{1,t} + a_{1,t} = (1 - \tau_{h,1})h_{1,t}w \quad (1.2)$$

and

$$c_{2,t+1} = (1 + r(1 - \tau_k))a_{1,t} + (1 - \tau_{h,2})\epsilon_2 h_{2,t+1}w \quad (1.3)$$

where $a_{1,t}$ is the amount young agents save, $\tau_{h,j}$ is the tax rate on labor income for an agent of age j , τ_k is the tax rate on capital income, w is the efficiency wage for labor services and r is the rental rate on capital. I assume that the tax rate on labor income can be conditioned on age; however the tax rate on capital income cannot.¹⁰ I combine equations 1.2 and 1.3 to form a joint intertemporal budget constraint

$$c_{1,t} + \frac{c_{2,t+1}}{1 + r(1 - \tau_k)} = w(1 - \tau_{h,1})h_{1,t} + \frac{w(1 - \tau_{h,2})\epsilon_2 h_{2,t+1}}{1 + r(1 - \tau_k)}. \quad (1.4)$$

The agent's problem is to maximize equation 1.1 subject to 1.4. The agent's first order conditions are

$$\frac{U_{h1}(t)}{U_{c1}(t)} = -w(1 - \tau_{h,1}) \quad (1.5)$$

$$\frac{U_{h2}(t+1)}{U_{c2}(t+1)} = -w\epsilon_2(1 - \tau_{h,2}) \quad (1.6)$$

and

$$\frac{U_{c1}(t)}{U_{c2}(t+1)} = \beta(1 + r(1 - \tau_k)) \quad (1.7)$$

where $U_{c1}(t) \equiv \frac{\partial U(c_{1,t}, h_{1,t})}{\partial c_{1,t}}$. Given a social welfare function, prices, and taxes, these first order conditions, combined with the intertemporal budget constraint, determine the optimal allocation of $(c_{1,t}, h_{1,t}, c_{2,t+1}, h_{2,t+1})$.

¹⁰Agents only live for two periods in the analytically tractable model so they choose not to save when they are old. Therefore, in this model the restriction on the capital tax policies is not binding.

Tax on Capital Mimics Age-Dependent Tax on Labor

In order to demonstrate why a tax on capital has an impact similar to an age-dependent labor income tax, I derive the intertemporal Euler equation by combining equations 1.5, 1.6 and 1.7

$$\epsilon_2 \frac{U_{h1}(t)}{U_{h2}(t+1)} = \beta(1 + r(1 - \tau_k)) \frac{1 - \tau_{h,1}}{1 - \tau_{h,2}}. \quad (1.8)$$

Equation 1.8 demonstrates that if the government wants to create a wedge on the marginal rate of substitution by varying the age-dependent labor income taxes, then τ_k is an alternative option. A positive tax on capital mimics the wedge imposed by a relatively higher tax rate on young labor income. Therefore, if the government has an incentive to condition taxes on age but age-dependent taxes are unavailable, then a non-zero tax on capital is optimal.

Primal Approach

In order to determine the optimal tax policy I use the primal approach.¹¹ I use a social welfare function that maximizes utility and discounts future generations with social discount factor θ ,

$$[U(c_{2,0}, h_{2,0})/\theta] + \sum_{t=0}^{\infty} \theta^t [U(c_{1,t}, h_{1,t}) + \beta U(c_{2,t+1}, 1 - h_{2,t+1})]. \quad (1.9)$$

The government maximizes this objective function with respect to three constraints: the implementability constraint; the resource constraint; and the government budget constraint. The implementability constraint is the agent's intertemporal budget constraint, with prices and taxes replaced by his first order conditions (equations 1.5, 1.6, and 1.7)

$$c_{1,t}U_{c1}(t) + \beta c_{2,t+1}U_{c2}(t+1) + h_{1,t}U_{h1}(t) + \beta h_{2,t+1}U_{h2}(t+1) = 0. \quad (1.10)$$

¹¹See Lucas and Stokey (1983) or Erosa and Gervais (2002) for a full description of the primal approach.

Including this constraint ensures that any allocation the government chooses can be supported by a competitive equilibrium. The resource constraint is

$$c_{1,t} + c_{2,t} + K_{t+1} - K_t + G_t = rK_t + w(h_{1,t} + h_{2,t}\epsilon_2). \quad (1.11)$$

I assume that the government is allowed to hold constant levels of savings or debt.¹² Therefore, due to Walras' Law, I include two of three constraints in the Lagrangian and leave out the government budget constraint. Including the benchmark utility specification, the Lagrangian the government maximizes is

$$\begin{aligned} \mathcal{L} = & \frac{c_{1,t}^{1-\sigma_1}}{1-\sigma_1} - \chi \frac{h_{1,t}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} + \beta \frac{c_{2,t+1}^{1-\sigma_1}}{1-\sigma_1} - \chi \frac{h_{2,t+1}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} \\ & - \rho_t(c_{1,t} + c_{2,t} + K_{t+1} - K_t + G_t - rK_t - w(h_{1,t} + h_{2,t}\epsilon_2)) \\ & - \rho_{t+1}\theta(c_{1,t+1} + c_{2,t+1} + K_{t+2} - K_{t+1} + G_{t+1} - rK_{t+1} - w(h_{1,t+1} + h_{2,t+1}\epsilon_2)) \\ & + \lambda_t(c_{1,t}^{1-\sigma_1} + \beta c_{2,t+1}^{1-\sigma_1} - \chi h_{1,t}^{1+\frac{1}{\sigma_2}} - \beta \chi h_{2,t+1}^{1+\frac{1}{\sigma_2}}) \end{aligned} \quad (1.12)$$

where ρ is the Lagrange multiplier on the resource constraint and λ is the Lagrange multiplier on the implementability constraint.

Optimal Tax Policy

I solve for the optimal tax policy in the analytically tractable exogenous model. The formulation of the government's problem and their first order conditions for this model can be found in appendix 1.10.1. Combining the government's first order conditions generates the following expression for optimal labor income taxes

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \frac{1 + \lambda_t(1 + \frac{1}{\sigma_2})}{1 + \lambda_t(1 + \frac{1}{\sigma_2})} = 1. \quad (1.13)$$

¹²The assumption that the equilibrium in the exogenous model can be decentralized relies heavily on the government being allowed to hold debt. If the government cannot hold debt then the additional constraint, $K_t = a_{1,t}$ must be added and the optimal tax on capital is no longer zero. See chapter 2 for further details.

Equation 1.13 demonstrates that the government has no incentive to condition labor income taxes on age when age-specific human capital is included exogenously.¹³

Utilizing the first order condition from the Lagrangian with respect to capital and consumption leads to the following equation,

$$\left(\frac{c_{1,t}}{c_{2,t+1}} \right)^{-\sigma_1} = \beta(1+r) \quad (1.14)$$

Applying the benchmark utility function to equation 1.7 provides the following relationship

$$\left(\frac{c_{1,t}}{c_{2,t+1}} \right)^{-\sigma_1} = \beta(1+r(1-\tau_k)). \quad (1.15)$$

Equations 1.14 and 1.15 demonstrate that in order for the household to choose the optimal allocation indicated by the primal approach, the tax on capital must equal zero.¹⁴ In the exogenous model the government has no incentive to condition labor income taxes on age and the optimal tax on capital is zero.

1.3.2 Learning-By-Doing

Including LBD Creates Motive for Age-Dependent Taxes on Labor Income

Next, I introduce LBD into the exogenous model. In the LBD model, age-specific human capital for a young agent is normalized to unity. Age-specific human capital for an old agent is determined by the function $s_2(h_{1,t})$. $s_2(h_{1,t})$ is a positive and concave function of the hours worked when young. In this model

¹³This result is specific to this utility function. See Garriga (2001) for further details.

¹⁴Regardless of whether the government can condition labor income taxes on age, in this model they do not want to tax capital because there is no desire to mimic an age-dependent tax on labor income. When the government cannot condition labor income taxes on age the Lagrangian includes an additional constraint

$$\epsilon_2 \frac{U_{h1}(t)}{U_{c1}(t)} = \frac{U_{h2}(t+1)}{U_{c2}(t+1)}. \quad (1.16)$$

However, in the analytically tractable model with exogenous human capital accumulation, this constraint is not binding and thus the Lagrange multiplier equals zero.

agents maximize

$$U(c_{1,t}, h_{1,t}) + \beta U(c_{2,t+1}, h_{2,t+1}) \quad (1.17)$$

subject to

$$c_{1,t} + a_{1,t} = (1 - \tau_{h,1})h_{1,t}w \quad (1.18)$$

and

$$c_{2,t+1} = (1 + r(1 - \tau_k))a_{1,t} + (1 - \tau_{h,2})s_2(h_{1,t})h_{2,t+1}w. \quad (1.19)$$

The agent's first order conditions are given by

$$\frac{U_{h_1}(t)}{U_{c_1}(t)} = -[w(1 - \tau_{h,1}) + \beta \frac{U_{c_2}(t+1)}{U_{c_1}(t)} w(1 - \tau_{h,2})h_{2,t+1}s_{h_1}(t+1)] \quad (1.20)$$

$$\frac{U_{h_2}(t+1)}{U_{c_2}(t+1)} = -ws_2(h_{1,t})(1 - \tau_{h,2}) \quad (1.21)$$

and

$$\frac{U_{c_1}(t)}{U_{c_2}(t+1)} = \beta(1 + r(1 - \tau_k)). \quad (1.22)$$

The first order conditions with respect to h_2 and a_1 are similar in the LBD and exogenous models (equations 1.21 and 1.22 are similar to equations 1.6 and 1.7). However, the first order condition with respect to h_1 is different in the two models (equation 1.20 is different from equation 1.5) because working has the additional human capital benefit in the LBD model. This human capital benefit also alters the implementability constraint. Suppressing the arguments of the skills function, the implementability constraint in the LBD model is

$$0 = c_{1,t}U_{c_1}(t) + \beta c_{2,t+1}U_{c_2}(t+1) + h_{1,t}U_{h_1}(t) - \frac{\beta h_{1,t}U_{h_2}(t+1)h_2s_{h_1}(t+1)}{s_2} + \beta h_{2,t+1}U_{h_2}(t+1), \quad (1.23)$$

where $s_{h_1}(t+1)$ represents the partial derivative of the skill function for an older agent with respect to hours worked when young.

The formulation for the government's problem and the resulting first order conditions (utilizing the benchmark utility function) are in appendix 1.10.1. Combining the first order conditions from the government's problem yields the

following ratio for optimal labor income taxes,

$$\frac{1-\tau_{h,1}}{1-\tau_{h,2}} = \frac{\left[1+h_{2,t+1}s_{h2}(t+1)\right] \left[1+\lambda_t\left(1+\frac{h_{1,t}s_{h2}(t+1)}{s_2}\right)\left(1+\frac{1}{\sigma_2}\right)\right]}{1+\lambda_t\left(1+\frac{1}{\sigma_2}\right)-\beta h_{2,t+1} h_{1,t} \left[\frac{s_{h2}(t+1)}{s_2} \left(1+\lambda_t\left(1+\frac{h_{1,t}s_{h2}(t+1)}{s_2}\right)\right)\left(1+\frac{1}{\sigma_2}\right)-h_1 \left(\left(\frac{s_{h2}(t+1)}{s_2}\right)^2 - \frac{s_{h2,h2}(t+1)}{s_2}\right)\right]} - \frac{h_{2,t+1}s_{h2}(t+1)}{1+r(1-\tau_k)}. \quad (1.24)$$

Equation 1.24 demonstrates that generally in the LBD model the government has an incentive to condition labor income taxes on age. This result contrasts with the exogenous model, in which the government has no incentive to condition labor income taxes on age (see equation 1.13).

LBD Enhances Motive for Postive Tax on Capital

In order to demonstrate why including LBD causes the optimal tax policy to include relatively larger taxes on young labor income, I solve for the intertemporal Euler equation (by combining equations 1.20, 1.21 and 1.22)

$$s_2(h_{1,t}) \frac{U_{h1}(t)}{U_{h2}(t+1)} = \beta(1+r(1-\tau_k)) \frac{1-\tau_{h,1}}{1-\tau_{h,2}} + \beta h_{2,t+1} s_{h1}(t+1). \quad (1.25)$$

Comparing equation 1.8 and equation 1.25, it is clear that the LBD intertemporal Euler equation has an extra term which is positive. Therefore, holding all else equal, the tax on young labor income must be relatively higher, in order to induce the same wedge on the marginal rate of substitution in the LBD model.¹⁵ If age-dependent taxes are not in the governments feasible set, then a larger tax on capital will be necessary to induce the same wedge.

By examining the Frisch elasticities in the exogenous and LBD models, it is clear why adding LBD increases the optimal relative tax on young labor income

¹⁵In the case of the benchmark utility function, the government does not want to introduce a wedge in the exogenous model. In this example, holding all else equal, the government needs to include a relatively larger labor income tax on young agents in the LBD model in order to induce the same zero wedge on the left hand side of the equation.

or tax on capital. These elasticities extend to a model where agents live for more than two periods, and I denote an agent's age with i . In the exogenous model the Frisch elasticity simplifies to

$$\bar{\Xi}_{\text{exog}} = \sigma_2 \quad (1.26)$$

The Frisch elasticity in the LBD model is

$$\Xi_{\text{LBD}} = \frac{\sigma_2}{1 - \frac{h_{i+1,t+1}w_{t+1}(h_{i,t}\sigma_2 s_{hi,hi}(t+1) - s_{hi}(t+1))}{s_{i,t}(1+r_t(1-\tau_k))w_t}}}.^{16} \quad (1.27)$$

The Frisch elasticity in the exogenous model is constant and valued at σ_2 . In the LBD model, the extra terms in Ξ_{LBD} increase the size of the denominator compared to the exogenous model, thus holding hours and consumption constant between the two models, $\bar{\Xi}_{\text{exog}} > \Xi_{\text{LBD}}$. Intuitively, the inclusion of the human capital benefit makes workers less responsive to a one period change in wages since the wage benefit is only part of their total compensation for working in the LBD model. The other benefit, the human capital benefit does not have a constant effect on an agent's Frisch elasticity over his lifetime. The relative importance of the human capital benefit decreases over an agent's lifetime because he has fewer periods to utilize his higher human capital as he ages.¹⁷ Therefore, adding LBD causes a young agent to supply labor relatively less elastically than an older agent. This shift in relative elasticities creates an incentive for the government to tax labor income of younger agents at a relatively higher rate. Thus, if the government cannot condition labor income taxes on age, then the optimal tax on capital is higher in the LBD model. I use the term elasticity channel to describe the impact on optimal tax policy caused by a change in the Frisch elasticity from including endogenous human capital. The elasticity channel is responsible for the change in optimal tax policy from including LBD.

¹⁶This is the Frisch elasticity with respect to a temporary increase in the wage. Therefore, one must distinguish between w_t and w_{t+1} .

¹⁷In order for the human capital benefit to decline over the lifetime it is necessary to assume agents work for a finite number of periods.

1.3.3 Learning-or-doing

Including LOD Creates Motive for Age-Dependent Taxes on Labor Income

I include LOD in the exogenous model to demonstrate that this form of endogenous age-specific human capital accumulation also creates a motive for the government to condition labor income taxes on age. Similar to the other models, age-specific human capital for a young agent is set to unity. Age-specific human capital for an old agent is determined by the function $s_2(n_{1,t})$. $s_2(n_{1,t})$ is a positive and concave function of the hours spent training when an agent is young ($n_{1,t}$). In this model agents maximize

$$U(c_{1,t}, h_{1,t} + n_{1,t}) + \beta U(c_{2,t+1}, h_{2,t+1}) \quad (1.28)$$

subject to

$$c_{1,t} + a_{1,t} = (1 - \tau_{h,1})h_{1,t}w \quad (1.29)$$

and

$$c_{2,t+1} = (1 + r(1 - \tau_k))a_{1,t} + (1 - \tau_{h,2})s_2(n_{1,t})h_{2,t+1}w. \quad (1.30)$$

The agent's first order conditions are given by

$$\frac{U_{h1}(t)}{U_{c1}(t)} = -[w(1 - \tau_{h,1})] \quad (1.31)$$

$$\frac{U_{h2}(t+1)}{U_{c2}(t+1)} = -ws_2(n_{1,t})(1 - \tau_{h,2}) \quad (1.32)$$

$$\frac{U_{c1}(t)}{U_{c2}(t+1)} = \beta(1 + r(1 - \tau_k)) \quad (1.33)$$

and

$$\frac{U_{n1}(t)}{U_{c2}(t+1)} = -\beta w(1 - \tau_{h,2})s_{n1}(n_{1,t})h_{2,t+1} \quad (1.34)$$

The first order conditions with respect to h_1 , h_2 and a_1 are similar in the LOD model and the exogenous model (equations 1.31, 1.32, and 1.33 are similar to

equations 1.5, 1.6, and 1.7).¹⁸ However, the first order condition with respect to n_1 (equation 1.34) is new in the LOD model. This new first order condition requires an additional constraint in the government's Lagrange that equates equations 1.32 and 1.34. This constraint simplifies to $U_n(t)s_2 = \beta U_{h_2}(t+1)h_{2,t+1}s_n(t+1)$. I use η_t as the lagrange multiplier on this new constraint. In the LOD model I need a utility function that incorporates training. I alter the benchmark utility specification so that it consistently incorporates the disutility to non-leisure activities, $\frac{c^{1-\sigma_1}}{1-\sigma_1} - \chi \frac{(h+n)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}}$.

The formulation of the government's problem and resulting first order conditions are provided in appendix 1.10.1. Combing the first order conditions yields the following relationship for optimal taxes on labor income,

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \frac{1 + \lambda_t \left(1 + \frac{h_{1,t}}{\sigma_2(h_{1,t} + n_{1,t})} \right) + \frac{\eta_t s_2}{\sigma_2(h_{1,t} + n_{1,t})}}{1 + \lambda_t \left(1 + \frac{1}{\sigma_2} \right) - \eta_t s_{n1}(t+1) \left(1 + \frac{1}{\sigma_2} \right)}. \quad (1.35)$$

Equation 1.35 demonstrates that the government has an incentive to condition labor income taxes on age when LOD is introduced into the model.

Although equation 1.35 shows that including LOD creates an incentive for the government to condition labor income taxes on age, it is unclear at which age the government wants to impose a relatively higher labor income tax. Comparing equations 1.13 and 1.35, there are two channels through which introducing LOD changes the optimal tax policy. The first channel results from using a utility function that is non-separable in training and labor. The non-separability affects the optimal tax policy through the elasticity channel since it causes LOD to alter the Frisch elasticity. The non-separability causes the numerator of the ratio to change. Whereas in the exogenous model, the term simplifies to one, the numerator now includes the additional term $\frac{h_{1,t}}{h_{1,t} + n_{1,t}}$. This new term causes the expression to decrease. The second channel results from the intertemporal link created by the additional constraints. I call this channel the savings channel because the intertemporal link exists because agents can use training as an alternative means

¹⁸Since the first order conditions with respect to hours and savings are the same in the LOD and exogenous model, the implementability constraints are the same.

to save. This second channel causes the additional terms $-\eta_t s_{n1}(t+1)\left(1 + \frac{1}{\sigma_2}\right)$ and $\frac{\eta_t s_2}{\sigma_2(h_{1,t} + n_{1,t})}$ in the denominator and numerator, respectively. Assuming that η_t is positive then this additional term causes the expression to increase.¹⁹ In this case, these two channels have opposing effects on the optimal tax policy, and the overall effect is unclear.²⁰

Examining the Frisch labor supply elasticities provides intuition for how the first channel affects the optimal tax policy. In the exogenous model, the Frisch elasticity for the benchmark utility specification is constant, σ_2 . Since the altered utility function is not additively separable in time spent working and training, the Frisch labor supply elasticity is not constant in the LOD model. The Frisch elasticity for the altered utility function is $\Xi_{\text{LOD}} = \frac{\sigma_2(h+n)}{h}$. This functional form implies that an agent supplies labor relatively more elastically when he spends a larger proportion of his non-leisure time training. Therefore, if an agent spends less time training as he ages then he will supply labor relatively more elastically when he is young and the government would want to tax the labor income from young agents at a relatively lower rate. One way to mimic this age-dependent tax is to decrease the tax on capital. Therefore, the effect of LOD through the elasticity channel is to decrease the tax on capital.

Examining an agent's first order condition with respect to training demonstrates how the savings channel affects the optimal tax policy. An agent optimizes his choices such that the marginal disutility of training when he is young equals marginal benefit of training ($U_{n1}(t) = \frac{U_{c1}(t)w(1-\tau_{h,2})h_{2,t+1}s_{n1}(t+1)}{1+r(1-\tau_k)}$). The marginal benefit is raised by increasing the tax on capital or decreasing the tax on older labor income. By adopting either of these changes the government makes it relatively more beneficial for the agent to use training in order to save, as opposed to using the risk free asset. Therefore, the government increases the tax on capital in order to increase the incentives for training.

¹⁹The value of η will depend on whether the government wants to increase the relative incentive to save with training or capital. I generally find in the computational simulations that η is positive and therefore treat it as positive in the exposition.

²⁰If an alternative utility function is used that is additively separable in training and hours then LOD only affects the optimal tax policy through the second channel and the ratio increases. An example of such a utility function is $\frac{c^{1-\sigma_1}}{1-\sigma_1} - \chi_1 \frac{(h)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} - \chi_2 \frac{(n)^{1+\frac{1}{\sigma_3}}}{1+\frac{1}{\sigma_3}}$.

Overall, adding LOD creates an incentive for the government to condition labor income taxes on age. Once again, if the government cannot condition labor income taxes on age, it would then want a non-zero tax on capital to mimic this tax. However, whether they would want a relatively higher labor income tax on young or old agents is not analytically clear in the LOD model.

1.4 Computational Model

In order to determine the direction and magnitude of the impact of endogenous human capital accumulation on optimal tax policy, I solve for the optimal tax policies in the LBD and LOD models and compare them to the exogenous model. The exogenous model is adapted from CKK, however I use a different benchmark utility function. Additionally, since the authors find that neither idiosyncratic earnings risk nor heterogeneous ability types are important motives for a positive tax on capital income, I exclude these sources of heterogeneity. In this section I describe the models and define the competitive equilibrium for each model.

1.4.1 Demographics

In the computational model, time is assumed to be discrete and there are J overlapping generations. Conditional on being alive at age j , Ψ_j is the probability of an agent living to age $j + 1$. All agents who live to an age of J die in the next period. If an agent dies with assets, the assets are confiscated by the government and distributed equally to all the living agents as transfers (Tr_t). All agents are required to retire at an exogenously set age j_r .

In each period a cohort of new agents is born. The size of the cohort born in each period grows at rate n . Given the population growth rate and conditional survival probabilities, the time invariant cohort shares, $\{\mu_j\}_{j=1}^J$, are given by

$$\mu_j = \frac{\Psi_{j-1}}{1+n} \mu_{j-1}, \text{ for } i = 2, \dots, J, \quad (1.36)$$

where μ_1 is normalized such that

$$\sum_{j=1}^J \mu_j = 1 \quad (1.37)$$

1.4.2 Individual

An individual is endowed with one unit of productive time per period which he divides between leisure and non-leisure activities. In the exogenous and LBD models the non-leisure activity is providing labor. In the LOD model the non-leisure activities include training as well as providing labor. An agent chooses how to spend his time endowment in order to maximize his lifetime utility

$$\sum_{j=1}^J \prod_{q=1}^j (\Psi_{q-1}) \beta^{j-1} u(c_j, h_j, n_j), \quad (1.38)$$

where c_j is the consumption of an agent at age j , h_j is the hours spent providing labor services, and n_j is the time spent training. Agents discount the next period's utility by the product of Ψ_j and β . β is the discount factor conditional on surviving, and the unconditional discount rate is $\beta\Psi_j$.

In the exogenous model an agent's age-specific human capital is ϵ_j . In the endogenous models, an agent's age-specific human capital, s_j , is endogenously determined. In the LBD model s_j is a function of a skill accumulation parameter, previous age-specific human capital, and time worked, denoted by $s_j = S_{\text{LBD}}(\Omega_{j-1}, s_{j-1}, h_{j-1})$. In the LOD model, s_j is a function of a skill accumulation parameter, previous age-specific human capital, and time spent training, denoted by $s_j = S_{\text{LOD}}(\Omega_{j-1}, s_{j-1}, n_{j-1})$. $\{\Omega_j\}_{j=1}^{j^r-1}$ is a sequence of calibration parameters that are set so that in the endogenous models, under the baseline fitted US tax policy, the agent's choices result in an agent having the same age-specific human capital as in the exogenous model. Individuals command a labor income of $h_j \epsilon_j w_t$ in the exogenous model and $h_j s_j w_t$ in the endogenous model.

Agents split their labor income between consumption and savings with a risk free asset. An agent's level of assets is denoted a_j and the asset pays a pre-tax

net return of r_t .

1.4.3 Firm

Firms are perfectly competitive with constant returns to scale production technology. Aggregate technology is represented by a Cobb-Douglas production function. The aggregate resource constraint is,

$$C_t + K_{t+1} - (1 - \delta)K_t + G_t \leq K_t^\alpha N_t^{1-\alpha}, \quad (1.39)$$

where K_t , C_t , and N_t represent the aggregate capital stock, aggregate consumption, and aggregate labor (measured in efficiency units), respectively. Additionally, α is the capital share and δ is the depreciation rate for physical capital. Unlike the analytically tractable model, I do not assume a linear production function in the computational model, so prices are determined endogenously and fluctuate with regard to the aggregate capital and labor.

1.4.4 Government Policy

The government consumes in an unproductive sector, G_t .²¹ The government has two fiscal instruments to finance this consumption. First, the government taxes capital income, $y_k \equiv r_t(a + Tr_t)$, according to a capital income tax schedule $T^K[y_k]$. Second, the government taxes each individual's taxable labor income. Part of the pre-tax labor income is accounted for by the employer's contributions to social security, which is not taxable under current US tax law. Therefore, the taxable labor income is $y_l \equiv w_t s_j h_j (1 - .5\tau_{ss})$, which is taxed according to a labor income tax schedule $T^l[y_l]$. I impose three restrictions on the labor and capital income tax policies. First, I assume human capital is unobservable, meaning that the government cannot tax human capital accumulation. Second, I assume the rates cannot be age-dependent. Third, both of the taxes are solely functions of the individual's relevant taxable income in the current period.

²¹Including G_t such that it enters the agent's utility function in an additively separable manner is an equivalent formulation.

In addition to raising resource for consumption in the unproductive sector, the government runs a pay-as-you-go (PAYGO) social security system. I include a simplified social security program in the model because chapter 2 demonstrates that excluding this type of program causes the government to include a negative tax on capital in order to mimic this welfare improving program. In the reduced form social security program, the government pays SS_t to all individuals that are retired. Social security benefits are determined such that retired agents receive an exogenously set fraction, b_t , of the average income of all working individuals. An agent's social security benefits are independent of his personal earnings history. Social security is financed by taxing labor income at a flat rate, $\tau_{ss,t}$. The payroll tax rate $\tau_{ss,t}$ is set to assure that the social security system has a balanced budget each period. The social security system is not considered part of the tax policy that the government optimizes.

1.4.5 Definition of Stationary Competitive Equilibrium

In this section I define the competitive equilibrium for the exogenous model. See appendix 1.10.2 for the definition of the competitive equilibriums in the endogenous models.

Exogenous Model

Given a social security replacement rate b , a sequence of exogenous age-specific human capital $\{\epsilon_j\}_{j=1}^{j_r-1}$, government expenditures G , and a sequence of population shares $\{\mu_j\}_{j=1}^J$, a stationary competitive equilibrium in the exogenous model consists of the following: a sequence of agent allocations, $\{c_j, a_{j+1}, h_j\}_{j=1}^J$, a production plan for the firm (N, K) , a government labor tax function $T^l : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, a government capital tax function $T^k : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, a social security tax rate τ_{ss} , a utility function $U : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$, social security benefits SS , prices (w, r) , and transfers Tr such that:

1. Given prices, policies, transfers, and benefits, the agent maximizes the

following

$$\sum_{j=1}^J \text{Max}_{c_j, h_j, a_{j+1}} \beta^{j-1} \left[\prod_{q=0}^{j-1} \Psi_q \right] u(c_j, h_j)$$

subject to

$$\begin{aligned} c_j + a_{j+1} &= w\epsilon_j h_j - \tau_{ss} w\epsilon_j h_j + (1+r)(a_j + Tr) \\ &\quad - T^l [w\epsilon_j h_j (1 - .5\tau_{ss})] - T^k [r(a_j + Tr)], \end{aligned}$$

for $j < j_r$, and

$$c_j + a_{j+1} = SS + (1+r)(a_j + Tr) - T^k [r(a_j + Tr)],$$

for $j \geq j_r$. Additionally,

$$c \geq 0, 0 \leq h \leq 1,$$

$$a_j \geq 0, a_1 = 0.$$

2. Prices w and r satisfy

$$r = \alpha \left(\frac{N}{\bar{K}} \right)^{1-\alpha} - \delta$$

$$w = (1 - \alpha) \left(\frac{K}{N} \right)^\alpha$$

3. The social security policies satisfy

$$SS = b \frac{wN}{\sum_{j=1}^{j_r-1} \mu_j}$$

$$\tau_{ss} = \frac{SS \sum_{j=j_r}^J \mu_j}{w \sum_{j=1}^{j_r-1} \mu_j}$$

4. Transfers are given by

$$Tr = \sum_{j=1}^J \mu_j (1 - \Psi_j) a_{j+1}$$

5. Government balances its budget

$$G = \sum_{j=1}^J \mu_j T^k [r(a_j + Tr)] + \sum_{j=1}^{j_r-1} \mu_j T^l [w \epsilon_j h_j (1 - .5\tau_{ss})]$$

6. The market clears

$$K = \sum_{j=1}^J \mu_j a_j$$

$$N = \sum_{j=1}^J \mu_j \epsilon_j h_j$$

$$\sum_{j=1}^J \mu_j c_j + \sum_{j=1}^J \mu_j a_{j+1} + G = K^\alpha N^{1-\alpha} + (1 - \delta)K$$

1.5 Calibration and Functional Forms

In order to determine the optimal tax policy it is necessary to choose functional forms and calibrate the model's parameters. Calibrating the models involves a two step process. The first step is choosing parameter values for which there are direct estimates in the data. Second, in order to calibrate the remaining parameters, values are chosen so that under the baseline fitted US tax policy certain target values are the same in the models and in the US economy.²²

Adding endogenous human capital accumulation to the model fundamentally changes the model. Accordingly, if the calibration parameters are the same, then the value of the targets will be different in the endogenous and exogenous models. In order to assure that all the models match the targets under the baseline fitted US tax policy, I calibrate the set of parameters based on targets separately in the three models. This implies that these calibration parameters are different in the exogenous and endogenous models.

As a sensitivity analysis, I do a numerical exercise, which I call sequential parametrization, where I determine the theoretical magnitude of the effect of

²²Since these are general equilibrium models, changing one parameter will alter all the values in the model that are used as targets. However, I present targets with the parameter which they most directly correspond.

adding endogenous human capital accumulation on optimal tax policy. In this exercise I hold all the parameter values constant in the endogenous and exogenous models. Therefore, under the baseline fitted US tax policy, many of the targets do not match in the three models. I present the parameter values and results for the model under sequential parametrization in section 1.8.2.

1.5.1 Demographics

In the model agents are born at a real world age of 20 which corresponds to a model age of 1. Agents are exogenously forced to retire at a real world age of 65. If an individual survives until the age of 100, he dies the next period. I set the conditional survival probabilities in accordance with the estimates in Bell and Miller (2002). I assume a population growth rate of 1.1%. The demographic parameters are listed in table 1.1.

Table 1.1: Demographics Parameters

Parameter	Value	Target
Retire Age: j_r	65	By Assumption
Max Age: J	100	By Assumption
Surv. Prob: Ψ_j	Bell and Miller (2002)	Data
Pop. Growth: n	1.1%	Data

1.5.2 Preferences

Agents have time-separable preferences over consumption and labor services, and conditional on survival, they discount their future utility by β . I use the benchmark utility function for the exogenous and LBD models

$$\frac{c^{1-\sigma_1}}{1-\sigma_1} - \chi \frac{(h)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}}.$$

I use an altered form of this utility function for the LOD model,

$$\frac{c^{1-\sigma_1}}{1-\sigma_1} - \chi \frac{(h+n)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}}.$$

I determine β such that the capital to output ratio matches US data of 2.7.²³ I determine χ such that under the baseline fitted US tax policy, agents spend on average a third of their time endowment in non-leisure activities.²⁴ Following CKK, I set $\sigma_1 = 2$, which controls the relative risk aversion.²⁵ Past micro-econometric studies estimate the Frisch elasticity to be between 0 and 0.5.²⁶ However, more recent research has shown that these estimates may be biased downward. Reasons for this bias include: utilizing weak instruments; not accounting for borrowing constraints; disregarding the life cycle impact of endogenous-age specific human capital; omitting correlated variables such as wage uncertainty; and not accounting for labor market frictions.²⁷ Furthermore, Rogerson and Wallenius (2009) show that because individuals make decisions with regards to labor on both the intensive and extensive margins “micro and macro elasticities need not be the same, and ... macro elasticities can be significantly larger.” Therefore, I set σ_2 such that the Frisch elasticity is at the upper bound of the range for the exogenous model (0.5).²⁸ The preference parameters are summarized in table 1.2.

1.5.3 Age-Specific Human Capital

The age-specific human capital parameters that require calibration are different in the exogenous and endogenous models. In the exogenous model, I set $\{\epsilon_j\}_{j=0}^{j_r-1}$ so that the sequence matches a smoothed version of the relative hourly

²³This is the ratio of fixed assets and consumer durable goods, less government fixed assets to GDP (CKK).

²⁴Using a target of one-third is standard in quantitative exercises. For examples, see CKK, Nakajima (2010), and Garriga (2001).

²⁵Although CKK’s utility specifications are different from my benchmark, they still have a parameter that corresponds to σ_1 .

²⁶For examples see Altonji (1986), MaCurdy (1981) and Domeij and Flodén (2006).

²⁷Some of these studies include Imai and Keane (2004), Domeij and Flodén (2006), Pistaferri (2003), Chetty (2009) and Contreras and Sinclair (2008)

²⁸If endogenous human capital accumulation is added to the model, the Frisch elasticity is no longer equal to σ_2 . I elect to hold σ_2 equal in the exogenous and LBD models.

Table 1.2: Preference Parameters

Parameter	Exog.	LBD	LOD	Target
Conditional Discount: β	0.994	0.993	0.996	$K/Y = 2.7$
Unconditional Discount: $\Psi_j \beta^a$	0.981	0.980	0.983	$K/Y = 2.7$
Risk aversion: σ_1	2	2	2	CKK
Frisch Elasticity: σ_2	0.5	0.5	0.5	Frisch = $\frac{1}{2}^b$
Disutility to Labor: χ	60.9	90	70	Avg. $h_j + n_j = \frac{1}{3}$

^aSince the value varies by age, this is the average value in the economy.

^bAdding endogenous human capital accumulation changes the Frisch elasticity. The average Frisch elasticity for all working agents under the baseline fitted US tax policy in the LBD and LOD models is .3545 and .5292 respectively.

earnings estimated by age in Hansen (1993). In the LBD model, agents accumulate age-specific human capital according to the following process,

$$s_{j+1} = \Omega_j s_{j,t}^{\Phi_1} h_j^{\Phi_2}, \quad (1.40)$$

where s_j is the age-specific human capital for an agent at age j , Ω_j is an age-specific calibration parameter, Φ_1 controls the importance of an agent's current human capital on LBD, and Φ_2 controls the importance of time worked on LBD. In the LOD model, agents accumulate human capital according to the following process,

$$s_{j+1} = \Omega_j s_{j,t}^{\kappa_1} n_j^{\kappa_2}, \quad (1.41)$$

where n_j is the percent of an agent's time endowment he spends training. In this formulation, κ_1 controls the importance of an agent's current human capital on LOD and κ_2 controls the importance of time training on LOD.²⁹ In the endogenous

²⁹Guvenen et al. (2009) use an alternative LOD accumulation specification that is additively separable in past skills and training. I find that that when I use this specification an agent does not accumulate any assets for the first 10-15 years of their working life, and instead tends to save using skill accumulation. In addition, during this time agents work only the necessary hours to finance consumption causing their labor supply profile to be low and flat (see Figure 5 in Guvenen et al. (2009)). Since the shape of these life cycle profiles does not match the data, I choose not to use this functional form.

models I do not set $\{s_j\}_{j=0}^{j_r-1}$ directly, rather I calibrate the sequence $\{\Omega_j\}_{j=1}^{j_r-1}$ such that the agent's equilibrium labor or training choices cause $\{s_j\}_{j=0}^{j_r-1}$ under the baseline fitted US tax code to match the age-specific human capital calibrated in the exogenous model ($\{\epsilon_j\}_{j=0}^{j_r-1}$).

In order to calibrate the rest of the LBD parameters, I rely on the estimates in Chang et al. (2002), setting $\Phi_1 = 0.407$ and $\Phi_2 = 0.326$. Following Hansen and Imrohorglu (2009), I set $\kappa_1 = 1$ and $\kappa_2 = 0.004$ in the LOD model. The value of $\kappa_1 = 1$ implies that there is zero depreciation of human capital when skill accumulation is the result of LOD.³⁰ The values of κ_2 and $\{\Omega_j\}_{j=1}^{j_r-1}$ imply that at the start of an agent's career the ratio of time spent training to working is approximately 10% and declines steadily until retirement. Through the agent's entire working life, the ratio of the average time spent training to market hours is about 6.25%. This average value is in line with the calibration target in Hansen and Imrohorglu (2009).³¹

1.5.4 Firm

I assume the aggregate production function is Cobb-Douglas. The capital share parameter, α , is set at .36. The depreciation rate is set to target the observed investment output ratio of 25.5%. These parameters are summarized in table 1.3.

Table 1.3: Firm Parameters

Parameter	Value	Target
α	.36	Data
δ	8.33%	$\frac{I}{Y} = 25.5\%$
A	1	Normalization

³⁰See Kuruscu (2006) and Heckman et al. (1998) for other examples of quantitative studies that assume zero depreciation.

³¹Mulligan (1995) provides empirical estimates of hours spent on employer financed training which are similar to the calibration target.

1.5.5 Government Policies and Tax Functions

To calibrate parameters based on the targets, it is necessary to use a baseline tax function that mimics the US tax code so that I can find the parameter values that imply the targets in the models match the values in the data. I use the estimates of the US tax code in Gouveia and Strauss (1994) for this tax policy, which I refer to as the baseline fitted US tax policy. The authors match the US tax code to the data using a three parameter functional form,

$$T(y; \lambda_0, \lambda_1, \lambda_2) = \lambda_0(y - (y^{-\lambda_1} + \lambda_2)^{-\frac{1}{\lambda_1}}) \quad (1.42)$$

where y represents the sum of labor or capital income. The average tax rate is principally controlled by λ_0 , and λ_1 governs the progressivity of the tax policy. λ_2 is left free in order to ensure that taxes satisfy the budget constraint. Gouveia and Strauss (1994) estimate that $\lambda_0 = .258$ and $\lambda_1 = .768$ when fitting the data. The authors do not fit separate tax functions for labor and capital income. Accordingly, I use a uniform tax system on both sources of income for the baseline fitted US tax policy. I calibrate government consumption, G , so that it equals 17% of output under the baseline fitted US tax policy, as observed in the US data.³² Therefore, λ_2 is determined as the value that equates government spending to 17% of GDP. When searching for the optimal tax policy, I restrict my attention to revenue neutral changes which imply that government consumption is equal under the baseline fitted US tax policy and the optimal tax policy.

In addition to government consumption, the government also runs a balanced budget social security program. Social security benefits are set so that the replacement rate, b , is 50%.³³ The payroll tax, τ_{ss} , is determined so that the social security system is balanced each period.

³²To determine the appropriate value for calibration, I focus on government expenditures less defense consumption.

³³The replacement rate matches the rate in CKK and Conesa and Krueger (2006). The Social Security Administration estimates that the replacement ratio for the median individual is 40% (see Table VI.F10 in the 2006 Social Security Trustees Report; available at <http://www.ssa.gov/OACT/TR/TR06/>). This estimate is lower than the replacement rate I use, however, if one also includes the benefits paid by Medicare then the observed replacement ratio would be higher.

Table 1.4: Government Parameters

Parameter	Exog	LBD	LOD	Target
λ_0	.258	.258	.258	Gouveia and Strauss (1994)
λ_1	.768	.768	.768	Gouveia and Strauss (1994)
G	0.137	0.139	0.127	17% of Y
b	0.5	0.5	0.5	CKK

1.6 Computational Experiment

The computational experiment is designed to determine the tax policy that maximizes a given social welfare function. I find that the optimal flat tax rates on capital and labor for all three models. I choose a social welfare function that corresponds to a Rawlsian veil of ignorance (Rawls (1971)). Since living agents face no earnings uncertainty, the social welfare is equivalent to maximizing the expected lifetime utility of a newborn,

$$SWF(\tau_h, \tau_k) = \sum_{j=1}^J \beta^{j-1} \left[\prod_{q=0}^{j-1} \Psi_q \right] u(c_j, h_j) \quad (1.43)$$

where τ_h is the flat tax rate on labor income and τ_k is the flat tax rate on capital income.

In order to determine the effects of endogenous human capital accumulation, I compare the tax policies that maximize the SWF in the three models. When I determine the optimal tax policy, I test different values of τ_h and determine values for τ_k so that the change in the tax policy are revenue neutral. Therefore, the experiment is to find τ_h that satisfies

$$\max_{\tau_h} SWF(\tau_h, \tau_k) \quad (1.44)$$

subject to,

$$G = \sum_{j=1}^J \mu_j \tau_k r (a_j + Tr) + \sum_{j=1}^{j_r-1} \mu_j \tau_h [ws_j h_j (1 - .5\tau_{ss})] \quad (1.45)$$

1.7 Results

In this section I quantitatively assess the effects on the optimal tax policy of including endogenous age-specific human capital accumulation in a life cycle model. I begin by determining the optimal tax policies in the exogenous, LBD and LOD models and then highlight the channels that cause the differences.

To fully understand the effects of endogenous human capital accumulation, I analyze the aggregate economic variables and life cycle profiles in all three models. I compare the aggregate economic variables and life cycle profiles in all three models under the baseline fitted US tax policy as well as the changes induced by implementing the optimal tax policies.

1.7.1 Optimal Tax Policies in Exogenous, LBD, and LOD Models

Table 1.5 describes the optimal tax policies in the three models. The optimal tax policy in the exogenous model is an 11.8% flat tax on capital income ($\tau_k = 11.8\%$) and a 24.7% flat tax on labor income ($\tau_h = 24.7\%$).³⁴ While the optimal tax on capital is much smaller in the exogenous model compared to CKK, it is not zero. The motives that cause a positive tax on capital in the exogenous model include: the inability of the government to borrow; agents being liquidity constrained and the government not being able to tax transfers at a separate rate from ordinary capital income. See chapter 2 for a thorough discussion of the relative strengths of these motives in a model similar to the exogenous model.

³⁴I checked whether a progressive tax on either capital or labor was optimal. However, I found that the optimal tax policies were always flat taxes. This result is similar to CKK who find that the optimal tax policies are flat in their model that is similar to the exogenous model. CKK find that a progressive tax on labor income is optimal only if the model includes within cohort heterogeneity. Since all the agents within a cohort are homogenous in my models, one would expect flat taxes to be optimal.

The optimal tax policy in the LBD is $\tau_k = 21.5\%$ and $\tau_h = 23.3\%$, and in the LOD model it is $\tau_k = 14.3\%$ and $\tau_h = 24.3\%$. Including either form of endogenous human capital accumulation increases the optimal tax on capital. The optimal tax on capital is more than eighty percent larger in the LBD model and more than twenty percent larger in the LOD compared to the exogenous model.

Table 1.5: Optimal Tax Policies in Benchmark Models

Tax Rate	Exog	LBD	LOD
τ_k	11.8%	21.5%	14.3%
τ_h	24.7%	23.3%	24.3%
$\frac{\tau_k}{\tau_h}$	0.48	0.92	0.59

With respect to LBD, the alteration in the Frisch labor supply elasticity profile is the principal reason that the optimal tax on capital increases. Figure 1.1 plots the lifetime Frisch labor supply elasticities in the LBD model and the exogenous model under the optimal tax policy. The lifetime labor supply elasticity is flat in the exogenous model and upward sloping in the LBD model. Adding LBD causes agents to supply labor relatively more elastically as they age because the human capital benefit decreases. The optimal tax on capital is higher in the LBD model in order to implicitly tax younger agents, who supply labor less elastically, at a higher rate.

In order to quantify the impact of the elasticity channel on the optimal tax policy in the LBD model, I alter the exogenous model so that the shape of the lifetime Frisch labor supply elasticity profile is the same as it is in the LBD model under the optimal tax policy. In order to match the shapes of the profiles, I vary σ_2 in the exogenous model by age.³⁵ I find that the optimal tax policy in this altered exogenous model is $\tau_k = 21\%$ and $\tau_h = 22.8\%$. The optimal tax policy in the LBD and the altered exogenous models are nearly identical, indicating that the larger optimal tax on capital in the LBD model is due to the elasticity channel.

³⁵I normalize the values of σ_2 in this altered model such that the average of the values is still 0.5.

In section 1.3.3 I show that both the elasticity channel and the savings channel affect the optimal tax on capital in the LOD model in opposite directions. The elasticity channel in the LOD model causes young agents to supply labor relatively more elastically. Figure 1.2 plots the Frisch elasticity profile in the exogenous model and the LOD model. This channel causes the optimal tax on capital to decrease, which implicitly taxes the younger agents who supply labor more elastically at a lower rate. However, with LOD, training is a form of savings which activates the savings channel. The savings channel causes the optimal tax on capital to increase in order to encourage agents to save via human capital accumulation as opposed to physical capital. The increase in the optimal tax on capital in the LOD model indicates that the savings channel dominates.

In order to quantify the impact of each channel I solve for the optimal tax policy in an alternative version of the LOD model that excludes the elasticity channel. I solve the LOD model with an alternative utility function, $\frac{c^{1-\sigma_1}}{1-\sigma_1} - \chi_1 \frac{(h)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} - \chi_2 \frac{(t)^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}}$, which is separable in training and hours worked. Since the utility function is separable, the Frisch elasticity is no longer a function of the time spent training and is constant at the value σ_2 . Therefore, using this utility function eliminates the elasticity channel. The optimal tax policy in this model with the alternative utility function is $\tau_k = 16.4\%$ and $\tau_h = 24\%$. These results indicate that the elasticity channel causes the optimal tax on capital to decrease 2.1 percentage points and that the savings channel causes it to increase by 4.6 percentage points compared to the exogenous model.

1.7.2 The Effects of Adding Endogenous Age-Specific Human Capital

This section analyzes the impact on the aggregate economic variables and life cycle profiles from adding LBD and LOD to the exogenous model under the baseline fitted US tax policy. Figure 1.3 plots the life cycle profiles of hours, consumption, assets and age-specific human capital in all three models. Table 1.6 describes the optimal tax policies and summarizes the aggregate economic

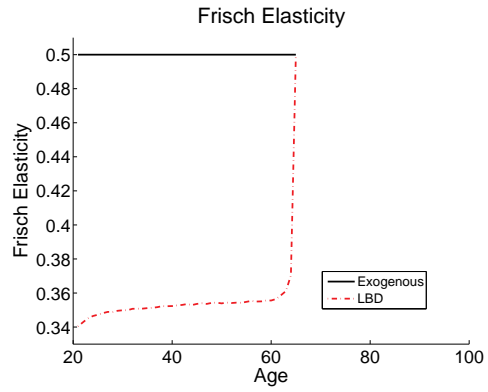


Figure 1.1: Life Cycle Frisch Labor Supply Elasticity in LBD Model

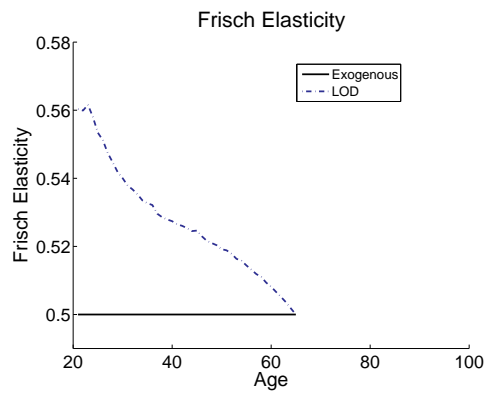


Figure 1.2: Life Cycle Frisch Labor Supply Elasticity in LOD Model

variables under both the baseline fitted US tax policy and optimal tax policies. The first, fourth, and seventh columns are the aggregate economic variables under the baseline fitted US tax policy in the exogenous, LBD, and LOD models, respectively. The second, fifth, and eighth columns are the aggregate economic variables under the optimal tax policies. The third, sixth, and ninth columns are the percentage changes in the aggregate economic variables induced from adopting the optimal tax policies.

Table 1.6: Aggregate Economic Variables

Aggregate	Exogenous			LBD			LOD		
	Baseline	Optimal	<i>% Change from Baseline to Optimal</i>	Baseline	Optimal	<i>% Change from Baseline to Optimal</i>	Baseline	Optimal	<i>% Change from Baseline to Optimal</i>
Y	0.79	0.82	3.5%	0.81	0.82	0.6%	0.75	0.77	2.7%
K	2.14	2.28	6.9%	2.16	2.21	2.5%	2.03	2.14	5.4%
N	0.46	0.46	1.6%	0.47	0.47	-0.4%	0.43	0.43	1.2%
Avg Hours	0.33	0.34	0.7%	0.34	0.34	0.2%	0.31	0.32	0.6%
w	1.12	1.14	1.8%	1.11	1.12	1.0%	1.12	1.14	1.5%
r	0.05	0.05	-8.4%	0.05	0.05	-4.7%	0.05	0.05	-6.8%
tr	0.03	0.03	7.0%	0.02	0.03	5.5%	0.03	0.03	6.6%
Value	-136.58	-135.85	0.5%	-148.93	-146.91	1.4%	-155.68	-154.74	0.6%
CEV			0.7%			2.0%			0.8%
Average Tax Rate	Baseline	Optimal		Baseline	Optimal		Baseline	Optimal	
Capital	14.9%	11.8%		17.6%	21.5%		15.0%	14.3%	
Labor	23.4%	24.7%		24.4%	23.3%		23.5%	24.3%	
Ratio	0.64	0.48		0.72	0.92		0.64	0.59	
Marginal Tax Rate	Baseline	Optimal		Baseline	Optimal		Baseline	Optimal	
Capital	18.9%	11.8%		21.1%	21.5%		18.9%	14.3%	
Labor	25.4%	24.7%		25.6%	23.3%		25.4%	24.3%	
Ratio	0.74	0.48		0.82	0.92		0.74	0.59	

Note: The average hours refers to the average percent of time endowment worked in the productive labor sector. Both the marginal and average tax rates vary with income under the baseline fitted US tax policy. The numbers reported are the population weighted averages.

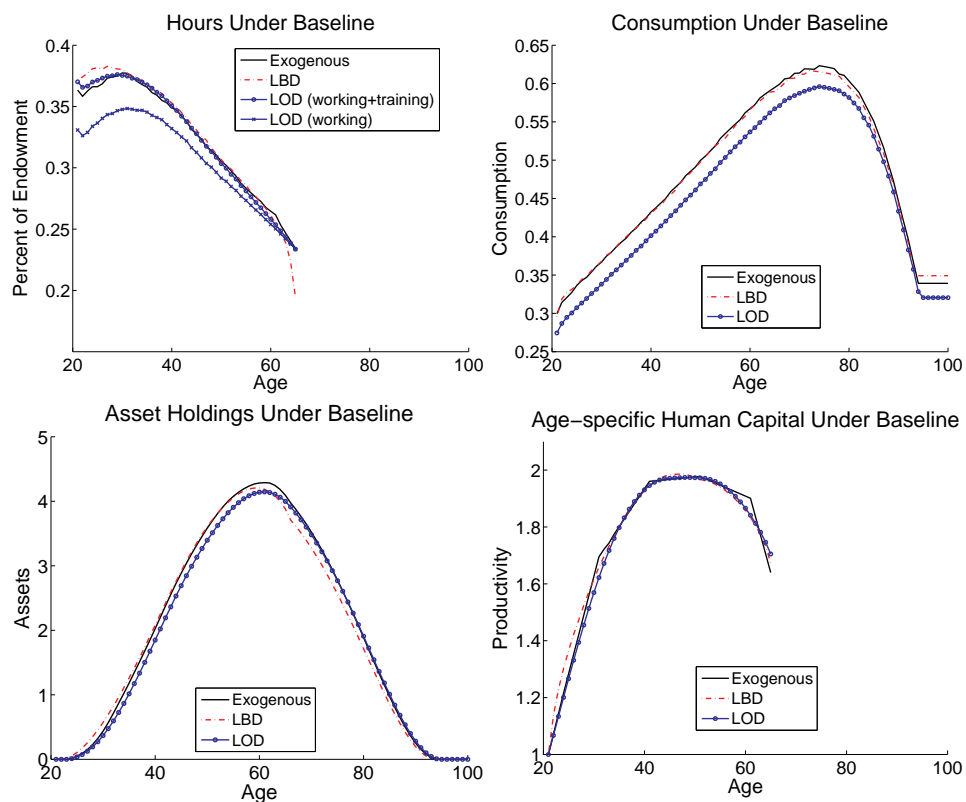
Comparing the first and fourth columns of table 1.6, it is clear that the levels of aggregate hours, labor supply, and aggregate capital are similar in the exogenous and LBD models. The calibrated parameters are determined so that under the baseline fitted US tax policy the models match certain targets from the data. Therefore, since many of the aggregate economic variables are targets and these calibration parameters are determined separately in the exogenous and LBD

models, the aggregates are similar in the two models.

Although adding LBD does not have a large impact on the aggregate economic variables, it does cause the life cycle profiles to differ in the models. Adding LBD causes agents to work relatively more at the beginning of their working life when the human capital benefit is larger, and less later in their working life when the benefit is smaller (see the solid black and dashed red lines in the upper left panel of figure 1.3). The upper right panel shows that the lifetime consumption profiles are similar in the exogenous and LBD models. Since adding LBD causes agents to work relatively less time in the middle of their lifetime compared to the exogenous model, agents' savings are also relatively smaller for the second half of their lifetime (see the lower left panel). The lifetime age-specific human capital profiles are similar in the exogenous and LBD models since the sequence of parameters $\{\Omega_j\}_{j=1}^{j_r-1}$ is calibrated so that age-specific human capital matches (see the lower right panel of figure 1.3).

The aggregate economic variables are not similar in the exogenous and LOD models because agents must spend time training in the LOD model. Comparing the first and seventh columns of table 1.6, aggregate hours, labor supply and capital are smaller in the LOD model because an agent spends part of his time endowment training. However, it is apparent that the relative ratios of the aggregates are similar in the two models since the factor prices are comparable.

Adding LOD also affects the life cycle profiles. Figure 1.3 plots two labor supply profiles for the LOD model - the first is solely hours spent working, and the second is the sum of hours spent working and training (see the blue lines in the upper left panel). The LOD labor supply profile that includes training is similar to the exogenous model; however the profile that excludes training is smaller. The difference between the two profiles is the amount of time spent training in the LOD model. It is clear that this gap shrinks as an agent ages, representing a decrease in the amount of time spent training. Agents spend less time training because the benefit decreases as they age since they have fewer periods to take advantage of their human capital. Since adding LOD causes the size of the economy to decrease, the life cycle profile for consumption also decreases. In the LOD model,



Note: These plots are life cycle profiles of the three calibrated models under the baseline fitted US tax policy. The upper left hand panel is a plot of the labor supply, the upper right hand panel is a plot of consumption, the bottom left panel is a plot of assets and the bottom right is a plot of skills. The solid black lines are in the exogenous model, the dashed red lines are in the LBD model and the blue marked lines are in the LOD model. There are two labor lines for the LOD model, one solely for hours worked and the other for hours worked plus hours spent training.

Figure 1.3: Life Cycle Profiles under Baseline Fitted US Tax Policy

agents can use their time endowment to accumulate human capital, which acts as an alternative form of savings from assets. Therefore, during their working lives, agents hold fewer ordinary capital assets, and opt instead to use human capital to supplement their savings. As an agent approaches retirement the value of the human capital decreases and the asset profile in the LOD converges to the profile in the exogenous model. Finally, similar to LBD, the lifetime age-specific human capital profiles are similar in the exogenous and LOD models since the profiles are a calibration target.

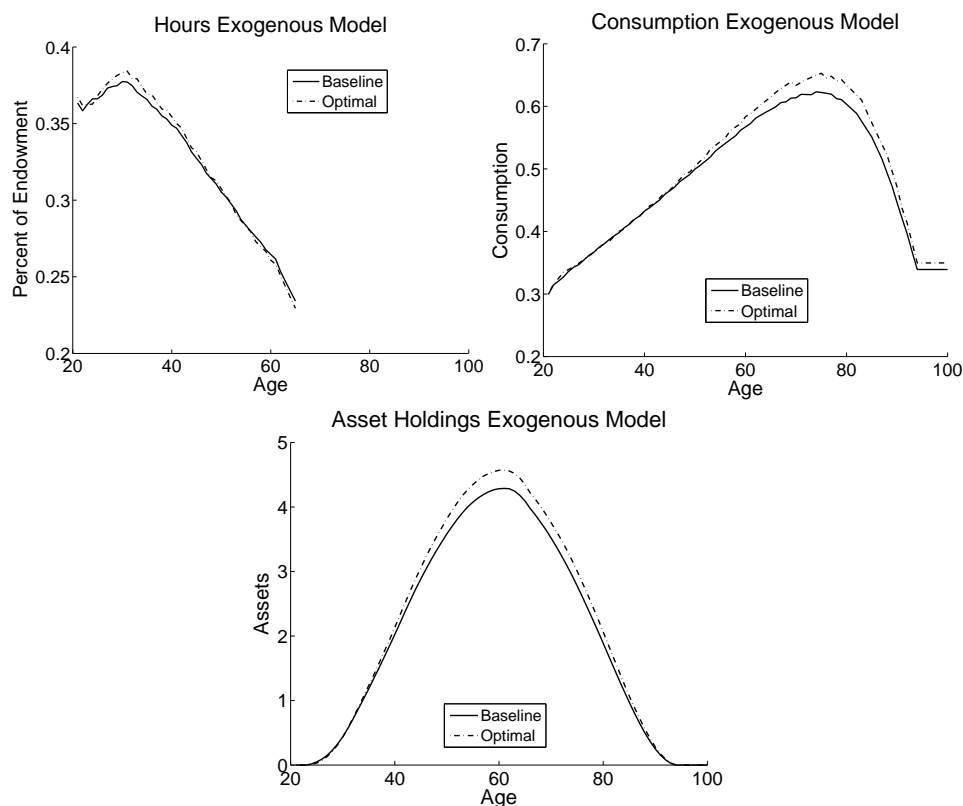
1.7.3 The Effects of the Optimal Tax Policy in the Exogenous Model

In the exogenous model, the optimal tax on capital is smaller than the tax under the baseline fitted US tax policy. Because of the smaller tax on capital, adopting the optimal tax policy causes an increase in aggregate capital (see columns one and two of figure 1.6). Since the optimal tax on labor is flat and the baseline is progressive, the average marginal tax on labor is less under the optimal tax policy. Therefore, agents work longer under the optimal tax policy. Although both average hours and the aggregate labor supply increase, they increase relatively less than the aggregate level of capital. Since aggregate labor increases relatively less than aggregate capital, the rental rate on capital decreases and the wage rate increases.

In order to compare the welfare effects of adopting the optimal tax policies in the models, I compute the consumption equivalent variation (CEV). The CEV is the uniform percentage increase in consumption, at each age, needed to make an agent indifferent between being born under the baseline fitted US tax policy and the optimal tax policy. Therefore, a positive CEV indicates a welfare increase due to tax reform. Overall, adopting the optimal tax policy in the exogenous model causes a welfare increase of 0.7% CEV.³⁶

³⁶Adopting the optimal tax policy causes two changes. First, it eliminates the progressivity of the baseline fitted US tax policy. Additionally, it adjusts the relative ratio of the tax on capital to labor. In order to test the relative significance of each change, I find the CEV between the baseline tax policy and a flat tax policy that raises the same amount of revenue from each income source as under the baseline tax policy. I find that eliminating progressivity represents a majority of the increase in the CEV.

Figure 1.4 plots the life cycle profiles for time worked, consumption, assets and age-specific human capital in the exogenous model under the baseline fitted US tax policies and the optimal tax policies. The solid lines are the profiles under the baseline fitted US tax policies and the dashed lines are the profiles under the optimal tax policies. Adopting the optimal tax policy in the exogenous model causes changes in all three life cycle profiles: (i) agents work relatively more early in their life; (ii) agents save more, especially during periods when they are wealthier; and (iii) the lifetime consumption profile steepens.



Note: The upper left hand panel is a plot of the labor supply, the upper right hand panel is a plot of consumption, and the bottom left panel is a plot of assets. Since the skills are the same in the exogenous models under the baseline fitted US tax policy and optimal tax policy they are not plotted. Solid lines are under the baseline fitted US tax policy while dashed lines are under the optimal tax system.

Figure 1.4: Life Cycle Profiles in the Exogenous Model

Comparing the profiles in the left panel of figure 1.4, it is evident that agents shift time worked from later to earlier years. This shift is a consequence of

the lower implicit tax on young labor income due to a decrease in the tax rate on capital income.

Implementing the optimal tax policy causes a decrease in both the tax on capital and the rental rate on capital. These changes have competing effects on the marginal after-tax return on capital. The drop in the tax rate is larger than the drop in the rental rate on capital so the average marginal after-tax return increases causing agents to save more.³⁷ The drop in the tax rate does not have a uniform effect on the agent's net return since the baseline fitted US tax on capital is progressive. The decrease is larger for agents who hold more savings since their marginal tax rate was higher under the progressive baseline fitted US tax policy. Therefore, the increase in savings is even larger for agents who hold more savings under the baseline fitted US tax policy (see the lower left panel of figure 1.4).

The change in the marginal after-tax return from adopting the optimal tax policy also affects the shape of the lifetime consumption profile. The intertemporal Euler equation controls the slope of consumption profile over an agent's lifetime. The relationship is,

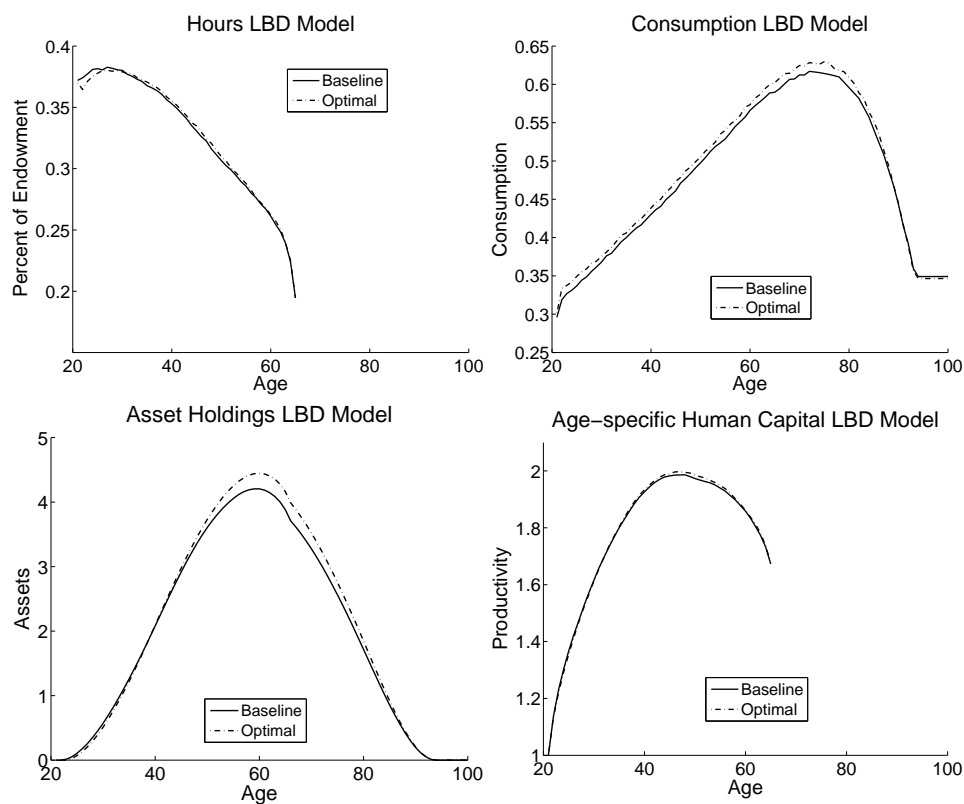
$$\left(\frac{c_{j+1}}{c_j}\right)^{\sigma_1} = \Psi_j \beta \tilde{r}_t \quad (1.46)$$

where \tilde{r}_t is the marginal after-tax return on capital. The marginal after-tax return on capital is larger in the optimal model, which means that the consumption profile (figure 1.4, upper right panel) is steeper under the optimal tax policy. Additionally, since the increase in the marginal after-tax return on capital is larger when agents hold more assets, the effect on the profile is more pronounced for ages when agents hold more assets.

1.7.4 The Effects of Optimal Tax Policy in the LBD Model

Since adopting the optimal tax policies causes the capital tax to change in different directions in the exogenous and LBD models, the aggregate economic variables react differently. The third and sixth column of table 1.6 describe the

³⁷When I describe the overall change in the average marginal after-tax return on capital, I am referring to the population weighted average marginal tax on capital.



Note: The upper left hand panel is a plot of the labor supply, the upper right hand panel is a plot of consumption, the bottom left panel is a plot of assets, and the bottom right panel is a plot of age-specific human capital. Solid lines are plotted using the baseline fitted US tax policy and dashed lines are plotted under the optimal tax system.

Figure 1.5: Life Cycle Profiles in the LBD Model

percentage changes in the aggregate economic variables induced by the optimal tax policies in the exogenous and LBD models, respectively. The optimal tax on capital is relatively lower in the exogenous model compared to the LBD model. Therefore, implementing the optimal tax policy causes a much smaller increase in the capital stock in the LBD model compared to the exogenous model. Neither average hours nor aggregate labor supply change by a significant amount as a result of adopting the optimal tax policy in the LBD model. The modest rise in aggregate capital and stable aggregate labor supply in the LBD model translates into an increase in the wages and a decrease in the rental rate on capital. The more dramatic rise in capital from adopting the optimal tax policy in the exogenous model induces a larger increase in the rental rate on capital and a larger decrease in the wage than in the LBD model. Despite the smaller change in prices, the CEV in the LBD model is 2.0%, indicating that the welfare gain from adopting the optimal tax policy is greater in the LBD model than in the exogenous model.

Implementing the optimal tax policies also causes different changes in the life cycle profiles in the exogenous and LBD models (see figures 1.4 and 1.5). Implementing the optimal tax policy in the LBD model causes changes in all three life cycle profiles: (i) agents shift time worked from early to later years; (ii) there is a uniform upward shift in the lifetime consumption profile; and (iii) the lifetime savings profile moves in opposite directions over the agent's lifetime.

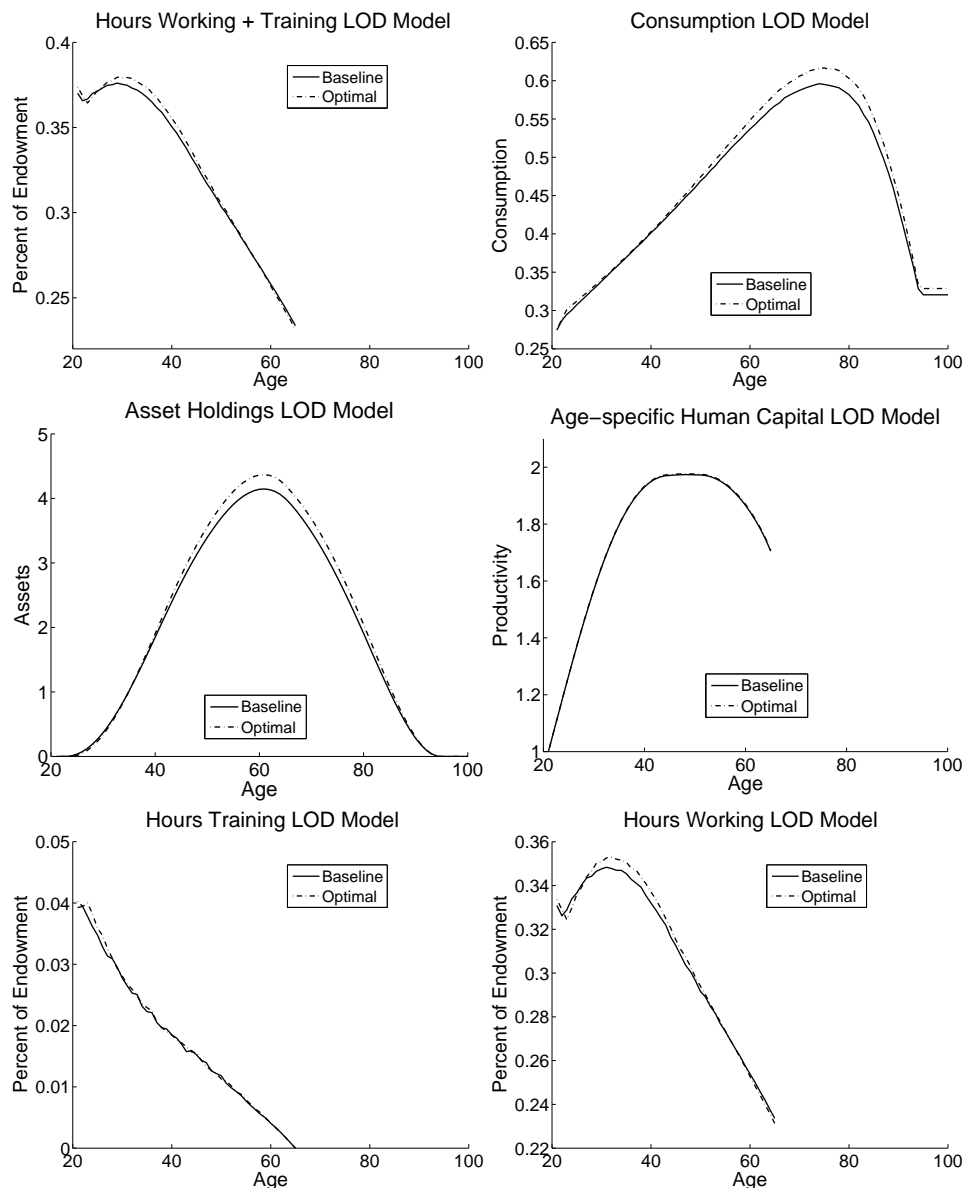
Adopting the optimal tax policy in the LBD model causes agents to shift hours from early in their lifetime to the remainder of their working years (see the upper right panel of figure 1.5). In the LBD model, adopting the optimal tax policy causes capital to be taxed at a relatively higher rate that implicitly taxes labor income from early years at a higher rate. This change results in the shift of time worked from earlier to later years. In the exogenous model, implementing the optimal tax policy decreases the tax on capital; agents accordingly shift time worked in the opposite direction.

Applying the optimal tax policy in the LBD model introduces two opposing effects on the agent's lifetime asset profile. First, in the LBD model, in which the economy is larger under the optimal tax policy compared to the fitted US tax policy,

the agents hold more assets. Second, the optimal tax policy decreases the average marginal after-tax return on capital, causing agents to hold fewer assets. The first effect is constant for all agents. The second effect is not constant for all agents, but it is negatively proportional to an agent's capital income because the baseline fitted US tax policy is progressive and the optimal tax policy is flat. The progressive baseline tax rate means that the increase in the tax rate on capital income under the optimal policy is less for agents who have higher capital income. Therefore, the overall decrease in the marginal after-tax return on capital from adopting the optimal tax policy is smaller for agents who have more savings. This means that the second effect is relatively stronger when agents save less and relatively weaker when they save more. As is apparent in the lower left panel of figure 1.5, adopting the optimal tax policy in the LBD model causes agents to save less at ages when they had lower savings under the baseline fitted US tax policy (early and late in life), and to save more at ages when they held larger savings under the baseline fitted US tax policy (in the middle of their life). This outcome is in contrast to the exogenous model, in which the optimal tax policy causes the assets profile to shift upward for all agents since the after tax return increases.

In the LBD model, implementing the optimal tax policy causes the consumption profile to uniformly shift upward (see the upper right panel). The profile shifts upward due to an increase in the overall size of the economy. The profile also shifts upward in the exogenous model, but the shift is not uniform.³⁸

In the LBD model the optimal tax policy causes agents to work more in their middle years. This shift in hours translates in higher age-specific human capital during those years (see the lower right panel). In the exogenous model agents cannot affect their level of age-specific human capital; implementing the optimal tax policy thus has no affect on the age-specific human capital profile.



Note: The upper left hand panel is a plot of the sum of labor and training, the upper right hand panel is a plot of consumption, the middle left panel is a plot of assets, the middle right panel is a plot of age-specific human capital, the bottom right panel is a plot of time spent training, and the bottom right is a plot of time spent working. Solid lines are plotted using the baseline fitted US tax policy, while dashed lines are plotted using the optimal tax system.

Figure 1.6: Life Cycle Profiles in the LOD Model

1.7.5 The Effects of Optimal Tax Policy in the LOD Model

Although the optimal tax on capital is larger in the LOD model than in the exogenous model, the changes in the tax rates from adopting the optimal tax policy are similar in the two models: a decrease in the tax on capital and an increase in the tax on labor. Therefore, the aggregate economic variables respond in a similar fashion in both models: capital increases, labor increase, wages increase and the rental rate decreases. When adopting the optimal tax policy, the drop in the capital tax is smaller in the LOD model, so the effects on the aggregate economic variables are muted. However, the effect on welfare is similar to the exogenous model, an increase of 0.8% CEV.

Adopting the optimal tax policy in the LOD induces changes in the life cycle profiles much like those in the exogenous model (see figures 1.4 and 1.6): (i) agents shift hours worked to earlier in their life, (ii) agents increase their savings, and (iii) agents increase their consumption at a faster rate throughout their life. The tax on capital is smaller in the optimal tax policy compared to the tax rate in baseline fitted US tax policy, meaning that the implicit tax on young labor income is smaller than old labor income. Therefore, agents react by shifting hours worked to earlier in their lifetime (see lower right panel of figure 1.6). Additionally, since the implicit tax goes down on young labor income, agents find it more valuable to acquire human capital when they are young so they spend more time training early in their lifetime (see lower left panel). This shift also results in agents having more human capital in the middle of their life (see the middle right profile).

As with the exogenous model, adopting the optimal tax policy causes a decrease in both the tax on capital and the rental rate on capital. These have counteracting effects on the agent's savings decisions. During the early and later years of an agent's life, the tax on capital falls less since the baseline fitted US tax policy is progressive; therefore the decrease in the rental rate dominates and the agent holds less savings. In the middle of an agent's life the tax on capital is larger under the baseline fitted US tax policy, so the drop in the tax from adopting the

³⁸The shift is uniform in the LBD model because the average marginal tax on capital is similar under the baseline fitted US tax policy and the optimal tax policy.

optimal policy dominates and agents hold more savings.

Adopting the optimal tax policy causes an agent's consumption profile to be steeper in both the exogenous and LOD models. The slope of the profile is controlled by the after tax return on capital. Therefore, the change in the slope is more pronounced for ages when agents hold more assets.

1.8 Sensitivity Analysis

Next I check the sensitivity of the results with respect to the utility specification and the procedure used to determine the values for the calibrated parameters.

1.8.1 Non-Separable Utility

In this section I determine the quantitative effects of adding endogenous human capital accumulation to the exogenous model with an alternative utility function,

$$U(c_{1,t}, 1 - h_{1,t}) = \frac{(c_{1,t}^\gamma (1 - h_{1,t})^{1-\gamma})^{1-\varsigma}}{1 - \varsigma}.$$

This utility function is the benchmark specification in CKK. I refer to this utility function as the non-separable utility function. This function includes two additional motives for a positive tax on capital. Atkeson et al. (1999), Erosa and Gervais (2002), Garriga (2001), and CKK demonstrate that non-separability creates an additional motive for a positive tax on capital. Additionally, under this utility specification, the labor supply elasticity is a negative function of hours worked. An agent's labor supply elasticity profile tends to slope upwards in simulations using this utility function since their labor supply profile generally slopes downward. The optimal tax on capital is therefore larger in order to implicitly tax younger labor income that is supplied less elastically at a higher rate. In the exogenous model the optimal tax on capital is also larger under the non-separable utility function than under the benchmark specification. I begin by presenting the new calibration parameters followed by the optimal tax policies in all three models.

Changes in Calibration

The non-separable utility function requires calibration of two new parameters. The new parameters are γ , which determines the comparative importance of consumption and leisure, and ς , which controls risk aversion. I can no longer target both the Frisch elasticity and average time worked since γ controls both of these values. Therefore, I calibrate γ to target the percentage of the time endowment worked and no longer use the Frisch elasticity as a target.

Table 1.7 lists the calibration parameters for the non-separable utility parameters and the Frisch elasticity (when hours = $\frac{1}{3}$). The Frisch elasticity for this utility function is $\frac{(1-h)}{h} \frac{1-\gamma(\varsigma-1)}{\varsigma}$ in the exogenous model. The Frisch elasticity is a decreasing function in hours, meaning it is no longer constant in the exogenous model (as long as hours worked vary over the lifetime). Additionally, the average Frisch elasticity implied by the calibration is more than twice as large as with the benchmark utility specification in the exogenous model. However, section 1.5.2 expresses reasons why a larger Frisch elasticity may be in line with unbiased empirical estimates.

Table 1.7: Preference Parameters

Parameter	Exog	LBD	LOD	Target
β	1.009	1.009	1.013	$K/Y = 2.7$
$\Psi_j \beta^a$	0.996	0.996	1.000	$K/Y = 2.7$
γ	0.36	0.27	0.34	Avg. $h_j + n_j = \frac{1}{3}$
ς	4	4	4	CKK
Frisch elasticity($h = \frac{1}{3}$)	1.29			

^aSince the value varies by age, this is the average value in the economy.

Adding LBD and LOD changes workers' incentives and choice variables, and the relevant preference parameters change accordingly. The main difference is that in the LBD model agents enjoy the human capital benefit. Because working is more valuable in the LBD model, the value for γ is smaller. In order to calibrate the model so that agents work equal percentages of their time endowment in the

exogenous and LBD models, γ must be lower in the LBD model to offset the extra benefit of working. A lower γ decreases the relative importance of consumption compared to leisure, which implies a lower Frisch elasticity in the LBD model than in the exogenous model. The rest of the parameters are similar in the exogenous and LBD models.

Adding LOD allows agents to acquire human capital which provides agents with an alternative method of saving. Since agents use human capital as part of their savings, in order to induce the ratio $\frac{K}{Y}$ to be the same in the LOD model β must be higher. A higher value for β encourages agents to be more patient, so they place a higher value on future consumption. In order to finance future consumption, agents increase their savings and the ratio increases. A higher value for β also implies that agents will spend more time training. Since γ is set in order to target the sum of time spent training and working, the value for γ must drop in order to keep agents working and training for one-third of their endowment.

Optimal Tax Policies in Non-Separable Models

The optimal tax policies for the non-separable model are listed in table 1.8. The optimal tax policy for the exogenous model is $\tau_k = 35.3\%$ and $\tau_h = 18.3\%$. The optimal tax policy in the LBD model is $\tau_k = 46.3\%$ and $\tau_h = 15.0\%$, and the optimal tax policy in the LOD model is $\tau_k = 36.5\%$ and $\tau_h = 18.7\%$. Once again, including LBD leads to an increase in the optimal tax on capital because it causes agents to become relatively more elastic over their lifetime, as the human capital benefit decreases. Although optimal tax policy reacts less to LBD with non-separable utility, the optimal tax on capital is still over thirty percent higher in the LBD model compared to the exogenous model. Adding LOD causes a small increase in the optimal tax on capital. In the case of LOD, the large tax on capital in the exogenous model means that altering the tax policy is less effective in increasing the incentives to save via human capital, so the savings channel is less important. Overall, even with the non-separable utility - in which the exogenous model contains a large motive for a tax on capital - adding either form of endogenous human capital causes the optimal tax on capital to increase. See

appendix 1.10.3 for the details of the economy in the non-separable models.

Table 1.8: Optimal Tax Policies in Non-Separable Models

Tax Rate	Exog	LBD	LOD
τ_k	35.3%	46.3%	36.5%
τ_h	18.3%	15.0%	18.7%
$\frac{\tau_k}{\tau_h}$	1.93	3.09	1.95

1.8.2 Sequential Parametrization

In this section I present the parametrization and results from the endogenous model with the benchmark utility function and sequential parametrization. This exercise is a numerical example which determines the theoretical magnitude of the effects of adding LBD and LOD on optimal tax policy. The sequential parametrization uses the parameter values from the exogenous model in the endogenous models. Since the parameter values are the same but the models are different, many of the targets under the baseline fitted US tax policy in the exogenous and LBD model vary.

Optimal Tax Policies in Sequential Parametrization

In the sequential parametrization, the parameter governing the disutility to labor, χ , is the same in both the exogenous and endogenous models. χ is determined by targeting the percentage of the time endowment that agents work in the exogenous model. However, in the LBD model agents receive the extra human capital benefit to working. Therefore, agents will generally work more in the LBD model than in the exogenous model. This translates into a larger economy in the LBD model than in the exogenous model. In the LOD model, agents generally spend the same amount of time on leisure as they do in the exogenous model. However, in the LOD model some of their non-leisure time is spent training whereas in the exogenous model it is all spent working. Therefore,

agents spend less time providing labor services in the LOD model than in the exogenous model. Accordingly, the economy is smaller in the LOD model than in the exogenous model.

G is held constant across models, but since there are changes to the relative sizes of the economy, the percentage of the economy that G accounts for is different in the three models. Adding LBD increases the size of the economy so G is a smaller percentage of output in the LBD model than in the exogenous model. Adding LOD has the opposite effect. The economy is smaller in the LOD model than in the exogenous model, so G is a larger percentage of output when compared to the exogenous model. Since the tax bases are different in the three models, holding all else equal, the tax rates in the LBD model will be smaller and the tax rates in the LOD model will be larger. Due to differing tax bases in the models, the most appropriate metric for comparison is the ratio of the optimal capital tax rate to the optimal labor tax rate.

The optimal tax policies are listed in Table 1.9. As with the calibrated model, adding both forms of endogenous human capital accumulation increases the optimal tax on capital. Comparing the ratio of the optimal tax on capital to the optimal tax on labor in the calibrated and sequentially parameterized models demonstrates that the magnitude of the effects of adding endogenous human capital are similar (see Table 1.5 and 1.9). Therefore, one can conclude that the results are not sensitive to the procedure used to determine the calibration parameters.

Table 1.9: Optimal Tax Policies in Sequential Parametrization

Tax Rate	Exog	LBD	LOD
τ_k	11.8%	22.3%	17.8%
τ_h	24.7%	19.5%	24.1%
$\frac{\tau_k}{\tau_h}$	0.48	1.14	0.74

1.9 Conclusion

In this paper I characterize the optimal capital and labor tax rates in three separate life cycle models in which age-specific human capital is accumulated exogenously, endogenously through LBD and endogenously through LOD. Analytically, I demonstrate that including endogenous human capital accumulation creates a motive for the government to condition labor income taxes on age and that if age-dependent taxes are unavailable, the tax on capital increases in order to mimic relatively higher labor income taxes on young agents. Quantitatively, I find that adding endogenous human capital accumulation in the model with the benchmark utility specification increases the the optimal tax on capital by approximately 70% in the LBD framework and 15% in the LOD framework. Furthermore, I find that the direction of the impact is robust to utility specification. Even under the non-separable utility function, which already contains a large motive for a tax on capital in the exogenous model, adding either form of endogenous human capital accumulation increases the optimal tax on capital. Many of the previous computational life cycle studies model age-specific human capital exogenously. Given my findings, this assumption is not innocuous.

LBD increases the motive for a tax on capital since it alters the lifetime labor supply elasticity profile. Adding LBD to the model causes younger agents to supply labor relatively less elastically since the human capital benefit decreases over an agent's lifetime. A larger tax on capital is optimal because it implicitly taxes younger labor supply income, which is supplied less elastically, at a higher rate. Adding LOD to the model has counteracting affects on the optimal tax policy. Including LOD causes younger agents to supply labor relatively more elastically because training is an imperfect substitute for working. This change in the elasticity motivates the government to decrease the tax on capital. However, the government increases the tax on capital in order to increase the agent's incentive to use training instead of physical capital in order to save. Overall, I find that in numerical simulations that the savings channel dominates and adding LOD causes the optimal tax on capital to increase.

In a standard life cycle model, I find that the optimal tax on capital falls

in the large range, depending on the model's assumptions with regard to how human capital is accumulated and the shape of the lifetime Frisch elasticity profile implied by the utility specification. In order for economists to reach more precise conclusions from the model, two empirical questions must be answered: What is the process by which agents acquire age-specific human capital once they start working? And what is the shape of the labor supply elasticity profile? I leave both of these questions for future research.

1.10 Appendix

1.10.1 Analytical Derivations

Exogenous

The Lagrangian for this specification is

$$\begin{aligned} \mathcal{L} = & \frac{c_{1,t}^{1-\sigma_1}}{1-\sigma_1} - \chi \frac{h_{1,t}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} + \beta \frac{c_{2,t+1}^{1-\sigma_1}}{1-\sigma_1} - \chi \frac{h_{2,t+1}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} \\ & - \rho_t (c_{1,t} + c_{2,t} + K_{t+1} - K_t + G_t - rK_t - w(h_{1,t} + h_{2,t}\epsilon_2)) \\ & - \rho_{t+1} \theta (c_{1,t+1} + c_{2,t+1} + K_{t+2} - K_{t+1} + G_{t+1} - rK_{t+1} - w(h_{1,t+1} + h_{2,t+1}\epsilon_2)) \\ & + \lambda_t (c_{1,t}^{1-\sigma_1} + \beta c_{2,t+1}^{1-\sigma_1} - \chi h_{1,t}^{1+\frac{1}{\sigma_2}} - \beta \chi h_{2,t+1}^{1+\frac{1}{\sigma_2}}) \end{aligned} \quad (1.47)$$

where ρ is the Lagrange multiplier on the resource constraint and λ is the Lagrange multiplier on the implementability constraint. The first order conditions with respect to labor, capital and consumption are

$$\rho_t = \chi h_{1,t}^{\frac{1}{\sigma_2}} \left(1 + \lambda_t \left(1 + \frac{1}{\sigma_2}\right)\right) \quad (1.48)$$

$$\rho_{t+1} \theta \epsilon_2 = \beta \chi h_{2,t+1}^{\frac{1}{\sigma_2}} \left(1 + \lambda_t \left(1 + \frac{1}{\sigma_2}\right)\right) \quad (1.49)$$

$$\rho_t = \theta (1+r) \rho_{t+1} \quad (1.50)$$

$$\rho_t = c_{1,t}^{-\sigma_1} + \lambda_t (1-\sigma_1) c_{1,t}^{-\sigma_1} \quad (1.51)$$

and

$$\theta \rho_{t+1} = \beta c_{2,t+1}^{-\sigma_1} + \beta \lambda_t (1 - \sigma_1) c_{2,t+1}^{-\sigma_1}. \quad (1.52)$$

Combining the first order equations for the governments problem with respect to capital and consumption yields

$$\left(\frac{c_{2,t+1}}{c_{1,t}} \right)^{\sigma_1} = \frac{\beta \rho_t}{\rho_{t+1} \theta} \quad (1.53)$$

where ρ is the Lagrange multiplier on the resource constraint and λ is the Lagrange multiplier on the implementability constraint. Taking the ratio of the agent's first order conditions, equations 1.5 and 1.6 under the benchmark utility specification gives

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \frac{1}{\epsilon_2} \left(\frac{c_{1,t}}{c_{2,t+1}} \right)^{-\sigma_1} \left(\frac{h_{2,t+1}}{h_{1,t}} \right)^{\frac{1}{\sigma_2}}. \quad (1.54)$$

Combining equation 1.53 and 1.54 yields

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \frac{1}{\epsilon_2} \left(\frac{\beta \rho_t}{\rho_{t+1} \theta} \right) \left(\frac{h_{2,t+1}}{h_{1,t}} \right)^{\frac{1}{\sigma_2}}. \quad (1.55)$$

The ratio of first order equations for the government with respect to young and old hours is

$$\frac{\rho_t \beta}{\epsilon_2 \rho_{t+1} \theta} \left(\frac{h_{2,t+1}}{h_{1,t}} \right)^{\frac{1}{\sigma_2}} = \frac{1 + \lambda_t (1 + \frac{1}{\sigma_2})}{1 + \lambda_t (1 + \frac{1}{\sigma_2})}. \quad (1.56)$$

Combining equation 1.56 and 1.55 generates the following expression for labor taxes

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \frac{1 + \lambda_t (1 + \frac{1}{\sigma_2})}{1 + \lambda_t (1 + \frac{1}{\sigma_2})} = 1. \quad (1.57)$$

LBD

The Lagrangian for this LBD specification is

$$\begin{aligned}
\mathcal{L} = & \frac{c_{1,t}^{1-\sigma_1}}{1-\sigma_1} - \chi \frac{h_{1,t}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} + \beta \frac{c_{2,t+1}^{1-\sigma_1}}{1-\sigma_1} - \chi \frac{h_{2,t+1}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} \\
& - \rho_t (c_{1,t} + c_{2,t} + K_{t+1} - K_t + G_t - rK_t - w(h_{1,t} + h_{2,t}s_2)) \\
& - \rho_{t+1} \theta (c_{1,t+1} + c_{2,t+1} + K_{t+2} - K_{t+1} + G_{t+1} - rK_{t+1} - w(h_{1,t+1} + h_{2,t+1}s_2)) \\
& + \lambda_t (c_{1,t}^{1-\sigma_1} + \beta c_{2,t+1}^{1-\sigma_1} - \chi h_{1,t}^{1+\frac{1}{\sigma_2}} + \frac{\chi \beta h_{2,t+1}^{1+\frac{1}{\sigma_2}} h_{1,t} s_{h1}(t+1)}{s_2} - \beta \chi h_{2,t+1}^{1+\frac{1}{\sigma_2}})
\end{aligned} \tag{1.58}$$

where ρ is the Lagrange multiplier on the resource constraint and λ is the Lagrange multiplier on the implementability constraint. The first order conditions with respect to labor, capital and consumption are

$$\begin{aligned}
\rho_t (1 + h_{2,t+1} s_{h1}(t+1)) = & \chi h_{1,t}^{\frac{1}{\sigma_2}} (1 + \lambda_t (1 + \frac{1}{\sigma_2})) - \theta \rho_{t+1} h_{2,t+1} s_{h1}(t+1) \\
& + \lambda_t \chi h_{2,t+1}^{1+\frac{1}{\sigma_2}} \beta h_{1,t} \left[\frac{s_{h1}(t+1)^2}{s_2^2} - \frac{s_{h2,h2}(t+1)}{s_2} \right]
\end{aligned} \tag{1.59}$$

$$\rho_{t+1} \theta s_2 = \beta \chi h_{2,t+1}^{\frac{1}{\sigma_2}} \left[1 + \lambda_t (1 + \frac{1}{\sigma_2}) + (1 + \frac{1}{\sigma_2}) \frac{h_{1,t} s_{h1}(t+1) \lambda_t}{s_2} \right] \tag{1.60}$$

$$\rho_t = \theta (1 + r) \rho_{t+1} \tag{1.61}$$

$$\rho_t = c_{1,t}^{-\sigma_1} + \lambda_t (1 - \sigma_1) c_{1,t}^{-\sigma_1} \tag{1.62}$$

and

$$\theta \rho_{t+1} = \beta c_{2,t+1}^{-\sigma_1} + \beta \lambda_t (1 - \sigma_1) c_{2,t+1}^{-\sigma_1}. \tag{1.63}$$

The first order conditions with respect to capital and consumption are the same in the exogenous (1.50, 1.51, and 1.52) and LBD models (1.61, 1.62, and 1.63). Therefore equation 1.14 still holds for this model and therefore the optimal tax on capital is still zero when the government can condition labor income taxes on age.

Combining the first order equations for the governments problem with re-

spect to capital and consumption yields

$$\left(\frac{c_{2,t+1}}{c_{1,t}}\right)^{\sigma_1} = \frac{\beta\rho_t}{\rho_{t+1}\theta} \quad (1.64)$$

Taking the ratio of the agent's first order conditions, equations 1.20 and 1.21 and combining with equation 1.64 yields

$$\frac{1 - \tau_{h,1}}{1 - \tau_{h,2}} = \left(\frac{h_{1,t}}{h_{2,t+1}}\right)^{\frac{1}{\sigma_2}} \left(\frac{\rho_{t+1}\theta s_2}{\beta\rho_t}\right) - \frac{h_{2,t+1}s_{h_2}(t+1)}{1+r(1-\tau_k)}. \quad (1.65)$$

Combining equations 1.65, 1.59 and 1.60 the ratio of the optimal taxes on labor is,

$$\begin{aligned} \frac{1-\tau_{h,1}}{1-\tau_{h,2}} = & \frac{\left[1+h_{2,t+1}s_{h_2}(t+1)\right] \left[1+\lambda_t\left(1+\frac{h_{1,t}s_{h_2}(t+1)}{s_2}\right)\left(1+\frac{1}{\sigma_2}\right)\right]}{1+\lambda_t\left(1+\frac{1}{\sigma_2}\right)-\beta h_{2,t+1} h_{1,t}^{\frac{1+\frac{1}{\sigma_2}}{\sigma_2}} \left[s_{h_2}(t+1)\left(1+\lambda_t\left(1+\frac{h_{1,t}s_{h_2}(t+1)}{s_2}\right)\right)\left(1+\frac{1}{\sigma_2}\right)-h_1\left(\left(\frac{s_{h_2}(t+1)}{s_2}\right)^2-\frac{s_{h_2}h_2(t+1)}{s_2}\right)\right]} \\ & - \frac{h_{2,t+1}s_{h_2}(t+1)}{1+r(1-\tau_k)}. \end{aligned} \quad (1.66)$$

LOD

The Lagrangian for the LOD model is

$$\begin{aligned} \mathcal{L} = & \frac{c_{1,t}^{1-\sigma_1}}{1-\sigma_1} - \chi \frac{(h_{1,t} + n_{1,t})^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} + \beta \frac{c_{2,t+1}^{1-\sigma_1}}{1-\sigma_1} - \chi \frac{h_{2,t+1}^{1+\frac{1}{\sigma_2}}}{1+\frac{1}{\sigma_2}} \\ & - \rho_t(c_{1,t} + c_{2,t} + K_{t+1} - K_t + G_t - rK_t - w(h_{1,t} + h_{2,t}s_2)) \\ & - \rho_{t+1}\theta(c_{1,t+1} + c_{2,t+1} + K_{t+2} - K_{t+1} + G_{t+1} - rK_{t+1} - w(h_{1,t+1} + h_{2,t+1}s_2)) \\ & + \lambda_t(c_{1,t}^{1-\sigma_1} + \beta c_{2,t+1}^{1-\sigma_1} - \chi h_{1,t}^{1+\frac{1}{\sigma_2}} - \beta \chi h_{2,t+1}^{1+\frac{1}{\sigma_2}}) \\ & + \eta_t(\chi h_{2,t+1}^{1+\frac{1}{\sigma_2}} s_{n_1}(t+1) - \chi(h_{1,t} + n_{1,t})^{\frac{1}{\sigma_2}}) \end{aligned} \quad (1.67)$$

where ρ is the Lagrange multiplier on the resource constraint, λ is the Lagrange multiplier on the implementability constraint and η is the Lagrange multiplier on the constraint equating the first order conditions with respect to training and work. The first order conditions with respect to labor, capital, consumption and training

are,

$$\rho_t = \chi(h_{1,t} + n_{1,t})^{\frac{1}{\sigma_2}} \left[1 + \lambda_t \left(1 + \frac{h_{1,t}}{\sigma_2(h_{1,t} + n_{1,t})} \right) + \frac{\eta_t s_2}{\sigma_2(h_{1,t} + n_{1,t})} \right] \quad (1.68)$$

$$\rho_{t+1} \theta s_2 = \beta \chi h_{2,t+1}^{\frac{1}{\sigma_2}} \left[1 + \lambda_2 \left(1 + \frac{1}{\sigma_2} \right) - \eta_t \left(1 + \frac{1}{\sigma_2} \right) s_{n1}(t+1) \right] \quad (1.69)$$

$$\rho_t = \theta(1+r)\rho_{t+1} \quad (1.70)$$

$$\rho_t = c_{1,t}^{-\sigma_1} + \lambda_t(1-\sigma_1)c_{1,t}^{-\sigma_1} \quad (1.71)$$

$$\theta\rho_{t+1} = \beta c_{2,t+1}^{-\sigma_1} + \beta\lambda_t(1-\sigma_1)c_{2,t+1}^{-\sigma_1} \quad (1.72)$$

and

$$\begin{aligned} \theta\rho_{t+1}h_{2,t+1}s_{n2}(t+1) = & \quad (1.73) \\ \frac{\chi(h_{1,t} + n_{1,t})^{\frac{1}{\sigma_2}} \left(\lambda_t h_{1,t} + \eta_t s_2 + \sigma_2(h_{1,t} + n_{1,t})(1 + \eta_t s_{n2}(t+1)) \right)}{\sigma_2(h_{1,t} + n_{1,t})} \\ - \frac{\beta\chi\eta_t\sigma_2 h_{2,t+1}^{1+\frac{1}{\sigma_2}}(h_{1,t} + n_{1,t})s_{n2,n2}(t+1)}{\sigma_2(h_{1,t} + n_{1,t})} \end{aligned}$$

The first order conditions with respect to capital and consumption are the same in the exogenous (1.50, 1.51, and 1.52) and LOD models (1.70, 1.71, and 1.72). Therefore equation 1.14 still holds for this model and therefore the optimal tax on capital is still zero when the government can condition labor income taxes on age.

Combining the first order equations for the governments problem with respect to capital and consumption yields

$$\left(\frac{c_{2,t+1}}{c_{1,t}} \right)^{\sigma_1} = \frac{\beta\rho_t}{\rho_{t+1}\theta} \quad (1.74)$$

Taking the ratio of the agent's first order conditions, equations 1.31 and 1.32 and combining with equation 1.74 yields

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \left(\frac{h_{2,t+1}}{h_{1,t} + n_{1,t}} \right)^{\frac{1}{\sigma_2}} \left(\frac{\beta\rho_t}{\rho_{t+1}\theta s_2} \right). \quad (1.75)$$

Taking the ratio of equations 1.68 and 1.69 yields,

$$\left(\frac{h_{2,t+1}}{h_{1,t} + n_{1,t}}\right)^{\frac{1}{\sigma_2}} \left(\frac{\beta \rho_t}{\rho_{t+1} \theta s_2}\right) = \frac{1 + \lambda_t \left(1 + \frac{h_{1,t}}{\sigma_2(h_{1,t} + n_{1,t})}\right) + \frac{\eta_t s_2}{\sigma_2(h_{1,t} + n_{1,t})}}{1 + \lambda_t \left(1 + \frac{1}{\sigma_2}\right) - \eta_t s_{n1}(t+1) \left(1 + \frac{1}{\sigma_2}\right)}. \quad (1.76)$$

Combining equations 1.75 and 1.76 generates the following expression for the ratio of the optimal labor taxes,

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \frac{1 + \lambda_t \left(1 + \frac{h_{1,t}}{\sigma_2(h_{1,t} + n_{1,t})}\right) + \frac{\eta_t s_2}{\sigma_2(h_{1,t} + n_{1,t})}}{1 + \lambda_t \left(1 + \frac{1}{\sigma_2}\right) - \eta_t s_{n1}(t+1) \left(1 + \frac{1}{\sigma_2}\right)}. \quad (1.77)$$

1.10.2 Competitive Equilibrium

LBD Model

Given a social security replacement rate b , a sequence of skill accumulations parameters $\{\Omega_j\}_{j=1}^{j_r-1}$, government expenditures G , and a sequence of population shares $\{\mu_j\}_{j=1}^J$, a stationary competitive equilibrium in the LBD model is a sequence of agent allocations, $\{c_j, a_{j+1}, h_j\}_{j=1}^J$, a production plan for the firm (N, K) , a government labor tax function $T^l : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, a government capital tax function $T^k : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, a social security tax rate τ_{ss} , a age-specific human capital accumulation function $S : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$, a utility function $U : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$, social security benefits SS , prices (w, r) , and transfers Tr such that:

1. Given prices, policies, transfers, and benefits the agent maximizes the following

$$\sum_{j=1}^J \text{Max}_{c_j, h_j, a_{j+1}} \beta^{j-1} \left[\prod_{q=0}^{j-1} \Psi_q \right] u(c_j, h_j)$$

subject to

$$\begin{aligned} c_j + a_{j+1} &= w s_j h_j - \tau_{ss} w s_j h_j + (1+r)(a_j + Tr) \\ &\quad - T^l[w s_j h_j (1 - .5\tau_{ss})] - T^k[r(a_j + Tr)], \end{aligned}$$

$$s_{j+1} = S_{\text{LBD}}(\Omega_j, s_j, h_j),$$

for $j < j_r$, and

$$c_j + a_{j+1} = SS + (1+r)(a_j + Tr) - T^k[r(a_j + Tr)],$$

for $j \geq j_r$. Additionally,

$$c \geq 0, 0 \leq h \leq 1,$$

$$a_j \geq 0, a_1 = 0.$$

2. Prices w and r satisfy

$$r = \alpha \left(\frac{N}{\bar{K}} \right)^{1-\alpha} - \delta$$

$$w = (1 - \alpha) \left(\frac{K}{N} \right)^\alpha$$

3. The social security policies satisfy

$$SS = b \frac{wN}{\sum_{j=1}^{j_r-1} \mu_j}$$

$$\tau_{ss} = \frac{SS \sum_{j=j_r}^J \mu_j}{w \sum_{j=1}^{j_r-1} \mu_j}$$

4. Transfers are given by

$$Tr = \sum_{j=1}^J \mu_j (1 - \Psi_j) a_{j+1}$$

5. Government budget balance:

$$G = \sum_{j=1}^J \mu_j T^k [r(a_j + Tr)] + \sum_{j=1}^{j_r-1} \mu_j T^l [w s_j h_j (1 - .5\tau_{ss})]$$

6. Market clearing:

$$K = \sum_{j=1}^J \mu_j a_j$$

$$N = \sum_{j=1}^J \mu_j s_j h_j$$

$$\sum_{j=1}^J \mu_j c_j + \sum_{j=1}^J \mu_j a_{j+1} + G = K^\alpha N^{1-\alpha} + (1 - \delta)K$$

LOD Model

Given a social security replacement rate b , a sequence of skill accumulations parameters $\{\Omega_j\}_{j=1}^{j_r-1}$, government expenditures G , and a sequence of population shares $\{\mu_j\}_{j=1}^J$, a stationary competitive equilibrium in the LBD model is a sequence of agent allocations, $\{c_j, a_{j+1}, h_j\}_{j=1}^J$, a production plan for the firm (N, K) ,

a government labor tax function $T^l : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, a government capital tax function $T^k : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, a social security tax rate τ_{ss} , a age-specific human capital accumulation function $S : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$, a utility function $U : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$, social security benefits SS , prices (w, r) , and transfers Tr such that:

1. Given prices, policies, transfers, and benefits the agent maximizes the following

$$\sum_{j=1}^J \text{Max}_{c_j, h_j, n_j, a_{j+1}} \beta^{j-1} \left[\prod_{q=0}^{j-1} \Psi_q \right] u(c_j, h_j, n_j)$$

subject to

$$\begin{aligned} c_j + a_{j+1} &= ws_j h_j - \tau_{ss} ws_j h_j + (1+r)(a_j + Tr) \\ &\quad - T^l[ws_j h_j(1 - .5\tau_{ss})] - T^k[r(a_j + Tr)], \end{aligned}$$

$$s_{j+1} = S_{\text{LOD}}(\Omega_j, n_j, h_j),$$

for $j < j_r$, and

$$c_j + a_{j+1} = SS + (1+r)(a_j + Tr) - T^k[r(a_j + Tr)],$$

for $j \geq j_r$. Additionally,

$$c \geq 0, 0 \leq h \leq 1,$$

$$a_j \geq 0, a_1 = 0.$$

2. Prices w and r satisfy

$$r = \alpha \left(\frac{N}{K} \right)^{1-\alpha} - \delta$$

$$w = (1 - \alpha) \left(\frac{K}{N} \right)^\alpha$$

3. The social security policies satisfy

$$SS = b \frac{wN}{\sum_{j=1}^{j_r-1} \mu_j}$$

$$\tau_{ss} = \frac{ss \sum_{j=j_r}^J \mu_j}{w \sum_{j=1}^{j_r-1} \mu_j}$$

4. Transfers are given by

$$Tr = \sum_{j=1}^J \mu_j (1 - \Psi_j) a_{j+1}$$

5. Government budget balance:

$$G = \sum_{j=1}^J \mu_j T^k [r(a_j + Tr)] + \sum_{j=1}^{j_r-1} \mu_j T^l [ws_j h_j (1 - .5\tau_{ss})]$$

6. Market clearing:

$$K = \sum_{j=1}^J \mu_j a_j$$

$$N = \sum_{j=1}^J \mu_j s_j h_j$$

$$\sum_{j=1}^J \mu_j c_j + \sum_{j=1}^J \mu_j a_{j+1} + G = K^\alpha N^{1-\alpha} + (1 - \delta)K$$

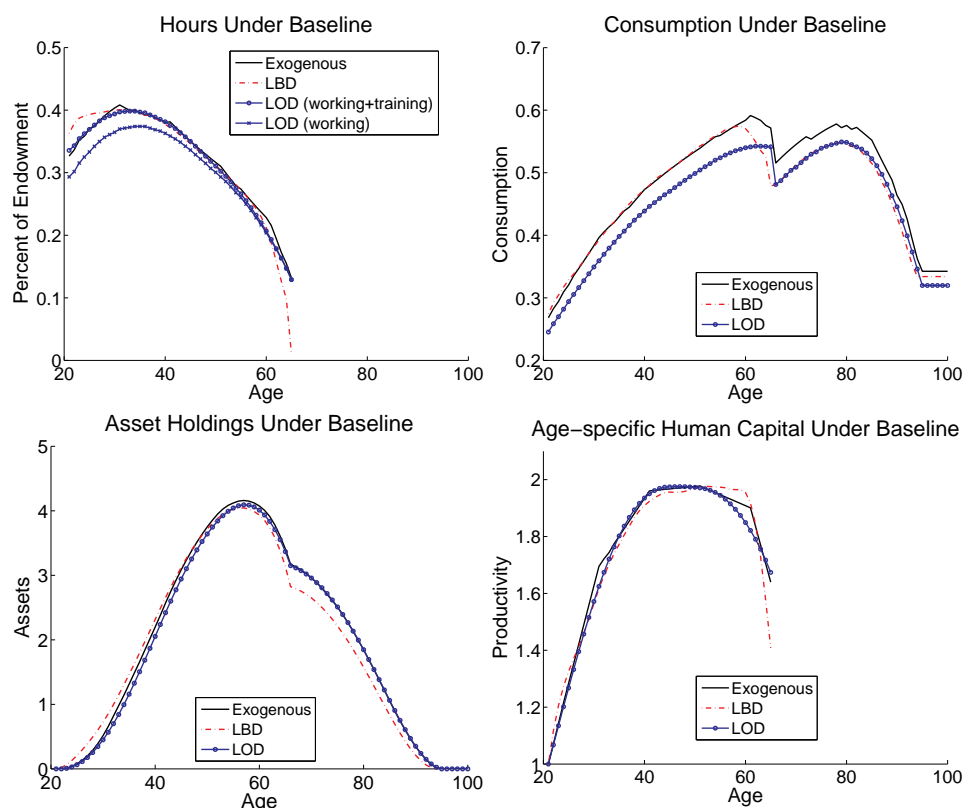
1.10.3 Non-separable Utility

Below are the tables and graphs of the economic aggregate variables and the life cycle profiles with the non-separable utility function. There is a discontinuity in consumption and assets at retirement, since labor supply changes discontinuously and the utility function is non-separable in consumption and hours worked.

Table 1.10: Aggregate Economic Variables (Non-separable Utility)

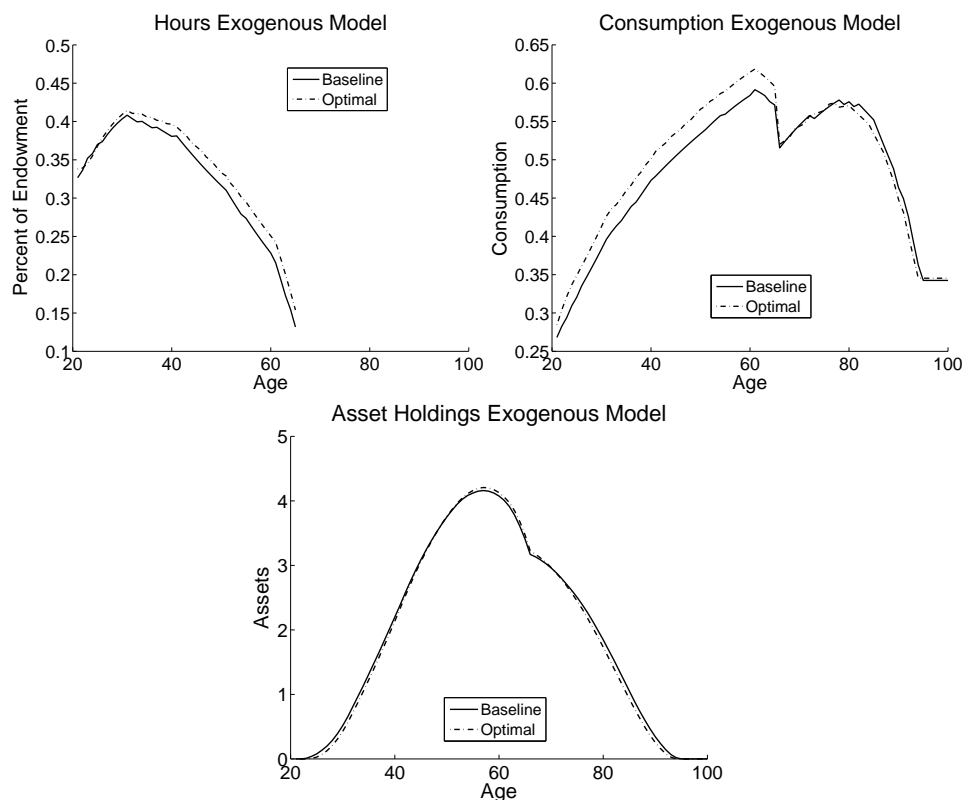
Aggregate	Exogenous			LBD			LOD		
	Baseline	Optimal	% Change from Baseline to Optimal	Baseline	Optimal	% Change from Baseline to Optimal	Baseline	Optimal	% Change from Baseline to Optimal
Y	0.80	0.82	2.8%	0.80	0.81	2.1%	0.76	0.77	1.8%
K	2.11	2.10	-0.8%	2.12	1.94	-8.5%	2.04	1.98	-3.0%
N	0.46	0.49	4.9%	0.46	0.50	8.6%	0.43	0.45	4.7%
Avg Hours	0.34	0.35	3.8%	0.34	0.35	4.8%	0.33	0.35	3.7%
w	1.10	1.08	-2.0%	1.11	1.04	-6.0%	1.12	1.09	-2.7%
r	0.05	0.06	9.3%	0.05	0.07	30.1%	0.05	0.06	13.2%
tr	0.02	0.02	-3.1%	0.02	0.02	-3.8%	0.02	0.02	-2.5%
Value	-106.85	-103.96	2.7%	-96.51	-94.44	2.1%	-121.60	-119.34	1.9%
CEV			2.6%			2.7%			1.9%
Average Tax Rate	Baseline	Optimal		Baseline	Optimal		Baseline	Optimal	
Capital	15.4%	35.3%		16.1%	46.3%		15.5%	36.5%	
Labor	23.4%	18.3%		23.4%	15.0%		23.5%	18.7%	
Ratio	0.66	1.93		0.69	3.09		0.66	1.95	
	0.64	0.48		0.72	0.92		0.64	0.59	
Marginal Tax Rate	Baseline	Optimal		Baseline	Optimal		Baseline	Optimal	
Capital	19.4%	35.3%		20.2%	46.3%		19.5%	36.5%	
Labor	25.4%	18.3%		25.3%	15.0%		25.4%	18.7%	
Ratio	0.77	1.93		0.80	3.09		0.77	1.95	

Notes: The average hours refers to the average percent of time endowment worked in the productive labor sector. Both the marginal and average tax rates vary with income under the baseline fitted US tax policy. The numbers reported are the population weighted averages.



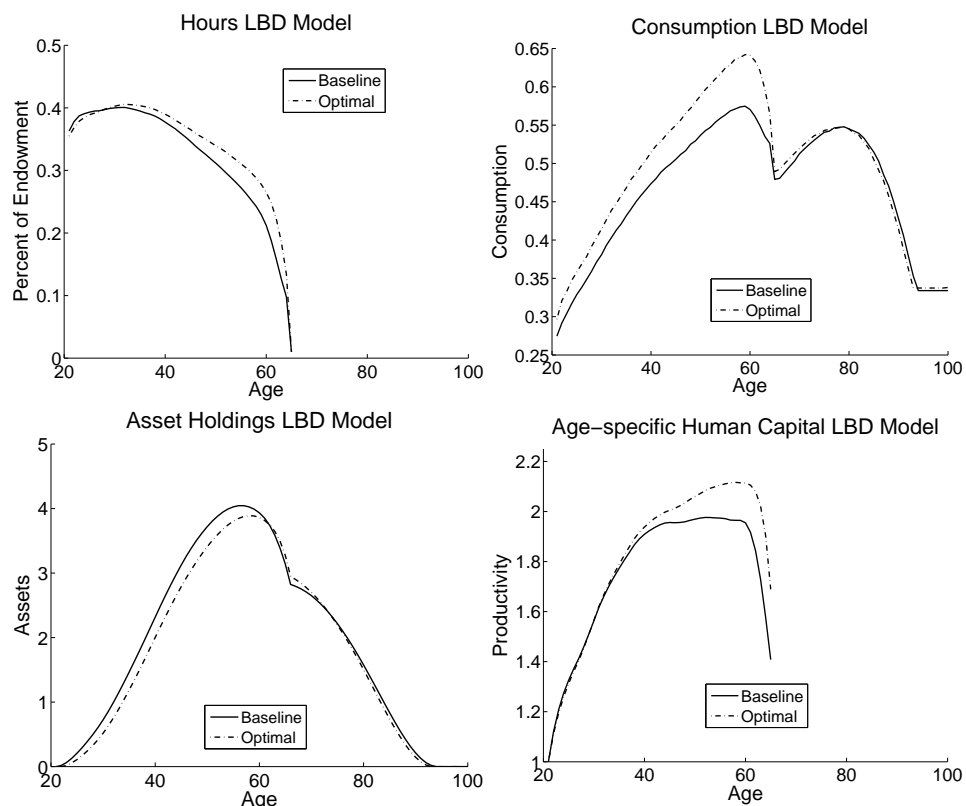
Notes: The plots are life cycle profiles of the three calibrated models under the baseline fitted US tax policy. The upper left hand panel is a plot of the labor supply, the upper right hand panel is a plot of consumption, the bottom left panel is a plot of assets and the bottom right is a plot of skills. The solid black lines are in the exogenous model, the dashed red lines are in the LBD model and the blue marked lines are in the LOD model. There are two labor lines for the LOD model, one for just hours worked and one for hours worked plus hours spent training.

Figure 1.7: Life Cycle Profiles under Baseline Fitted US Tax Policy (Non-separable Utility)



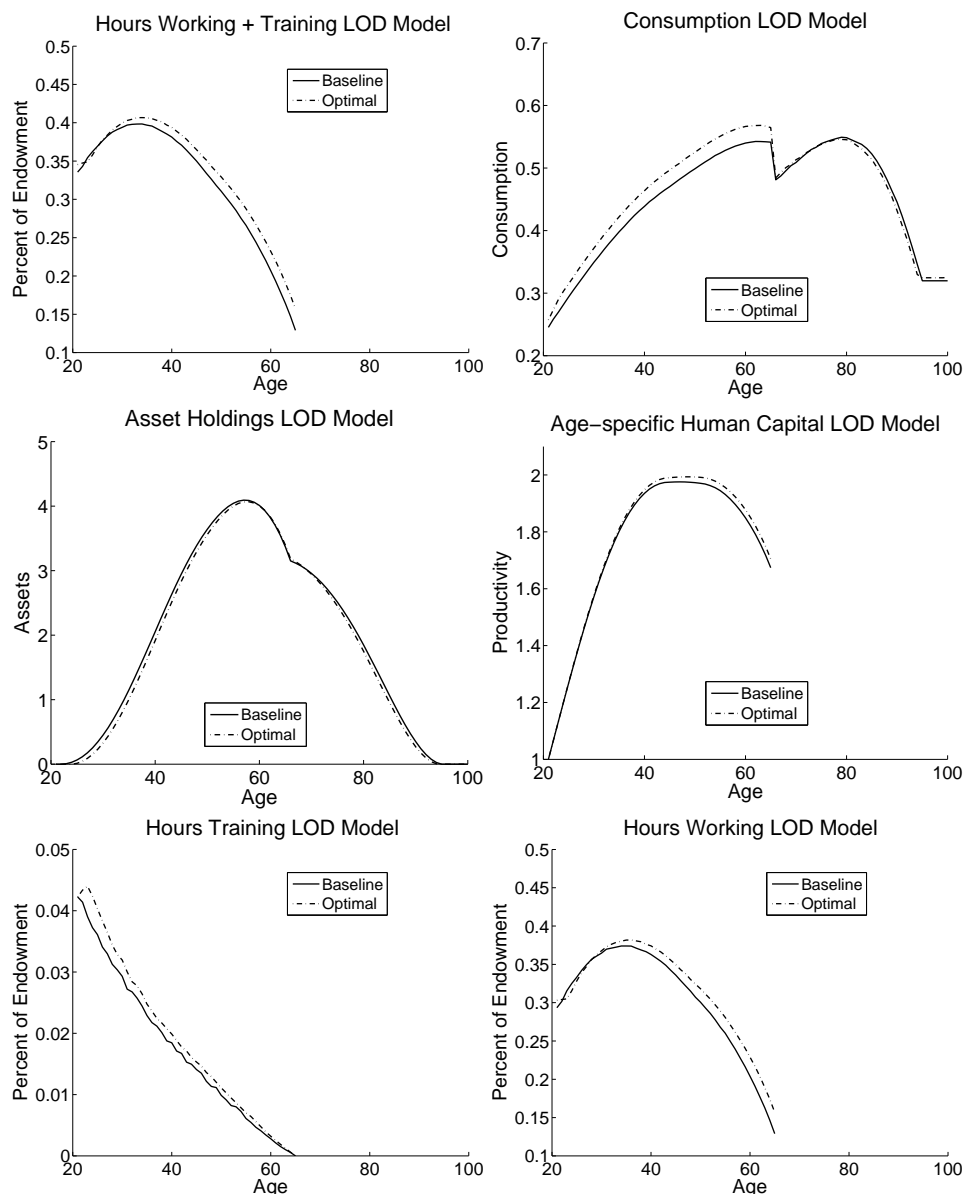
Notes: The upper left hand panel is a plot of the labor supply, the upper right hand panel is a plot of consumption, and the bottom left panel is a plot of assets. Since the skills are the same in the exogenous models under the baseline fitted US tax policy and optimal tax policy they are not plotted. Solid lines are under the baseline fitted US tax policy while dashed lines are under the optimal tax system.

Figure 1.8: Life Cycle Profiles in the Exogenous Model (Non-separable Utility)



Notes: The upper left hand panel is a plot of the labor supply, the upper right hand panel is a plot of consumption, the bottom left panel is a plot of assets, and the bottom right panel is a plot of age-specific human capital. Solid lines are under the baseline fitted US tax policy while dashed lines are under the optimal tax system.

Figure 1.9: Life Cycle Profiles in the LBD Model (Non-separable Utility)



Notes: The upper left hand panel is a plot of the sum of labor and training, the upper right hand panel is a plot of consumption, the middle left panel is a plot of assets, the middle right panel is a plot of age-specific human capital, the bottom right panel is a plot of time spent training, and the bottom right is a plot of time spent working. Solid lines are under the baseline fitted US tax policy while dashed lines are under the optimal tax system.

Figure 1.10: Life Cycle Profiles in the LOD Model (Non-separable Utility)

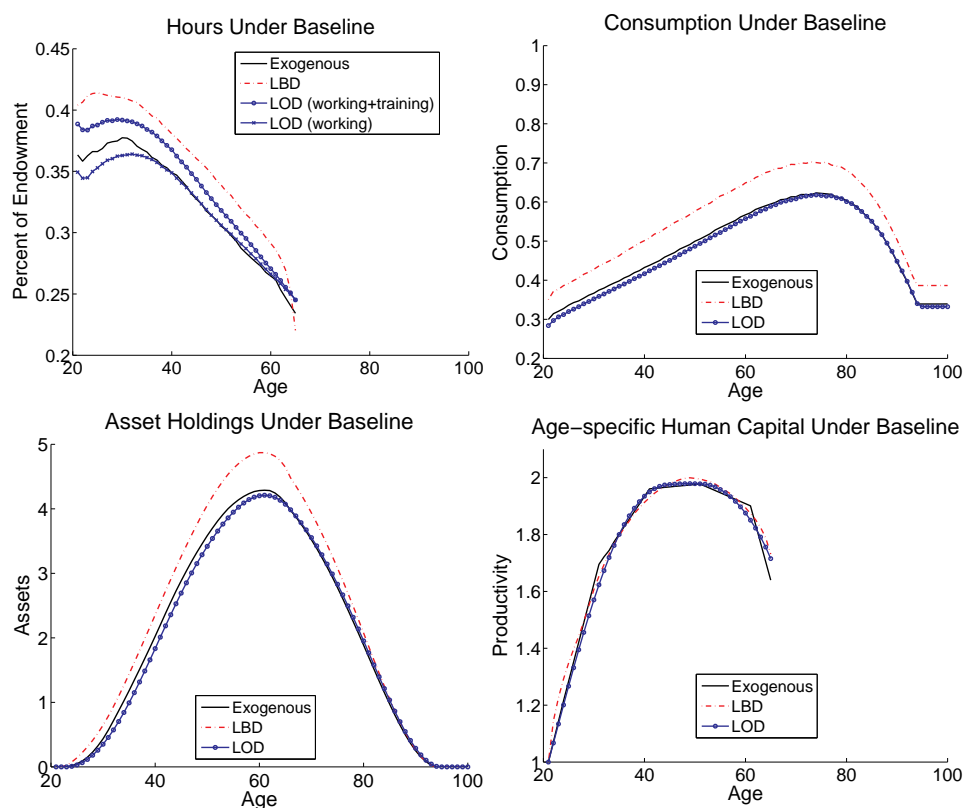
1.10.4 Sequential Parametrization

This section provides the aggregate economic variables and the life cycle profiles for the model that is sequentially parameterized. There are no plots for the effect of the optimal tax policy in the exogenous model because they are the same as in the calibrated model (see Figure 1.6 and 1.4)

Table 1.11: Aggregate Economic Variables (Sequential Parametrization)

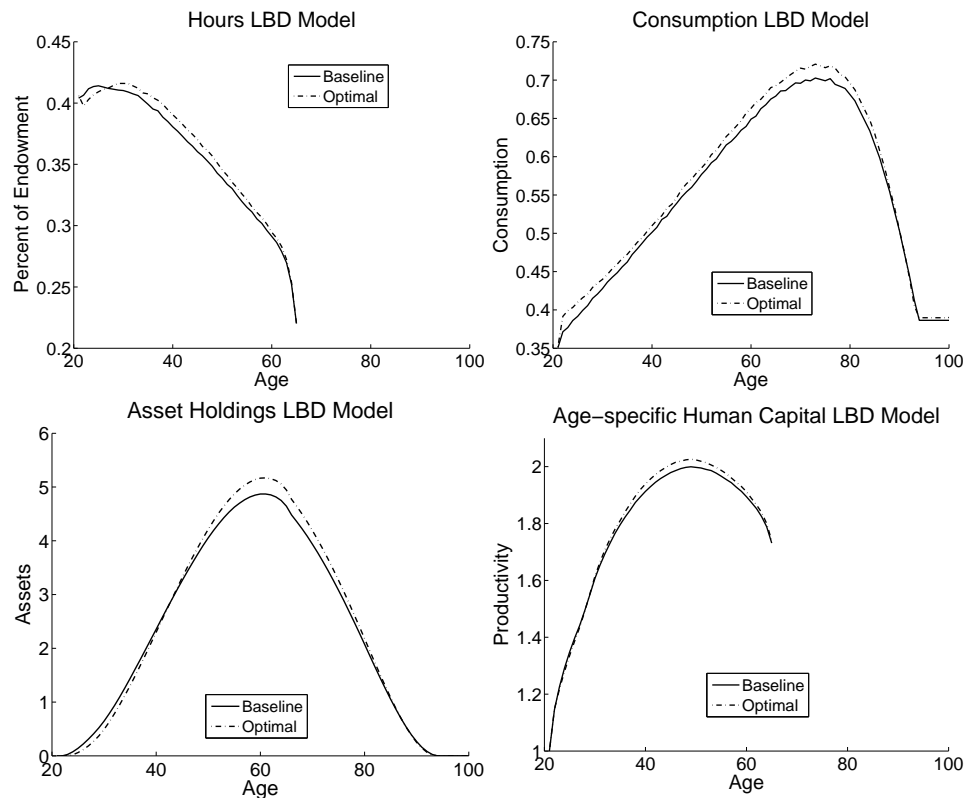
Aggregate	Exogenous			LBD			LOD		
	Baseline	Optimal	<i>% Change from Baseline to Optimal</i>	Baseline	Optimal	<i>% Change from Baseline to Optimal</i>	Baseline	Optimal	<i>% Change from Baseline to Optimal</i>
Y	0.79	0.82	3.5%	0.91	0.92	1.3%	0.78	0.80	1.9%
K	2.14	2.28	6.9%	2.51	2.52	0.6%	2.06	2.14	3.9%
N	0.46	0.46	1.6%	0.51	0.52	1.6%	0.45	0.46	0.8%
Avg Hours	0.33	0.34	0.7%	0.37	0.37	1.3%	0.35	0.35	0.6%
w	1.12	1.14	1.8%	1.13	1.13	-0.4%	1.10	1.12	1.1%
r	0.05	0.05	-8.4%	0.05	0.05	1.8%	0.05	0.05	-4.9%
tr	0.03	0.03	7.0%	0.03	0.03	3.6%	0.03	0.03	5.8%
Value	-136.58	-135.85	0.5%	-130.43	-129.62	0.6%	-143.03	-142.10	0.7%
CEV			0.7%			2.6%			2.1%
Average Tax Rate	Baseline	Optimal		Baseline	Optimal		Baseline	Optimal	
Capital	14.9%	11.8%		11.5%	22.3%		15.9%	17.8%	
Labor	23.4%	24.7%		21.3%	19.5%		23.9%	24.1%	
Ratio	0.64	0.48		0.54	1.14		0.67	0.74	
Marginal Tax Rate	Baseline	Optimal		Baseline	Optimal		Baseline	Optimal	
Capital	18.9%	11.8%		15.9%	22.3%		19.6%	17.8%	
Labor	25.4%	24.7%		24.6%	19.5%		25.5%	24.1%	
Ratio	0.74	0.48		0.65	1.14		0.77	0.74	

Notes: The average hours refers to the average percent of time endowment worked in the productive labor sector. Both the marginal and average tax rates vary with income under the baseline fitted US tax policy. The numbers reported are the population weighted averages.



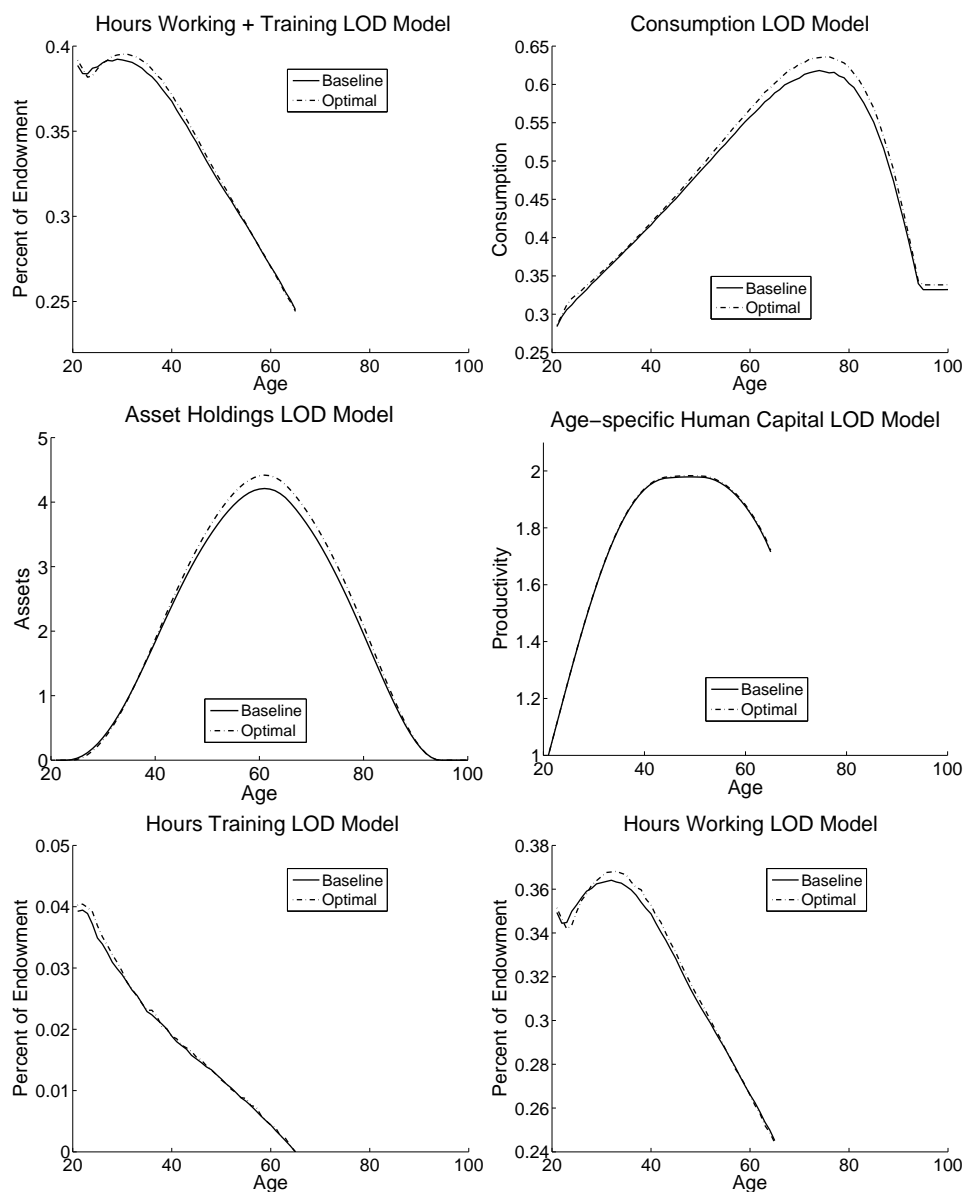
Notes: The plots are life cycle profiles of the three calibrated models under the baseline fitted US tax policy. The upper left hand panel is a plot of the labor supply, the upper right hand panel is a plot of consumption, the bottom left panel is a plot of assets and the bottom right is a plot of skills. The solid black lines are in the exogenous model, the dashed red lines are in the LBD model and the blue marked lines are in the LOD model. There are two labor lines for the LOD model, one for just hours worked and one for hours worked plus hours spent training.

Figure 1.11: Life Cycle Profiles under Baseline Fitted US Tax Policy (Seq. Parametrization)



Notes: The upper left hand panel is a plot of the labor supply, the upper right hand panel is a plot of consumption, the bottom left panel is a plot of assets, and the bottom right panel is a plot of age-specific human capital. Solid lines are under the baseline fitted US tax policy while dashed lines are under the optimal tax system.

Figure 1.12: Life Cycle Profiles in the LBD Model (Seq. Parametrization)



Notes: The upper left hand panel is a plot of the sum of labor and training, the upper right hand panel is a plot of consumption, the middle left panel is a plot of assets, the middle right panel is a plot of age-specific human capital, the bottom right panel is a plot of time spent training, and the bottom right is a plot of time spent working. Solid lines are under the baseline fitted US tax policy while dashed lines are under the optimal tax system.

Figure 1.13: Life Cycle Profiles in the LoD Model (Seq. Parametrization)

Chapter 2

Determining the Motives for a Positive Optimal Tax on Capital

Abstract

Previous literature demonstrates that in a computational life cycle model the optimal tax on capital is positive and large. In a standard overlapping generations model, this paper measures the relative strength of the motives generally understood to produce a large optimal tax on capital. I focus on the impact of changing two common assumptions in a benchmark model that generates a large optimal tax on capital similar to the model in Conesa et al. (2009). First, the utility function is altered such that it implies an agent's Frisch labor supply elasticity is constant, as opposed to increasing, over his lifetime. Second, the government is allowed to tax accidental bequests at a separate rate from ordinary capital income. The main finding of this paper is that these two changes cause the optimal tax on capital to drop by more than seventy percent. Quantifying the impact of these assumptions in the benchmark model is important because the first has limited empirical evidence and the second, although included for tractability, confounds a motive for taxing capital with a motive for taxing accidental bequests.

2.1 Introduction

Total receipts from taxes on individuals' capital income (capital gains and dividends) in 2005 were approximately 140 billion dollars, or 15% of total income tax receipts.¹ Based on the sizable tax receipts from capital income in the US economy and savings disincentives created by a capital tax, considerable research has been devoted to determining whether a non-zero tax on capital income is optimal.² In the seminal works on this topic, Chamley (1986) and Judd (1985) conclude that it is not optimal to tax capital in a model where individuals are infinitely lived and face no idiosyncratic risk. Atkeson et al. (1999) show that the optimal tax on capital is still zero in a two-period overlapping generations model when the government is allowed to condition the labor income tax on age. Other works, such as Aiyagari (1995), Hubbard and Judd (1986), İmrohoroğlu (1998), Erosa and Gervais (2002), Conesa et al. (2009), Garriga (2001), Jones et al. (1997) and Correia (1996), identify theoretical conditions under which it is optimal to tax capital. Five such conditions are when individuals face uninsurable risk, an individual's earnings increase over their lifetime and they face borrowing constraints, the government is not allowed to borrow or save, the government cannot tax all factors of production or sources of income at a separate rate, and certain life cycle features like retirement and social security are included in the model.

When determining the optimal tax on capital, the policymaker must weigh the relevant benefits versus the distortions imposed by the tax. Since a tax on capital discourages saving it is important to analyze the tax in an overlapping generations (OLG) model that includes these life cycle factors that motivate saving. In one such study, Conesa et al. (2009) using a calibrated life cycle model find that the optimal tax policy consists of flat tax rates on capital and labor income of 34% and 14%, respectively.³ Given the computational complexities of these OLG

¹See <http://www.treas.gov/offices/tax-policy/library/capgain3-2008.pdf> and <http://www.irs.gov/taxstats/indtaxstats/article/0,,id=129270,00.html>.

²I define an optimal tax policy as one that maximizes the expected lifetime utility of a newborn in a stationary equilibrium, holding tax revenue constant.

³This is model M4 in Conesa et al. (2009) which excludes idiosyncratic risk. See chapter 1, Smyth (2006), and İmrohoroğlu (1998) for additional OLG studies that demonstrate the optimal

models, it is helpful to determine the economic factors driving these results. This paper quantifies the relative importance of each of these modeling assumptions that motivate a positive tax on capital in a canonical OLG model.

According to the theoretical conditions stated above, there are five common features in OLG models that motivate a non-zero optimal tax on capital.⁴ These features are: (i) a varying lifetime Frisch labor supply elasticity,⁵ (ii) the inability of the government to tax accidental bequests at a different rate from ordinary capital income, (iii) the inability of individuals to borrow, (iv) the inability of the government to hold savings or debt, and (v) the inability of the government to provide a social security program. I start by solving for optimal tax policy in a benchmark model similar to the model in Conesa et al. (2009) that includes all these features and find the optimal tax on capital is large. Next, I solve for the optimal tax policy in a model in which I eliminate two assumptions. First, I no longer use a utility function that implies the Frisch elasticity varies over the life cycle. Second I no longer force the government to tax accidental bequests at the same rate as it taxes ordinary capital income. I test the impact these two features because the first assumption has limited empirical evidence and the second assumption, although included for tractability, confounds a motive for taxing capital with a motive to confiscate accidental bequests.⁶ The main finding of this paper is that these two assumptions are responsible for over seventy percent of the positive optimal tax on capital in a standard OLG model. When these two assumptions are removed from the model the optimal tax on capital is reduced to less than nine percent. Even if I only eliminate the second assumption that the government is forced to tax accidental bequests and ordinary capital income at the same rate, the optimal tax on capital is not zero.

⁴See Conesa et al. (2009), Smyth (2006), Conesa and Krueger (2006), Guvenen et al. (2009), Fuster et al. (2008), Garriga (2001), Erosa and Gervais (2002), and Nakajima (2010) for examples of papers that include similar assumptions when analyzing tax policy in an OLG framework.

⁵The Frisch labor supply elasticity is the labor supply elasticity holding the marginal utility of wealth constant.

⁶One exception is French (2005), in which the authors estimate that the labor supply elasticity is more than three times larger for sixty year old individuals than forty year old individuals. However, the author notes that social security and pension incentives are responsible for this change. Therefore, the change in elasticity results from changes on the extensive margin and not the intensive margin. Since retirement is considered exogenous in my model, I am interested in the Frisch elasticity on the intensive margin.

tax on capital drop by more than fifty percent.

Given these stark results, it becomes necessary to quantify the individual impact of all five modeling features within a common framework. I solve for the optimal tax policy in six other models with one of the five features that motivate a non-zero optimal tax on capital changed in order to determine the impact of each feature.⁷ In addition to the the non-constant Frisch elasticity and the government not being able to tax accidental bequests at a separate rate, I find that the inability of the government to save is also a large motivation for a positive tax on capital. Individual liquidity constraints do not have much impact on the optimal tax policy. Furthermore, if the reduced form social security program is excluded from the model, the optimal tax on capital drops to a large negative number. The tax on capital drops because a negative tax on capital mimics a welfare improving social security program. Therefore, it is important for economists to include a social security program when examining optimal tax policy otherwise the motive for a social security program will be confounded with the motive for a negative tax on capital. Finally, this paper analyzes how the impact of the features change when the model is calibrated to match different targets for the Frisch elasticity. I check the sensitivity of the results since there is a large variance in the empirical estimates of the Frisch elasticity. Generally, I find that these five features have a larger impact on optimal tax policy when the model is calibrated to match a medium or low Frisch elasticity as opposed to a high value.

This exercise is related to Conesa et al. (2009). However there are three important differences. First, I exclude inter-cohort heterogeneity as a motive for a positive tax on capital. I abstract from this type of heterogeneity because Conesa et al. (2009) demonstrate that it does not affect the level of the optimal tax on capital.⁸ Second, I examine how prohibiting the government from taxing accidental bequests at a different rate than ordinary capital income affects the optimal

⁷There are six models because I test both the impact of the government being able to borrow and save.

⁸The authors find that including idiosyncratic uninsurable income shocks and productivity differences affect the progressivity of the optimal labor tax policy but not the optimal level of the tax on labor or capital. Therefore, this paper abstracts from these sources of heterogeneity and focuses on models where agents are homogenous within the cohort.

tax policy. Third, I take an alternative approach to discern the effect of the government's inability to condition labor income taxes on age. In order to determine its impact on optimal tax policy, Conesa et al. (2009) eliminate the desire for the government to condition labor income taxes on age by holding the labor supply exogenously constant. The authors use a utility specification in which the agent's Frisch labor supply elasticity is negatively related to hours worked. Therefore, holding the labor supply constant implies that the Frisch elasticity no longer varies over the life-cycle so the government no longer wants to condition taxes on age. However, using this method abstracts from the general equilibrium effects of endogenously determined labor supply. This paper takes an alternative approach that does not require an exogenously determined labor supply. Instead, I eliminate the government's desire to condition labor income taxes on age by using a utility specification that implies the Frisch labor supply elasticity is constant.⁹ The advantage of this approach is that it isolates the impact of the government wanting to use age-dependent taxes while including general equilibrium effects.

This paper is organized as follows: Section 2 examines a simplified version of the model in order to provide analytical insights into some of the channels that drive the computational results. I introduce the computational model, and present the competitive equilibrium in section 3. Section 4 describes the functional forms and calibration parameters. Section 5 sets up the computational experiment and section 6 reports the results of the computational experiment. Section 7 examines the sensitivity of the results with respect to the target that the Frisch elasticity is calibrated to match. Finally, section 8 summarizes the paper's findings.

2.2 Analytical Model

In order to derive some intuition for when it is optimal to tax capital, I first examine an analytically tractable version of the computational model where agents live with certainty for two periods. In this model I abstract from retirement, population growth, progressive tax policy, and conditional survivability. I begin by

⁹See Garriga (2001) for an analytical derivation of why it is not optimal to condition labor income taxes on age with this type of utility function.

setting up the household problem and illustrating the primal approach for solving for the optimal tax policy. Using the primal approach, I solve for the optimal tax policies in several versions of the simplified model in order to analyze two different assumptions that lead to a non-zero optimal tax on capital. The two assumptions I examine are the inability of the government to borrow or save as well as the inability of the government to condition labor income taxes on age when the Frisch elasticity is not constant. I do not address the other features of the model that motivate a non-zero tax on capital in this section since they cannot be assessed analytically.

After demonstrating how the primal approach is applied to the simple model, I solve for the optimal tax policy in the simple model when the government can condition labor income tax on age. Generally, an efficient labor tax will tax inelastically-supplied labor at a higher rate. I show that in a model with a utility function such that the Frisch labor supply elasticity is a function of time spent working, it is optimal to condition labor income tax on age. Next, I show that in the same model if the government cannot condition the labor income tax on age, it is optimal to tax capital.¹⁰ I show that when the utility function is such that the Frisch labor supply elasticity is constant then the optimal tax policy is not dependent on age and includes a zero tax on capital income.¹¹ Next, I focus on the effect of the government not being allowed to borrow or save. I show that in a model where the government is not allowed to save or borrow and labor income taxes can be conditioned on age, the optimal tax on capital is non-zero. In all models, I solve for the optimal tax policy in the steady state. Therefore, if the government is permitted to save in the model, then its level of savings must be constant. This implies that in the iterations of the model where the government can save or borrow, it cannot run a deficit or a surplus since this would violate definition of the steady state conditions.

In order to demonstrate when the policymaker wants to condition labor

¹⁰Conesa et al. (2009), Garriga (2001), Erosa and Gervais (2002) and Atkeson et al. (1999) demonstrate similar analytic results.

¹¹Garriga (2001) demonstrates a more general result. He shows that if the utility function is both separable and homothetic in consumption and labor then it is not optimal to condition labor income taxes on age.

income taxes on age, I solve the model with two different utility functions:

$$U_{\text{constant Frisch}} = \frac{c^{1-\varsigma_1}}{1-\varsigma_1} - \chi \frac{(h)^{1+\frac{1}{\varsigma_2}}}{1+\frac{1}{\varsigma_2}}$$

$$U_{\text{non-constant Frisch}} = \frac{c^{1-\sigma_1}}{1-\sigma_1} + \chi \frac{(1-h)^{1-\sigma_2}}{1-\sigma_2}.$$

I refer to the first utility function as the “constant Frisch” utility function because the Frisch labor supply elasticity, ς_2 , is not a function of time worked and is constant throughout the agent’s life. I refer to the second utility function as the “non-constant Frisch” utility function because as long as hours are not constant, the Frisch labor supply elasticity, $\frac{1-h}{\sigma_2 h}$, varies over the life cycle.

2.2.1 Households’ Problem

The analytical model is a simplified model where agents live with certainty for two periods and their preferences over consumption and leisure are given by

$$U(c_{1,t}, 1 - h_{1,t}) + \beta U(c_{2,t+1}, 1 - h_{2,t+1}) \quad (2.1)$$

where β is the discount rate, $c_{j,t}$ is the consumption of an age j agent at time t and $h_{j,t}$ is the percent of his time endowment he works.¹² Age-specific human capital is normalized to unity when the agent enters the model. At age 2, age-specific human capital is ϵ_2 . The agent maximizes equation 2.1 with respect to consumption and hours subject to the following constraints

$$c_{1,t} + a_{1,t} = (1 - \tau_{h,1})h_{1,t}w_t \quad (2.2)$$

and

$$c_{2,t+1} = (1 + r_t(1 - \tau_k))a_{1,t} + (1 - \tau_{h,2})\epsilon_2 h_{2,t+1} w_{t+1} \quad (2.3)$$

¹²Time working is measured as a percentage of endowment and not in hours. However for notational convenience, I sometimes refer to $h_{j,t}$ as hours.

where $a_{1,t}$ is the amount saved at age 1, $\tau_{h,j}$ is the tax rate on labor income for an agent of age j , τ_k is the tax rate on capital income, w_t is the efficiency wage for labor services and r_t is the rental rate on capital. I assume that the tax rate on labor income can be conditioned on age in some of the models; however, the tax rate on capital income cannot. I combine equations 2.2 and 2.3 to form a joint intertemporal budget constraint

$$c_{1,t} + \frac{c_{2,t+1}}{1 + r_t(1 - \tau_k)} = w_t(1 - \tau_{h,i})h_{1,t} + \frac{w_{t+1}(1 - \tau_{h,2})\epsilon_2 h_{2,t+1}}{1 + r_t(1 - \tau_k)}. \quad (2.4)$$

The agent's problem is to maximize equation 2.1 subject to 2.4. The agent's first order conditions are

$$\frac{U_{h1}(t)}{U_{c1}(t)} = -w_t(1 - \tau_{h,1}) \quad (2.5)$$

$$\frac{U_{h2}(t+1)}{U_{c2}(t+1)} = -w_{t+1}\epsilon_2(1 - \tau_{h,2}) \quad (2.6)$$

and

$$\frac{U_{c1}(t)}{U_{c2}(t+1)} = \beta(1 + r_t(1 - \tau_k)) \quad (2.7)$$

where $U_{c1}(t) \equiv \frac{\partial U(c_{1,t}, 1 - h_{1,t})}{\partial c_{1,t}}$. Given prices and taxes, these first order conditions together with the intertemporal budget constraint determine the optimal allocation of $(c_{1,t}, h_{1,t}, c_{2,t+1}, h_{2,t+1})$.

2.2.2 Primal Approach

In order to determine the optimal tax policy, I use the primal approach.¹³ I assume that the benevolent government discounts future generations with social discount factor θ . The government maximizes the objective function,

$$[U(c_{2,0}, 1 - h_{2,0})/\theta] + \sum_{t=0}^{\infty} \theta^t [U(c_{1,t}, 1 - h_{1,t}) + \beta U(c_{2,t+1}, 1 - h_{2,t+1})], \quad (2.8)$$

with respect to the implementability constraint and the resource constraint. The implementability constraint is the agent's intertemporal budget constraint with

¹³See Lucas and Stokey (1983) or Erosa and Gervais (2002) for a more in depth discussion of the primal approach.

the prices and taxes replaced by his first order conditions (equations 2.5, 2.6, and 2.7)

$$c_{1,t}U_{c1}(t) + \beta c_{2,t+1}U_{c2}(t+1) + h_{1,t}U_{h1}(t) + \beta h_{2,t+1}U_{h2}(t+1) = 0. \quad (2.9)$$

The resource constraint is

$$c_{1,t} + c_{2,t} + K_{t+1} - K_t + G_t = rK_t + w(h_{1,t} + h_{2,t}\epsilon_2). \quad (2.10)$$

I assume the technology is such that the marginal products of capital and labor are constant.¹⁴ This assumption allows me to focus on the life cycle elements of the model because changes to the tax system do not affect the pre-tax wage or rate of return. Since there is no variation in the factor prices, I suppress the time subscripts on the factor prices.

There are two additional constraints that need to be included for some of the models. When the government cannot condition labor income taxes on age then the tax on labor income must be constant and following constraint is included,

$$\epsilon_2 \frac{U_{h1}(t)}{U_{c1}(t)} = \frac{U_{h2}(t+1)}{U_{c2}(t+1)} \quad (2.11)$$

in order to ensure agents are equating their marginal rate of substitution between labor and consumption across ages. If the government is not allowed to borrow or save then the balanced budget constraint is included,

$$G_t = \tau_{h,1}wh_{1,t} + \tau_{h,2}wh_{2,t}\epsilon_2 + \tau_k r a_{1,t}. \quad (2.12)$$

Using Walras' law this constraint simplifies to $a_{1,t} = K_{t+1}$, which can be rewritten as (using equations 2.2 and 2.5)

$$K_{t+1} = -\frac{U_{h1}(t)}{U_{c1}(t)}h_{1,t} - c_{1,t}. \quad (2.13)$$

¹⁴In the computational model I relax this assumption.

The Lagrangian excluding the additional constraints (equations 2.13 and 2.11) is

$$\begin{aligned}
\mathcal{L} = & U(c_{1,t}, 1 - h_{1,t}) + \beta U(c_{2,t+1}, 1 - h_{2,t+1}) \\
& - \rho_t(c_{1,t} + c_{2,t} + K_{t+1} - K_t + G_t - rK_t - w(h_{1,t} + h_{2,t}\epsilon_2)) \\
& - \rho_{t+1}(c_{1,t+1} + c_{2,t+1} + K_{t+2} - K_{t+1} + G_{t+1} - rK_{t+1} - w(h_{1,t+1} + h_{2,t+1}\epsilon_2)) \\
& + \lambda_t(c_{1,t}U_{c1}(t) + \beta c_{2,t+1}U_{c2}(t+1) + h_{1,t}U_{h1}(t) + \beta h_{2,t+1}U_{h2}(t+1)).
\end{aligned} \tag{2.14}$$

where ρ is the Lagrange multiplier on the resource constraint and λ is the Lagrange multiplier on the implementability constraint.

2.2.3 Mimicking Age-dependent tax on labor

Motivation for age-dependent tax on labor income

In order to demonstrate the motivation for an age-dependent tax on labor income I start by solving for the optimal tax policy in the simple model with the non-constant Frisch utility function where the government can condition labor income taxes on age and the government is allowed to borrow or save. I refer to this as the benchmark simple model. The formulation of the government's problem and resulting first order conditions for the benchmark simple model can be found in appendix 2.9.1.

Combining the household's and government's first order conditions simplifies to the following ratio

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \frac{1 + \lambda_t(1 + \frac{\sigma_2 h_{1,t}}{1 - h_{1,t}})}{1 + \lambda_t(1 + \frac{\sigma_2 h_{2,t+1}}{1 - h_{2,t+1}})}. \tag{2.15}$$

Equation 2.15 demonstrates that in a model with the non-constant Frisch utility function, the optimal tax on labor income varies by age if $h_{1,t} \neq h_{2,t+1}$. In the steady state, if $h_1^* > h_2^*$ then the optimal tax on labor income is such that $\tau_{h,1}^* > \tau_{h,2}^*$. Recall that Frisch labor supply elasticity for the non-constant Frisch utility function is $\frac{1-h}{\sigma_2 h}$. Therefore, one can interpret this result as $\tau_{h,1}^* > \tau_{h,2}^*$ when the Frisch labor supply elasticity rises over the agent's lifetime. The government prefers a higher

tax on the labor that is supplied less elastically as it limits the distortions imposed by the tax policy.

Zero tax on capital with age-dependent labor taxes

Utilizing the first order conditions from the Lagrangian with respect to capital and consumption leads to the following equation,

$$\left(\frac{c_{1,t}}{c_{2,t+1}} \right)^{-\sigma_1} = \beta(1+r). \quad (2.16)$$

Applying the non-constant Frisch utility function to equation 2.7 provides the following relationship

$$\left(\frac{c_{1,t}}{c_{2,t+1}} \right)^{-\sigma_1} = \beta(1+r(1-\tau_k)). \quad (2.17)$$

Equations 2.16 and 2.17 demonstrate that for the household to choose the optimal allocation from the primal approach the tax on capital must equal zero. Therefore, if the government can condition labor income taxes on age, and the government can borrow or save, then the optimal tax on capital is zero for the non-constant Frisch utility specification.

Non-zero tax on capital without age-dependent labor taxes

Next, I demonstrate that if the government would like to condition labor income taxes on age but is not allowed to do so then the optimal tax on capital is non-zero. I solve for the optimal tax policy in a model that is similar to the benchmark simple model but the government is not allowed to condition labor income taxes on age. The Lagrangian and first order conditions for this model are in appendix 2.9.1

Combining the governments first order conditions with respect to capital

and consumption leads to

$$\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\sigma_1} = \beta(1+r) \left(\frac{1 + \lambda_t(1 - \sigma_1) - \frac{\eta_t \epsilon_2 \sigma_1 (1-h_{1,t})^{-\sigma_2}}{c_{2,t+1}}}{1 + \lambda_t(1 - \sigma_1) + \frac{\eta_t \sigma_1 (1-h_{2,t+1})^{-\sigma_2}}{c_{1,t}}}\right). \quad (2.18)$$

Equations 2.17 and 2.18 demonstrate that for most consumption and labor profiles in order for the household to choose the optimal consumption profile from the primal approach the tax on capital is not zero. In the absence of the ability to condition labor income taxes on age, mimicking such an age-dependent tax on labor income is a motive for a non-zero tax on capital.

A tax on capital mimics age-dependent tax on labor

In order to demonstrate intuition as to why a capital tax is optimal when the government cannot condition the labor income tax on age, I derive the intertemporal Euler equation by combining equations 2.5, 2.6, and 2.7

$$\epsilon_2 \frac{U_{h1}(t)}{U_{h2}(t+1)} = \beta(1+r(1-\tau_k)) \frac{1-\tau_{h,1}}{1-\tau_{h,2}}. \quad (2.19)$$

Equation 2.19 demonstrates that if the government wants to tax labor income at different rates, then τ_k is an imperfect alternative. A tax on capital imperfectly mimics an age-dependent tax on labor income by creating a similar wedge on the marginal rate of substitution in the Euler equation.¹⁵ Specifically, a positive tax on capital mimics a relatively higher tax rate on young labor income since it creates a similar impact on the right hand side of equation 2.19.

Impact of constant Frisch elasticity on optimal tax policy

Next, I examine the optimal tax policy in a model with the constant Frisch utility function where the government is allowed to borrow or save and can also condition labor taxes on age. This model is similar to the benchmark model but uses the constant Frisch utility function instead of the non-constant utility function.

¹⁵A non-zero tax on capital can only imperfectly mimic age-dependent taxes on labor income because the former provides one less degree of freedom so the government can no longer independently determine both the wedge and the overall revenue from the tax policy.

Making this change to the model isolates the effect of a varying Frisch elasticity on optimal tax policy. The formulation of the government's problem and their first order conditions for this model can be found in appendix 2.9.1

Combining the household's and government's first order equations generates the following expression for the optimal labor taxes

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \frac{1 + \lambda_t(1 + \frac{1}{\varepsilon_2})}{1 + \lambda_t(1 + \frac{1}{\varepsilon_1})} = 1. \quad (2.20)$$

Equation 2.20 demonstrates that even if the government could condition taxes on age then they would tax labor income for different aged individuals at the same rate. Under the constant Frisch utility specification, it is not optimal to vary the labor income tax rate based on age since the Frisch elasticity is constant. Therefore, using this utility function eliminates the motive of taxing capital in order to mimic an age-dependent tax.

Further utilizing the government's first order conditions leads to the same expression as 2.16. Applying the constant Frisch utility function yields equation 2.17. Once again, these two equations imply that the optimal tax on capital is zero for the constant Frisch utility specification.¹⁶ In this model, with a utility specification where the Frisch elasticity is constant, there is no desire to tax capital in order to mimic an age-dependent tax on labor income.

2.2.4 Government balanced budget leads to non-zero tax on capital

When solving for the optimal tax policy, it is commonly assumed that the government cannot borrow or save. Therefore, in order to demonstrate the impact of the government being allowed to borrow or save, I solve for the optimal tax policy in a model similar to the simple benchmark model but one in which the government cannot borrow or save. Comparing this model to the simple benchmark

¹⁶If the government cannot condition labor income taxes on age the Lagrangian includes the constraint from equation 2.11. However, with the constant Frisch utility specification this constraint is not binding so the optimal tax on capital is still zero. See Conesa et al. (2009) for further discussion.

model demonstrates that not allowing the government to borrow or save is an additional motive for a non-zero tax on capital. I solve this model using the non-constant Frisch utility function with age-dependent taxes on labor income in order to demonstrate that this motive for a non-zero tax on capital is independent of the motive to mimic an age-dependent tax on labor income. Appendix 2.9.1 contains the government's problem and the first order conditions.

Combining the government's first order conditions with respect to capital and consumption yields

$$\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\sigma_1} = \beta(1+r) \left(\frac{1 + \lambda_{t+1}(1 - \sigma_1)}{1 + \lambda_t(1 - \sigma_1) - \varphi_t\left(\frac{1}{c_{1,t}^{\sigma_1}} + \frac{\chi\sigma_1(1-h_{1,t})^{-\sigma_2}}{c_{1,t}}\right)}\right). \quad (2.21)$$

Comparing equations 2.21 and 2.17 demonstrates that the optimal tax on capital is no longer zero in this model as long as the government balanced budget constraint is binding.

2.2.5 Summary of analytic results

In this section I demonstrate that with the non-constant Frisch utility function, it is optimal to condition labor income taxes on age. If the government is not able to condition labor income taxes on age then it is optimal to tax capital to mimic an age-dependent labor income tax. Even if the government can condition taxes on age, it is still optimal to tax capital if they are not allowed to borrow or save. In order to test the relevant magnitude of these and the other motives for a non-zero tax on capital I examine a computational calibrated overlapping generations model.

2.3 Computational Model

In order to analyze the motivating factors for the robust positive optimal tax on capital, I begin by computationally solving for the optimal tax policy in a benchmark overlapping generations general equilibrium model. Next, I examine

the optimal tax policy in a model where the government can consume accidental bequests and under the constant Frisch utility function. I choose to examine these two features because the first confounds a motive for a positive tax on capital with the desire of the government to consume accidental bequests and the second has only limited empirical motivation. Finally, I eliminate each of the potential motives for a non-zero tax on capital from the benchmark model. I describe these candidates in section 2.5. In this section, I describe the computational model (focusing on the benchmark model) and the definition of a stationary competitive equilibrium.

2.3.1 Demographics

In the computational model, time is assumed to be discrete and there are J overlapping generations. Ψ_j is the probability of an agent living to age $j + 1$ conditional on being alive at age j . All agents who live to an age of J die the next period. Agents retire at an exogenously set age j_r .

In each period a continuum of new agents is born. The population of new agents born each period grows at rate n . Given the population growth rate and conditional survival probabilities, the time invariant cohort shares, $\{\mu_j\}_{j=1}^J$, are given by

$$\mu_j = \frac{\Psi_{j-1}}{1+n} \mu_{j-1}, \text{ for } i = 2, \dots, J, \quad (2.22)$$

where μ_1 is normalized such that

$$\sum_{j=1}^J \mu_j = 1 \quad (2.23)$$

I use two different treatments for accidental bequests. In the benchmark model, the government is not allowed to consume accidental bequests and instead redistributes them to all remaining agents (Tr_t). The government taxes the capital income from these transfers at the same rate as ordinary capital income. In this treatment, the agents have no control over the transfers that they receive so they are equivalent to inelastically supplied capital income. Therefore, the optimal tax

policy would be for the government to confiscate and consume these transfers. As part of the experiment, the second treatment I use is to allow the government to consume these assets removing this motive for a positive tax on capital.

2.3.2 Individual

An individual is endowed with one unit of productive time per period which he splits between providing labor services and leisure in order to maximize his lifetime utility

$$\left\{ \sum_{j=1}^J \beta^j \prod_{q=1}^{j-1} \Psi_{q-1} u(c_j, h_j) \right\}, \quad (2.24)$$

where c_j is the consumption of an agent at age j and h_j is the hours spent providing labor services. Agents discount the next period's utility by the product of Ψ_j and β . The discount factor conditional on surviving is β and the unconditional discount rate is $\beta\Psi_j$.

An agent's age-specific human capital is ϵ_j so he receives labor income of $h_j\epsilon_jw_t$. Agents split their labor income between consumption and savings. An agent can save by purchasing a risk free asset. An agent's level of assets are denoted by a_j and he receives a pre-tax net return of r_t on the assets per period. Agents being liquidity constrained early in their life is another potential motive for a positive tax on capital. In some of the iterations, I test this motive's strength by allowing agents to borrow. In these iterations of the model, agents pay the actuarially fair interest rate of $r_{b,j,t} = \frac{r_t}{\Psi_j}$ to borrow.

2.3.3 Firm

Firms are perfectly competitive with constant returns to scale production technology. Aggregate technology is represented by a Cobb-Douglas production function. The aggregate resource constraint is,

$$C_t + K_{t+1} - (1 - \delta)K_t + G_t \leq K_t^\alpha N_t^{1-\alpha}, \quad (2.25)$$

where K_t , C_t , and N_t represent the aggregate capital stock, aggregate consumption, and aggregate labor (measured in efficiency units), respectively. Additionally, α is the capital share and δ is the depreciation rate for physical capital.

2.3.4 Government Policy

The government consumes resources in an unproductive sector, G_t .¹⁷ The government has two fiscal instruments to finance their consumption in the benchmark model. First, the government taxes capital income, $y_k \equiv r_t(a + Tr_t)$, according to a capital income tax schedule $T^K[y_k]$. Second, the government taxes each individual's taxable labor income. Part of the pre-tax labor income is accounted for by the employer's contributions to social security, which is not taxable under current US tax law. Therefore, the taxable labor income is $y_l \equiv w_t s_j h_j (1 - .5\tau_{ss})$, which is taxed according to a labor income tax schedule $T^L[y_l]$. I impose two restrictions on the labor and capital income tax policies. First, I assume anonymity of the tax code so the rates cannot be personalized, nor can they be age-dependent. Second, both of the taxes are functions only of the individual's relevant taxable income in the current period.

In some iterations of the model I allow the government to borrow or save in order to quantify its impact on optimal tax policy. I solve for the optimal tax policies under a steady state equilibrium so the government's level of savings or debt can not change over time. Therefore, in these iterations the government holds a fixed level of savings or debt but is still not allowed to run a deficit or surplus. When the government holds savings the return on its capital is used to offset government consumption and when the government is in debt it finances the interest payments on its debt by taxing individuals.

In addition to taxing income in order to finance G_t , the government runs a pay-as-you-go social security system in the benchmark model. The government pays SS_t to all individuals that are retired. Social security benefits are such that retired agents receive an exogenously determined fraction, b_t , of the average income

¹⁷A formulation that induces the same optimal tax policy is if the G_t enters the agents utility function in an additively separable manner.

of all working individuals. An agent's social security benefits are independent of his personal earnings history. Social security is financed by taxing labor income at a flat rate, $\tau_{ss,t}$. The payroll tax rate $\tau_{ss,t}$ is set to assure the social security system has a balanced budget each period. The social security system is not considered part of the tax policy that the government optimizes. In other iterations of the model, I eliminate the social security program in order to determine its impact on the optimal tax policy.

2.3.5 Definition of Stationary Competitive Equilibrium

In this section, I define the competitive equilibria for the benchmark model. I do not define the competitive equilibrium for the other iterations of the model since it is easy to determine the alternative competitive equilibria from the benchmark equilibrium.

Given a social security replacement rate b , government expenditures G , and a sequence of population shares $\{\mu_j\}_{j=1}^J$, a stationary competitive equilibrium is a sequence of agent allocations, $\{c_j, a_{j+1}, h_j\}$, a production plan for the firm (N, K) , a government labor tax function $T^l : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, a government capital tax function $T^k : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, a social security tax rate τ_{ss} , a utility function $U : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$, social security benefits SS , prices (w, r) , and transfers Tr such that:

1. Given prices, policies, transfers, and benefits the agent maximizes the following

$$\sum_{j=1}^J \text{Max}_{c_j, h_j, a_{j+1}} \beta^{j-1} \left[\prod_{q=0}^{j-1} \Psi_q \right] u(c_j, h_j)$$

subject to

$$c_j + a_{j+1} = w\epsilon_j h_j - \tau_{ss} w s_j h_j + (1+r)(a_j + Tr) - T^l[w\epsilon_j h_j(1 - .5\tau_{ss})] - T^k[r(a_j + Tr)] \text{ for } j < j_r,$$

$$c_j + a_{j+1} = SS + (1+r)(a_j + Tr) - T^k[r(a_j + Tr)], \text{ for } j \geq j_r$$

$$c \geq 0, 0 \leq h \leq 1, a_j \geq 0, \text{ and } a_1 = 0.$$

2. Prices w and r satisfy:

$$r = \alpha \left(\frac{N}{K} \right)^{1-\alpha} - \delta \text{ and } w = (1 - \alpha) \left(\frac{K}{N} \right)^\alpha$$

3. The social security policies satisfy:

$$SS = b \frac{wN}{\sum_{j=1}^{j_r-1} \mu_j} \text{ and } \tau_{ss} = \frac{ss \sum_{j=j_r}^J \mu_j}{w \sum_{j=1}^{j_r-1} \epsilon_j \mu_j}.$$

4. Transfers are given by:

$$Tr = \sum_{j=1}^J \mu_j (1 - \Psi_j) a_{j+1}.$$

5. Government budget balance:

$$G = \sum_{j=1}^J \mu_j T^k [r(a_j + Tr)] + \sum_{j=1}^{j_r-1} \mu_j T^l [w \epsilon_j h_j (1 - .5\tau_{ss})].$$

6. Market clearing:

$$K = \sum_{j=1}^J \mu_j a_j, \quad N = \sum_{j=1}^J \mu_j \epsilon_j h_j \text{ and}$$

$$\sum_{j=1}^J \mu_j c_j + \sum_{j=1}^J \mu_j a_{j+1} + G = K^\alpha N^{1-\alpha} + (1 - \delta)K.$$

2.4 Calibration

In this section, I describe the the functional forms and calibration. Calibration involves two steps. The first step is choosing parameter values for which there are direct estimates in the data. Second, in order to calibrate the remaining parameters, I choose values such that under the baseline fitted US tax policy certain target values are the same in the models and the US economy. When possible, I hold the value of the calibration parameters constant between different iterations of the model. The one notable exception is when the utility parameters change because I use different utility specification. Table 2.1 lists all the parameter values.

Table 2.1: Calibration Parameters

Parameter	Non-const. Calibration	Constant Calibration	Target
Demographics			
Retire Age: j_r	65	65	By Assumption
Max Age: J	100	100	By Assumption
Surv. Prob: Ψ_j	Bell and Miller (2002)	Bell and Miller (2002)	Data
Pop. Growth: n	1.1%	1.1%	Data
Preferences			
Conditional Discount: β	0.993	0.993	$K/Y = 2.7$
Risk aversion: σ_1, ς_1	2	2	Conesa et al. (2009)
Frisch Elasticity: σ_2, ς_2	3	$\frac{2}{3}$	Frisch = $\frac{2}{3}$
Disutility to Labor: χ	1.9	35.1	Avg. $h_j = \frac{1}{3}$
Technology			
α	.36	.36	Data
δ	8.33%	8.33%	$\frac{I}{Y} = 25.5\%$
A	1	1	Normalization
Government			
λ_0	.258	.258	Gouveia and Strauss (1994)
λ_1	.768	.768	Gouveia and Strauss (1994)
G	0.135	0.135	17% of Y

2.4.1 Demographics

In the model, agents are born at a real world age of twenty which corresponds to a model age of one. Agents are exogenously forced to retire at a real world age of sixty-five. If an individual survives until one hundred (model age 80) then he dies the next period. I use Bell and Miller (2002) to determine the conditional survival probabilities. I assume a population growth rate of 1.1%.

2.4.2 Individual

As a benchmark specification I use the non-constant Frisch utility function. In order to determine the impact of the desire to mimic an age-dependent tax on the optimal tax policy, I also find the optimal tax policy using the constant

Frisch utility function. This utility function eliminates the desire to condition labor income taxes on age.

I determine β such that the capital to output ratio matches US data of 2.7 in the benchmark model.¹⁸ I determine χ such that under the baseline fitted US tax policy agents work on average a third of their time endowment in the benchmark model. Following Conesa et al. (2009) I set $\varsigma_1 = \sigma_1 = 2$ which controls the relative risk aversion. I set $\sigma_2 = 3$ for the non-constant Frisch utility function which implies a Frisch labor supply elasticity of $\frac{2}{3}$ when agents are working a third of their time endowment. Under the constant Frisch utility function, I set $\varsigma = \frac{2}{3}$ which also implies a Frisch elasticity of two thirds. Past micro-econometric studies estimate the Frisch elasticity between 0 and 0.5. For examples see Altonji (1986), MaCurdy (1981) and Domeij and Flodén (2006). However, more recent research has identified that these estimates may be biased downward. Some of the reasons for the bias are: utilizing weak instruments, not accounting for borrowing constraints, disregarding the life cycle impact of endogenous-age specific human capital and omitting correlated variables such as wage uncertainty. Some of these studies include Imai and Keane (2004), Domeij and Flodén (2006), Pistaferri (2003) and Contreras and Sinclair (2008) Rogerson and Wallenius (2009) show that because individuals choose their labor supply on both the intensive and extensive margin “micro and macro elasticities need not be the same, and that macro elasticities can be significantly larger.” Furthermore, Chetty (2009) shows that small frictions in the labor market can lead the observed Frisch elasticity to be much smaller. Since there is some uncertainty about this value, I test the sensitivity of the results with regards to this parameter in section 2.7. I calibrate $\{\epsilon_j\}_{j=0}^{j_r-1}$ such that the sequence matches a smoothed version of the relative hourly earnings estimated by age in Hansen (1993).

¹⁸This is the ratio of fixed assets and consumer durable goods less government fixed assets to GDP.

2.4.3 Firm

I assume the capital share parameter, α , is .36. The depreciation rate is set to target the observed investment-output ratio of 25.5%.

2.4.4 Government Policy

In order to calibrate the parameters, I need a benchmark tax function to use when matching the targets in the models to the values in the data. I calibrate the model under a baseline tax function that mimics the US tax code. I refer to this tax function as the fitted US tax policy. I use the estimates from Gouveia and Strauss (1994) to determine the fitted US tax policy. The authors match the US tax code to the data using a three parameter functional form,

$$T(y; \lambda_0, \lambda_1, \lambda_2) = \lambda_0(y - (y^{-\lambda_1} + \lambda_2)^{-\frac{1}{\lambda_1}}) \quad (2.26)$$

where y represents the sum of labor or capital income. The average tax rate is principally controlled by λ_0 , and λ_1 governs the progressivity of the tax policy. λ_2 is left free in order to ensure that the tax policy satisfy the budget constraint.

Gouveia and Strauss (1994) estimate values of $\lambda_0 = .258$ and $\lambda_1 = .768$ from the US data. The authors do not fit separate tax functions for labor and capital income. Therefore, I use a uniform tax system on both sources of income for the baseline fitted US tax policy. I calibrate government consumption, G , such that it equals 17% of output under the baseline fitted US tax policy, as observed in the US data.¹⁹ Therefore, I set λ_2 (for both sources of income) at the value that equates government consumption to 17% of GDP. When determining the optimal tax policy, I restrict my attention to revenue neutral changes to the tax policy where the optimal tax policy is a separate flat tax rate on capital income and on labor income (τ_k and τ_h). This experiment restricts the government consumption such that it is equal under the baseline fitted US tax policy and the optimal tax policies.

In the benchmark model, the social security system is chosen so that the

¹⁹To determine this target I used government expenditures less defense consumption.

Table 2.2: Models

Model	Non-Const. Frisch	Gov't Dist. Tr	Ind. Liq. Const.	No Gov't Saving	No Gov't Borrowing	SS Program
A0:	Yes	Yes	Yes	Yes	Yes	Yes
B0:	No	No	Yes	Yes	Yes	Yes
A1:	No	Yes	Yes	Yes	Yes	Yes
A2:	Yes	No	Yes	Yes	Yes	Yes
A3:	Yes	Yes	No	Yes	Yes	Yes
A4:	Yes	Yes	Yes	No	Yes	Yes
A5:	Yes	Yes	Yes	Yes	No	Yes
A6:	Yes	Yes	Yes	Yes	Yes	No

Note: Model A0 is the benchmark, B0 is the alternative specification, A1 is the constant Frisch, A2 is the government consumes transfers, A3 is individual borrowing, A4 is government saves, A5 is government borrows, and A6 is no social security program.

replacement rate, b , is 50%.²⁰ The payroll tax, τ_{ss} , is determined such that the social security system is balanced each period.

2.5 Computational Experiment

The computational experiment aims to determine the relative strengths of the motives for a large optimal tax on capital in the benchmark model. I begin by solving for the optimal tax policy in my benchmark model. Next, I solve for a model that has a utility function that implies a constant Frisch elasticity and allows the government to consume accidental bequests. Finally, I examine the strength of each of the motives by eliminating each of them from the benchmark model. The aspects of the benchmark model that I change are: a varying Frisch labor supply elasticity profile, no separate tax rates on accidental bequests and ordinary capital income, individual borrowing constraints, not allowing the government to save, not allowing the government to borrow, and including a reduced form social security program. I solve for the optimal tax policy in a total of eight different iterations of the model. Table 2.2 list the features in each iteration of the model.

In order to quantify the effect of each of these features, I need to solve for the

²⁰The replacement rate matches the rate in Conesa et al. (2009) and Conesa and Krueger (2006).

optimal tax policy in each iteration of the model. When searching for the optimal tax policy, I limit my attention to flat taxes instead of searching over progressive tax policies. Conesa et al. (2009) and chapter 1 solve for the optimal tax policies in a model similar to the benchmark model. They both find that the optimal tax policies are flat taxes in models that do not include within cohort heterogeneity. Therefore, I restrict my attention to flat taxes since all the agents within a cohort are homogenous.

To quantify the optimal tax policy, I need a social welfare function. I choose a social welfare function that corresponds to a Rawlsian veil of ignorance (Rawls (1971)). Because living agents face no earnings uncertainty, given a stationary competitive equilibrium, the social welfare is equal to the expected lifetime utility of a newborn,

$$SWF(\tau_h, \tau_k) = \sum_{j=1}^J \beta^{j-1} \left[\prod_{q=0}^{j-1} \Psi_q \right] u(c_j, h_j) \quad (2.27)$$

where τ_h is the flat tax rate on labor income and τ_k is the flat tax rate on capital income. When I determine the optimal tax policy, I search over τ_h and leave τ_k free to satisfy the government's budget constraint.²¹ I require that any change in the tax policy is revenue neutral.

2.6 Results

In this section, I start by solving for the optimal tax policy in the benchmark model (A0) and the alternative specification (B0) in order to test the impact of the varying Frisch elasticity and accidental bequests assumptions on the optimal tax policy. Next, I change one of the features of the benchmark model and solve for the optimal tax policy (models A1, A2, A3, A4, A5, and A6) in order to determine each feature's individual impacts on optimal tax policy.

²¹Even when I do not require the government to balance their budget, I am able to solve for a unique τ_k because I solve for a tax policy with a specific level of government savings or debt.

2.6.1 Impact of Frisch Elasticity and Accidental Bequests

Table 2.3 lists the optimal tax policies from the benchmark model and the alternative specification model. I find that when I change the utility specification and allow the government to consume transfers the optimal tax on capital drops over twenty percentage points. Although the tax on capital remains positive, it is no longer large. Therefore, one assumption that has limited empirical motivation and one that confounds a motive for a positive tax on capital with the governments desire to consume accidental bequests are responsible for over seventy percent of the large optimal tax on capital. The next section documents the individual impact of all the model features on optimal tax policy and the economy.

Table 2.3: Optimal Tax Policies

Model	τ_k	τ_h	$\frac{\tau_k}{\tau_h}$
A0 (Benchmark):	29.4%	21.3%	1.4
B0 (Alternative Specification):	8.5%	19.2%	0.4

2.6.2 Determining Impact of Each Assumption

Table 2.5 describes the optimal tax policies and the aggregate economic variables in the seven iterations of the model that test each feature individually. Table 2.6 contains the percent change of the economic aggregate variables from the benchmark model. In this section I examine the effect of each feature on the optimal tax policy, the aggregate economic variables and life cycle profiles.

Table 2.4: Optimal Tax Policies

Model	τ_k	τ_h
A0 (Benchmark):	29.4%	21.3%
A1 (Constant Frisch):	19.5%	23.6%
A2 (Gov't Consumes Tr):	13.9%	18.2%
A3 (Ind. Borrowing):	22.5%	22.8%
A4 (Gov't Saves):	14.1%	19.3%
A5 (Gov't Borrows):	34.7%	28%
A6 (No SS):	-39.1%	24.9%

Table 2.5: Aggregate Economic Variables

Model	Y	K	N	w	r	tr	τ_{ss}
A0 (Benchmark):	0.81	2.16	0.47	1.11	0.052	0.024	11.7%
A1 (Constant Frisch):	0.81	2.22	0.46	1.13	0.049	0.026	11.7%
A2 (Gov't Consumes Tr):	0.85	2.41	0.47	1.15	0.044	0.028	11.7%
A3 (Ind. Borrowing):	0.82	2.21	0.47	1.12	0.05	0.025	11.7%
A4 (Gov't Saves):	0.86	2.59	0.46	1.19	0.036	0.02	11.7%
A5 (Gov't Borrows):	0.76	1.78	0.47	1.04	0.07	0.03	11.7%
A6 (No SS):	0.99	3.66	0.47	1.34	0.014	0.058	0%

Table 2.6: Percent Changes Induced by New Optimal Tax Policy in Model

Model	Y	K	N	w	r	tr
A1:	0.5%	3.1%	-0.9%	1.5%	-6.7%	4.5%
A2:	4.8%	11.8%	1.1%	3.7%	-16.3%	15.8%
A3:	0.7%	2.5%	-0.3%	1%	-4.5%	4.1%
A4:	6.2%	20.1%	-0.9%	7.2%	-30%	-16.3%
A5:	-6.5%	-17.3%	0.2%	-6.7%	33.9%	21.7%
A6:	22%	69.8%	1.3%	20.4%	-73.1%	135.7%

Notes: Each row is the percent change from the benchmark model (A0). For example, A1 is the percent change between A0 and A1.

2.6.3 Desire to mimic age-dependent tax

The first assumption I alter is changing the utility function such that the Frisch elasticity is constant. I demonstrate in section 2.2.3 that utilizing the constant Frisch utility function instead of the non-constant Frisch utility function eliminates the government's desire to condition taxes on age. Models A0 and A1 are the same other than A0 uses the non-constant Frisch utility function and A1 uses the constant Frisch utility function. I find that eliminating this channel for a tax on capital causes the optimal tax on capital to drop approximately ten percentage points (see table 2.4). In response to the drop in the tax on capital, the tax on labor income rises just over two percentage points. Generally the aggregate economic variables look similar in the two models with only modest changes. There is a small drop in aggregate labor and a rise in aggregate capital which lead to a small rise in the wage rate and a drop in the rental rate.

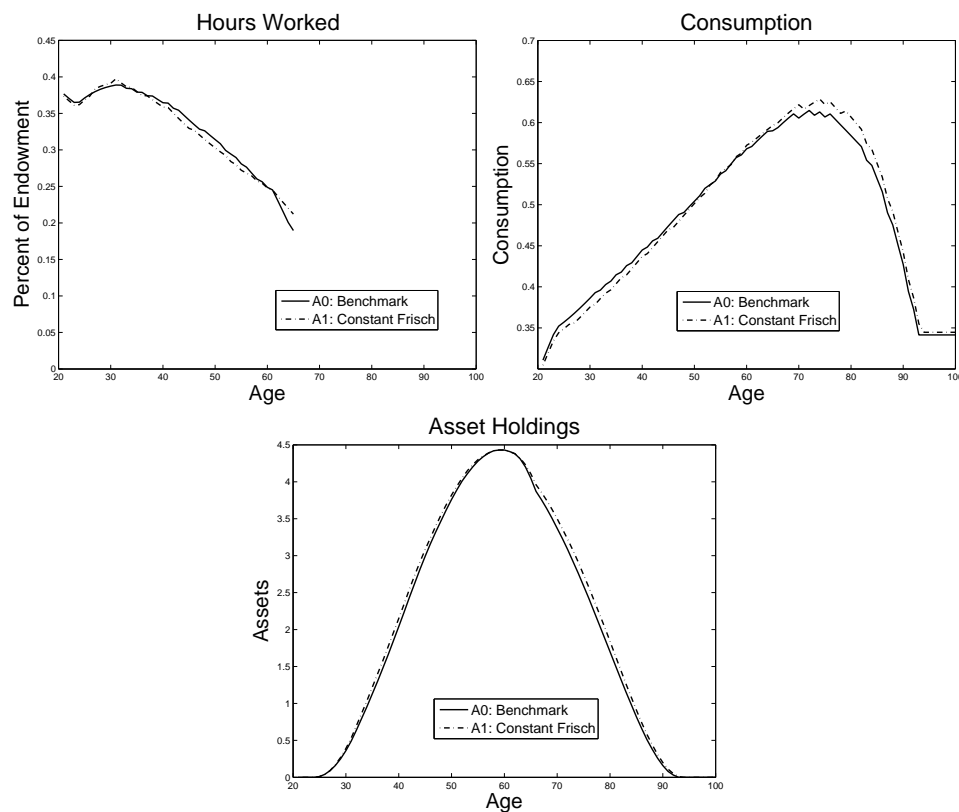


Figure 2.1: Life Cycle Profiles in Model A0 and A1

Figure 2.1 plots the life cycle profiles for labor supply, consumption and savings in the benchmark model (A0) and the model that eliminates the desire to condition taxes on age (A1). Generally, the life cycle profiles in the two models look similar. In the benchmark model, an agent's Frisch labor supply elasticity is negatively related to the hours they work. Therefore, in the benchmark model, agents become more elastic towards the end of their life when their hours decrease. In model A1, an agent's Frisch labor supply elasticity is constant. Therefore, agents tend to be relatively less elastic in their middle years and more elastic in their later years in model A0 compared to in model A1. This difference in relative elasticity combined with the relatively lower implicit tax on older labor income leads agents to work less hours in their middle years and more hours in their later working years in model A1 (see upper left panel of figure 2.1).

The change in the marginal after-tax return in A1 also affects the shape of

the lifetime consumption profile. The intertemporal Euler equation controls the slope of consumption profile over an agent's lifetime. The relationship is,

$$\left(\frac{c_{j+1}}{c_j}\right)^{\sigma_1} = \Psi_j \beta \tilde{r}_t \quad (2.28)$$

where \tilde{r}_t is the marginal after-tax return on capital. Since the marginal after-tax return on capital is larger in A1 than in A0, the consumption profile, in the upper right panel of figure 2.1, is steeper. The larger after-tax return on capital causes agents to increase savings (see bottom panel of figure 2.1).

2.6.4 Government consumption of accidental bequests

Next, I examine the effect of the government being restricted to tax accidental bequests, or transfers, at the same rate as ordinary capital income on the optimal tax policy. The optimal tax policy will tax inelastically supplied income at a relatively higher rate. Accidental bequests are inelastic income. However, in model A0, the government is restricted to taxing the returns on these transfers at the same rate as other capital income. Additionally, the government is not allowed to tax the principal of the bequests. Since the tax on capital in the benchmark model (A0) is a hybrid tax on ordinary capital income and accidental bequests, the optimal tax rate is a weighted average of the optimal rates on each income. In model A2, I allow the government to tax these incomes at different rates. In model A2, the government fully consumes accidental bequests and the optimal tax on capital only represents the optimal tax on ordinary capital income.

Comparing line one and three in table 2.4 shows that eliminating this motive for a positive tax on capital causes the optimal tax on capital income to drop by just over fifteen percentage points. In addition, the tax on labor income also falls. Both tax rates drop in this model because the government has to raise less revenue from income taxes since they have the additional revenue generated by fully confiscating and consuming the accidental bequests. However, the tax rate on capital income fall much more than the tax rate on labor income. Even in this model, where I include a varying Frisch elasticity the optimal tax on capital

drop by over fifty percent. Therefore, the large optimal tax on capital result is not robust to allowing the government a richer policy set.²²

Examining tables 2.5 and 2.6, the decreases in the tax rates cause agents to save more and work almost the same amount so there is an increase in aggregate output in model A2. The increase in capital and decrease in labor causes the return on capital to drop and the wage rate to increase. Since agent's have larger levels of savings, their accidental bequests also increase.²³

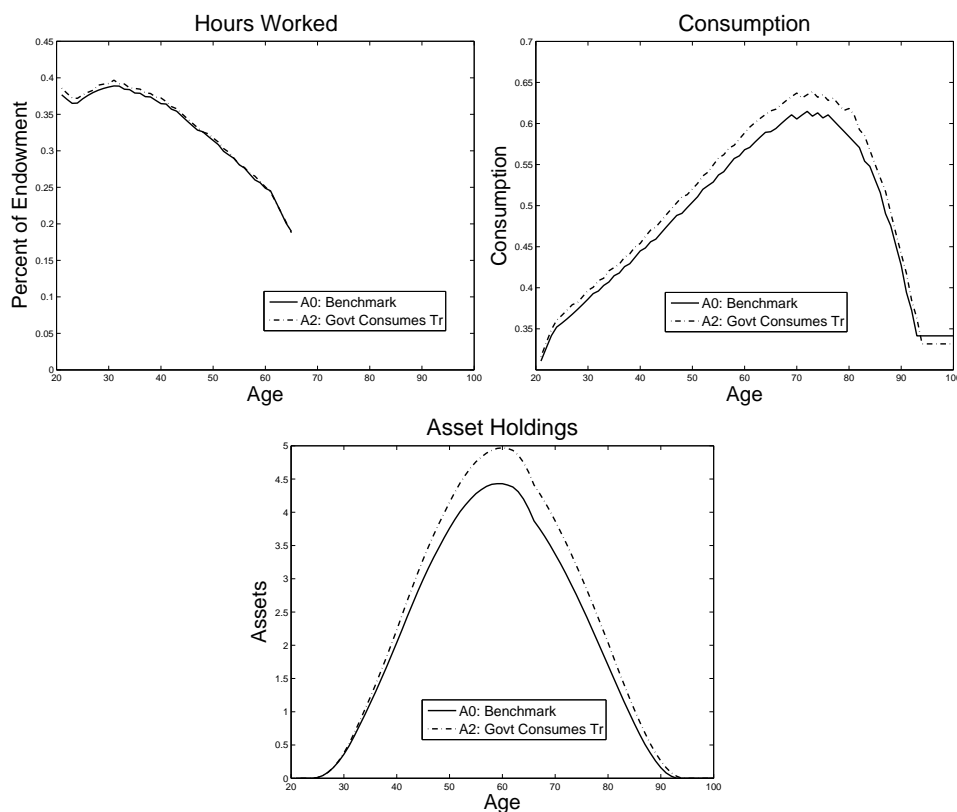


Figure 2.2: Life Cycle Profiles in Model A0 and A2

Figure 2.2 plots the life cycle profiles in model A0 and A2. Generally the shapes of the hours profiles look similar, although there is an upward shift in hours worked in model A2 since the economy is bigger. Additionally, in model A2, agents

²²I find that an individual's utility in model A2 is higher than A0 so the government using accidental bequests to finance government consumption is welfare improving

²³Note, in model A2 these bequests are consumed by the government as opposed to being redistributed to living agents.

no longer receive income from transfers each period. Agents compensate for this lost income by accumulating more assets. The larger economy causes consumption to generally increase.

2.6.5 Individual Liquidity Constraints

There are two forces changing the optimal tax on capital when I change the model to allow individual's to borrow. First, agents prefer to smooth their consumption. Therefore, when an agent faces a hump-shaped lifetime earnings profile he would prefer to smooth his consumption by borrowing against earnings from later years to facilitate consumption in earlier years. Borrowing constraints hinder an agent's ability to shift consumption, creating a role for tax policy to help facilitate this shift. Since an individual typically accumulates more assets later in their life, increasing the tax on capital income and decreasing the tax on labor income will allocate more of the lifetime tax burden to an individual's later years, which facilitates consumption smoothing. Therefore, restricting agents from borrowing can motivate a positive tax on capital. Second, when agents are allowed to borrow I find that agents decrease their labor supply early in their life because they are able to utilize borrowing (see the upper left panel of figure 2.3). This shift in hours affects the labor supply elasticity in model A3, causing young agents to supply labor more elastically than in model A0. This change in relative elasticity leads to a decrease in the desire to mimic an age-dependent tax on labor income. In order to determine the overall affect, I compare the optimal tax policies in a model where agents are not able to borrow (A0) and one where agents can borrow at the actuarially fair rate (A3). I find that when I eliminate individual liquidity constraints the optimal tax on capital falls just under seven percentage points.

Overall, allowing agents to borrow does not have a dramatic impact on the models. Comparing the third row of table 2.6 to the other rows, it is clear that eliminating liquidity constraints has a relatively small impact on the aggregate economic variables.

The lower panel of figure 2.3 demonstrates that an agent's borrowing constraint is only binding in the first few years of their life. Therefore, eliminating

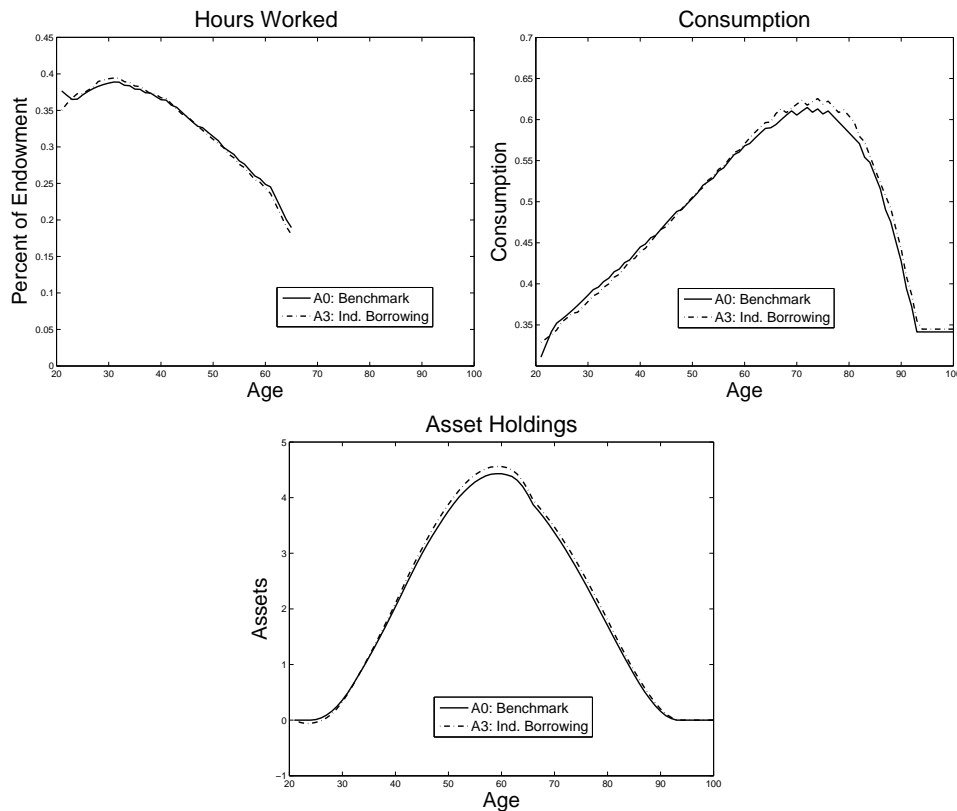


Figure 2.3: Life Cycle Profiles in Model A0 and A3

borrowing constraints only alters an agents hours and consumption decisions in the first few years of their life (see the upper left panel of 2.3). After the first five years, the life cycle profiles in model A0 and A3 look similar. Because eliminating borrowing constraints has a minimal impact on agents, the optimal tax policies are similar in the two models.

2.6.6 Government Savings and Debt

In section 2.2.4, I analytically demonstrate that restricting the government from saving or borrowing motivates a non-zero tax on capital. In order to quantify the strength of this motive, I examine the optimal tax policy when the government saves (A4) and when they borrows (A5). I examine the model when the government savings or government debt equals 550% of their annual consumption. I use this

number because the relative government debt to government expenditures (less defense consumption) is approximately 550%.

Savings

Comparing models A0 and A4, it is clear that this level of government savings has a large impact on the economy. It causes over a fifteen percentage point drop in the optimal tax on capital and a two percentage point drop in the optimal tax on labor income (see table 2.4). The aggregate capital increases over twenty percent because the government now holds capital in addition to agents (see table 2.6). The decrease in the overall tax burden that the an agent faces leads to approximately a one percent decrease in aggregate labor. The large rise in capital coupled with a small decrease in labor causes output to rise just more than six percent. Since aggregate labor stays relatively constant while aggregate capital rises, wages rise approximately seven percent and the rental rate decreases approximately thirty percent. The decrease in the tax on capital and decrease in the rental rate on capital have opposing affects on the after tax return on capital so it only drops a modest amount. Although overall capital increases in model A4, the level of private savings decreases approximately fifteen percent. The lower level of private savings causes transfers to decrease.

Examining the bottom panel of figure 2.4, it is clear that the fall in the after tax return on capital causes agents to hold less savings. The fall in the rental rate also causes agents to value savings relatively less in their early years leading them to shift hours to later years and consumption to earlier years. The fall in the after tax return on capital causes an agent's consumption profile to flatten.

Debt

Including government debt in the model causes an opposite reaction to the optimal tax policy compared to when the government holds savings. The tax on capital increases over five percentage points and the tax on labor increases almost seven percentage points (see table 2.4) in model A5 Comparing the fourth and fifth line of table 2.6, it is clear that the government holding debt causes the economic

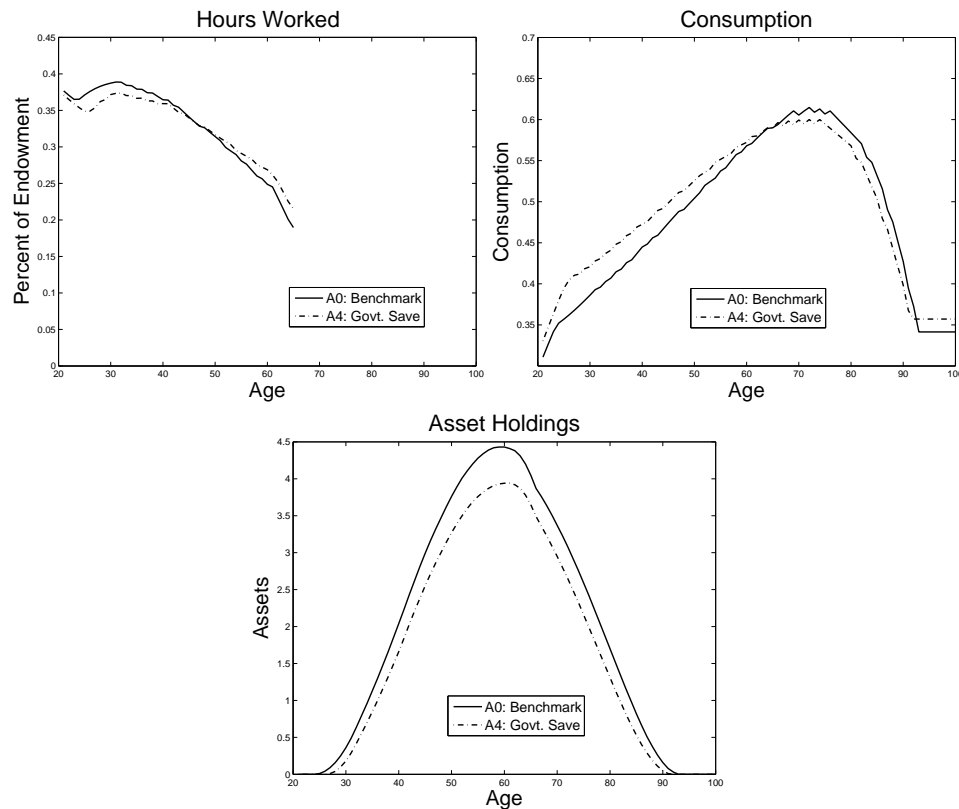


Figure 2.4: Life Cycle Profiles in Model A0 and A4

aggregate variables to have an opposite reaction (with a similar magnitude) to the government holding savings.

The impact of the government holding debt on the life cycle profiles is also opposite to the effect when the government holds savings (compare figures 2.4 and 2.5). The government holding debt causes the after tax return on capital to increase which leads to a steepening of the consumption profile (see the upper right panel of figure 2.5). Because agents do not consume as much early in their life, they are no longer liquidity constrained in their early years. Since agents are not liquidity constrained, they no longer work less hours early in their life so the labor supply profile is smoother in those years.

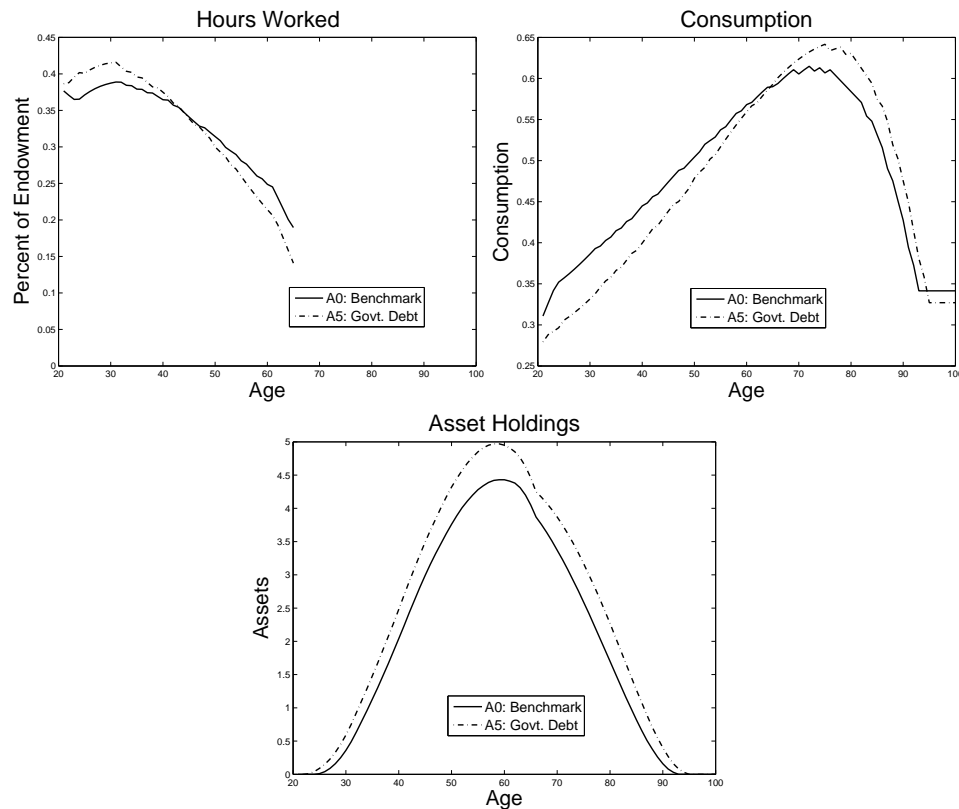


Figure 2.5: Life Cycle Profiles in Model A0 and A5

2.6.7 Social Security Program

Model A6 examines the impact of the social security program on the optimal tax policy by eliminating it from model A0. Eliminating the social security program has a large impact on the optimal tax policy. When the program is excluded, the optimal tax on capital decreases from approximately seventy percentage points to negative thirty-nine percent. Without social security, the government provides a rebate on capital income in order to imperfectly mimics a welfare increasing social security program.

Since agents must finance their own retirement with personal savings and the tax on capital decreases, aggregate capital increases almost seventy percent (see table 2.6). The rise in capital coupled with a small rise in labor causes aggregate output to rise, the wage rate to increase and the rental rate on capital to fall.

Excluding the social security program causes the life cycle consumption

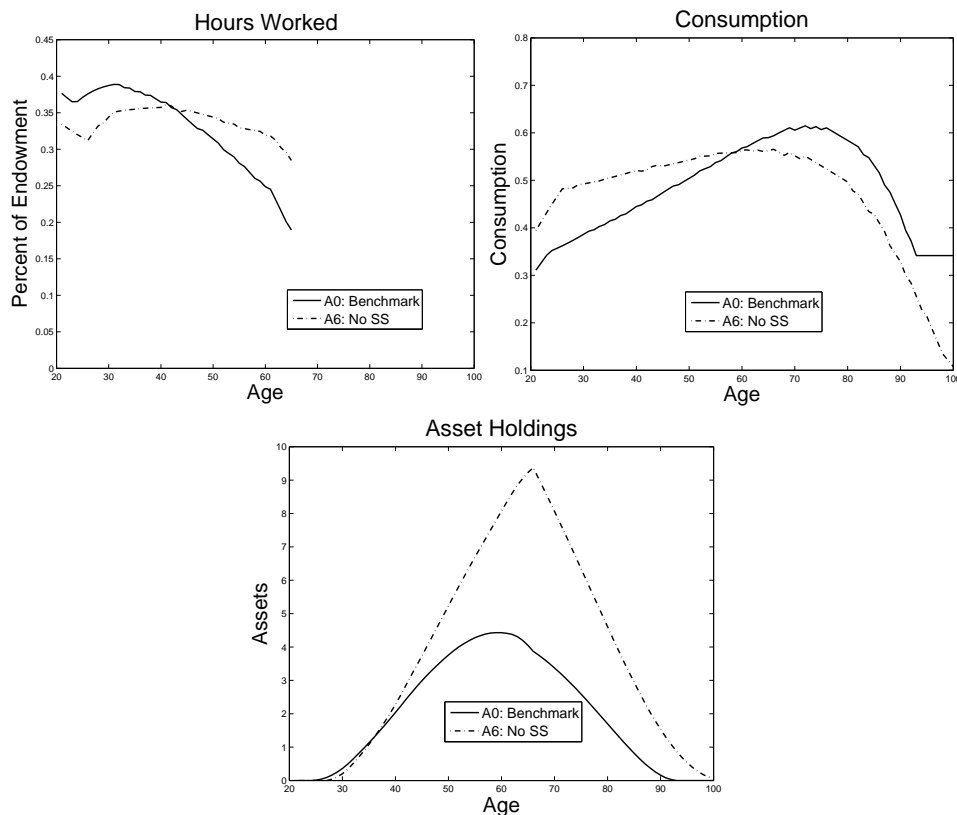


Figure 2.6: Life Cycle Profiles in Model A0 and A6

and savings profiles to have less realistic shapes. The upper left and right panel in figure 2.6 demonstrate that eliminating the social security program causes the labor profile and consumption profile to be flatter while an agent is working. The profiles are flatter because the rental rate on capital is lower. Since agents face lifetime uncertainty and finance their own retirement consumption in model A6, their consumption falls much more dramatically towards the end of their life. Additionally, agents accumulate more assets to finance retirement so their lifetime savings profile shifts upwards in model A6.

2.6.8 Summary of Results

Overall, I find that assuming the government is not allowed to save, the Frisch elasticity is non-constant, and the government cannot consume accidental

bequests are significant motives for a positive tax on capital. Removing individual budget constraints has a minimal impact on optimal tax policy. Additionally, including a reduced form social security program is important because it causes the optimal tax policy and life cycle profiles to be more realistic.

2.7 Sensitivity Analysis

Next, I check the sensitivity of the results with respect to the parameter value that governs the Frisch labor supply elasticity. I choose to examine the sensitivity of the results with respect to this parameter because there is some uncertainty about the actual value of the Frisch elasticity. In this section I test how using different Frisch elasticity parameters affects the optimal tax policy in the benchmark model and the impact of each of the model features on optimal tax policy. I solve for the optimal tax policy in model A0 with three different value of σ_2 of 6, 3 and 2. These values imply a Frisch elasticity of $\frac{1}{3}$, $\frac{2}{3}$ and 1, respectively, if an agent works a third of their endowment. I also solve for models A1-A6 with the different values.²⁴ Prior to solving the models, I need to calibrate the benchmark model with the three different targets for the Frisch elasticity. Table 2.7 lists these parameters.

Table 2.7: Calibration Parameters

Parameter	Frisch= $\frac{1}{3}$	Frisch= $\frac{2}{3}$	Frisch= 1
Conditional Discount: β	0.996	0.993	0.99
Risk aversion: σ_1, ς_1	2	2	2
Non-constant Frisch Utility			
Frisch Elasticity: σ_2	6	3	2
Disutility to Labor: χ	0.59	1.9	2.8
Constant Frisch Utility			
Frisch Elasticity: ς_2	$\frac{1}{3}$	$\frac{2}{3}$	1
Disutility to Labor: χ	182.2	35.1	20

²⁴In model A1 I use the values of $\frac{1}{3}$, $\frac{2}{3}$ and 1 for ς .

2.7.1 Effect on Optimal Tax Policies in Benchmark Models

Table 2.8 presents the optimal tax policies in the benchmark model calibrated to target the three different Frisch elasticities. I find that the optimal tax on capital is larger for the models with a Frisch elasticity of $\frac{2}{3}$ and 1. There are two reasons that the tax on capital increases with these two values for the parameter. When the government is deciding between taxing capital and labor income, they are weighing the relative distortions that each tax induces on the economy. An agent will be more reactive to a tax on labor income when the Frisch elasticity is higher. Therefore, the government prefers to reduce the tax on labor income and increase the tax on capital under the two calibrations that target a higher Frisch elasticity.

Table 2.8: Optimal Tax Policies in Benchmark Models with Different Frisch Elasticities

Tax Rates	Frisch= $\frac{1}{3}$	Frisch= $\frac{2}{3}$	Frisch= 1
τ_k	22.1%	29.4%	29.3%
τ_h	22.9%	21.3%	21.2%

The second reason that the optimal tax on capital increases when the model is calibrated with the higher Frisch elasticities is that it enhances the motive for an age-dependent tax on labor income. Figure 2.7 plots the life cycle profiles for the three different calibrations. The upper left panel of the figure demonstrates that as the model is calibrated to match the higher Frisch elasticities the relative change between the hours he works when he is young and old is larger.²⁵ In the medium and high model a larger drop in hours enhances the motive for an age-dependent tax on labor income. Since the government cannot condition labor income taxes on ages, they increase the tax on capital.

²⁵Although an agent work less hours when he is old in the model calibrated to target a high elasticity compared to the medium model, he also works less hours when he is young. Therefore, the relative drop in hours between young and old agents is similar in the medium and high model.

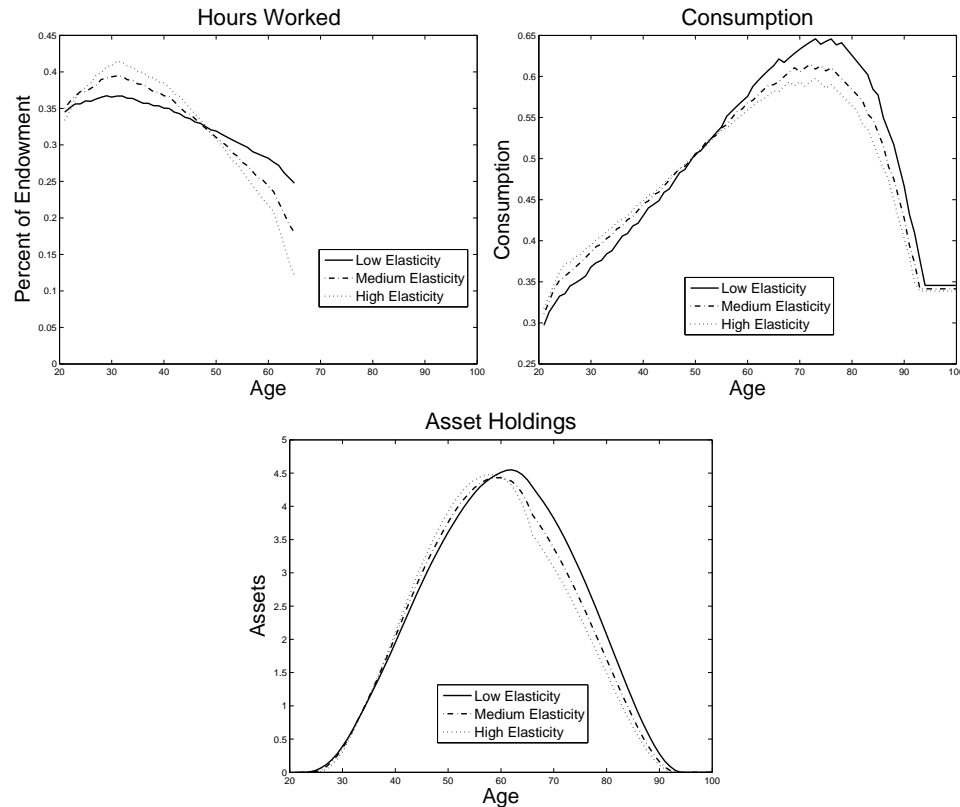


Figure 2.7: Life Cycle Profiles in Benchmark Models with Different Elasticities

2.7.2 Frisch Elasticity effect on Channels' Impact

In order to determine how changing the Frisch elasticity alters the impact on optimal tax policy of each of the channels I solve for models A0, A1, A2, A3, A4, A5 and A6 under the three different calibration parameters. Table 2.9 describes the optimal tax policies in the six models under the three different calibrations. Table 2.10 presents the percentage point changes in the optimal tax policies between the benchmark model (A0) and the various models (A1-A6) under all three calibrations. Generally, each of the channels have a larger impact on the optimal tax policy when the model is calibrated to match a medium or low elasticity.

The optimal tax on capital decreases in models A1, A2, A4 and A6 for all the calibrations. The optimal tax on capital increases in model A3 when the Frisch

Table 2.9: Optimal Tax Policy in Sensitivity Analysis Under Different Calibrations

Model	τ_k	τ_h	τ_k	τ_h	τ_k	τ_h
	(Low)		(Medium)		(High)	
A0:	22.1%	22.9%	29.4%	21.3%	29.3%	21.2%
A1:	12.1%	24.8%	19.5%	23.6%	19.3%	23.5%
A2:	10%	18.4%	13.9%	18.2%	22.3%	16.9%
A3:	22.6%	22.8%	22.5%	22.8%	27.3%	21.7%
A4:	-1.6%	21.5%	14.1%	19.3%	9.2%	19.9%
A5:	29.2%	29.9%	34.7%	28%	40.7%	26%
A6:	-84.6%	26.8%	-39.1%	24.9%	-31.2%	24.8%

Notes: Each row is the optimal tax policy for a model similar to the benchmark model with one channel removed. Low is calibrated with a target Frisch elasticity of $\frac{1}{3}$, medium it calibrated with a target Frisch elasticity of $\frac{2}{3}$ and high is calibrated with a target Frisch elasticity of 1.

Table 2.10: Percentage Point Changes Induced by Change in Model

Model	τ_k	τ_h	τ_k	τ_h	τ_k	τ_h
	(Low)		(Medium)		(High)	
A1 (Constant Frisch):	-10.1	2	-10	2.4	-10	2.3
A2 (Gov't Consumes Tr):	-12.1	-4.5	-15.5	-3.1	-7	-4.4
A3 (Ind. Borrowing):	0.5	-0.1	-6.9	1.6	-1.9	0.5
A4 (Gov't Saves):	-23.7	-1.4	-15.3	-2	-20.1	-1.3
A5 (Gov't Borrows):	7.1	7	5.3	6.7	11.5	4.8
A6 (No SS):	-106.7	4	-68.5	3.7	-60.5	3.6

Notes: Each row is the percentage point change form the benchmark model. For example, A1 is the percentage point change in the optimal tax policy between A0 and A1.

elasticity is calibrated to a low value, but increases for the other two calibrations. Model A3 eliminates individual borrowing constraints which has two counteracting effects on the optimal tax policy. First, borrowing decreases the overall capital so the government would need to increase the tax on capital in order to raise the same amount of revenue. Second, agents decrease the hours they work early in their life since they can borrow to finance consumption in the early periods when their labor is less productive (see figure 2.3). This shift in hours reduces the motive for a tax on capital because it decreases the disparity between the labor supply elasticity of an agent when he is young versus old. The second effect of removing liquidity constraints is diminished when the model is calibrated to match a lower Frisch

elasticity. Therefore, the first effect dominates in the model calibrated to a low elasticity and the second effect dominates in the models calibrated to a medium and high elasticity.

2.8 Conclusion

Through an analysis of the optimal tax on capital in a standard life cycle model, this paper concludes that if one alters the utility function such that the Frisch elasticity profile is constant and allows the government to tax accidental bequests at a separate rate from ordinary capital income then the optimal tax on capital falls from 29.4% to 8.5%. It is important to quantify the impact of these two model features because there is a lack of empirical motivation for an upward sloping elasticity profile and prohibiting the government taxing accidental bequests at a different rate confounds the government's desire to confiscate the bequests with a positive optimal tax on capital. Although the optimal tax on capital is not zero in the model without these features, it is no longer large.

I also find that if the government holds savings (debt) then the optimal tax on capital decreases (increases). Individual liquidity constraints have a small effect on optimal tax policy, and the direction of their impact depends on the specific calibration targets. I show that it is important to include at least a reduced form social security program in a life cycle analysis of optimal tax policy otherwise the optimal policy will try to replicate a welfare improving social security program with a negative tax on capital. If the model is calibrated to match a medium or low Frisch elasticity instead of a high value, it will enhance the impact of a varying Frisch elasticity profile, the government not being able to tax accidental bequests at a separate rate from ordinary capital income, the government holding savings, and excluding a social security program.

When modeling certain aspects of the economy, economists try to balance realism and tractability. I demonstrate that some of these simplifying assumptions have a sizable impact on optimal tax policy. Specifically, prohibiting the government from borrowing or saving, simplifying the estate tax system, and excluding

a reduced form social security program have large implications for optimal tax policy. Therefore, further research should focus on modeling these features more realistically and their impacts on optimal tax policy. Additionally, the shape of the Frisch labor supply elasticity profile has a large impact on the optimal tax policy. Since there is little existing empirical evidence addressing whether the Frisch elasticity varies, it is an important question for future work to address.

2.9 Appendix

2.9.1 Analytical Derivations

Benchmark Simple Model

The Lagrangian for the benchmark simple model is

$$\begin{aligned} \mathcal{L} = & \frac{c_{1,t}^{1-\sigma_1}}{1-\sigma_1} + \chi \frac{(1-h_{1,t})^{1-\sigma_2}}{1-\sigma_2} + \beta \frac{c_{2,t+1}^{1-\sigma_1}}{1-\sigma_1} + \chi \frac{(1-h_{2,t+1})^{1-\sigma_2}}{1-\sigma_2} \\ & - \rho_t (c_{1,t} + c_{2,t} + K_{t+1} - K_t + G_t - rK_t - w(h_{1,t} + h_{2,t}\epsilon_2)) \\ & - \rho_{t+1} \theta (c_{1,t+1} + c_{2,t+1} + K_{t+2} - K_{t+1} + G_{t+1} - rK_{t+1} - w(h_{1,t+1} + h_{2,t+1}\epsilon_2)) \\ & + \lambda_t (c_{1,t}^{1-\sigma_1} + \beta c_{2,t+1}^{1-\sigma_1} - \chi(1-h_{1,t})^{-\sigma_2} h_{1,t} - \beta \chi (1-h_{2,t+1})^{-\sigma_2} h_{2,t+1}). \end{aligned} \quad (2.29)$$

The first order conditions with respect to $h_{1,t}$, $h_{2,t+1}$, K_{t+1} , $c_{1,t}$ and $c_{2,t+1}$ are

$$\rho_t = \chi(1-h_{1,t})^{-\sigma_2} \left[1 + \lambda_t \left(1 + \frac{\sigma_2 h_{1,t}}{(1-h_{1,t})} \right) \right], \quad (2.30)$$

$$\rho_{t+1} \theta = \chi(1-h_{2,t+1})^{-\sigma_2} \frac{\beta}{\epsilon_2} \left[1 + \lambda_t \left(1 + \frac{\sigma_2 h_{2,t+1}}{(1-h_{2,t+1})} \right) \right], \quad (2.31)$$

$$\rho_t = \theta(1+r)\rho_{t+1}, \quad (2.32)$$

$$\rho_t = c_{1,t}^{-\sigma_1} + \lambda_t(1-\sigma_1)c_{1,t}^{-\sigma_1}, \quad (2.33)$$

and

$$\theta \rho_{t+1} = \beta c_{2,t+1}^{-\sigma_1} + \beta \lambda_t(1-\sigma_1)c_{2,t+1}^{-\sigma_1}, \quad (2.34)$$

Combining the first order equations for the government's problem with respect to consumption (equations 2.33 and 2.34) yields

$$\left(\frac{c_{2,t+1}}{c_{1,t}}\right)^{\sigma_1} = \frac{\beta\rho_t}{\rho_{t+1}\theta}. \quad (2.35)$$

Further, combining the household's first order conditions, equations 2.5, and 2.6, under the non-constant Frisch utility specification yields

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \frac{1}{\epsilon_2} \left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\sigma_1} \left(\frac{1 - h_{2,t+1}}{1 - h_{1,t}}\right)^{-\sigma_2}. \quad (2.36)$$

Combining equation 2.36 and 2.35 gives

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \left(\frac{\beta\rho_t}{\epsilon_2\rho_{t+1}\theta}\right) \left(\frac{1 - h_{2,t+1}}{1 - h_{1,t}}\right)^{-\sigma_2}. \quad (2.37)$$

Next, I combine the first order conditions for the government with respect to young and old hours,

$$\frac{1 + \lambda_t\left(1 + \frac{\sigma_2 h_{1,t}}{1 - h_{1,t}}\right)}{1 + \lambda_{t+1}\left(1 + \frac{\sigma_2 h_{2,t+1}}{1 - h_{2,t+1}}\right)} = \frac{\beta\rho_t}{\epsilon_2\rho_{t+1}\theta} \left(\frac{1 - h_{2,t+1}}{1 - h_{1,t}}\right)^{-\sigma_2}. \quad (2.38)$$

Therefore, equation 2.37 and equation 2.38 simplify to

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \frac{1 + \lambda_t\left(1 + \frac{\sigma_2 h_{1,t}}{1 - h_{1,t}}\right)}{1 + \lambda_t\left(1 + \frac{\sigma_2 h_{2,t+1}}{1 - h_{2,t+1}}\right)}. \quad (2.39)$$

No Age Conditional Labor Income Taxes

When the government cannot condition labor income taxes on age, the equation 2.11 must be included as a constraint in the Lagrangian. The Lagrangian

for this model is,

$$\mathcal{L} = \frac{c_{1,t}^{1-\sigma_1}}{1-\sigma_1} + \chi \frac{(1-h_{1,t})^{1-\sigma_2}}{1-\sigma_2} + \beta \frac{c_{2,t+1}^{1-\sigma_1}}{1-\sigma_1} + \chi \frac{(1-h_{2,t+1})^{1-\sigma_2}}{1-\sigma_2} \quad (2.40)$$

$$\begin{aligned} & - \rho_t (c_{1,t} + c_{2,t} + K_{t+1} - K_t + G_t - rK_t - w(h_{1,t} + h_{2,t}\epsilon_2)) \\ & - \rho_{t+1} \theta (c_{1,t+1} + c_{2,t+1} + K_{t+2} - K_{t+1} + G_{t+1} - rK_{t+1} - w(h_{1,t+1} + h_{2,t+1}\epsilon_2)) \\ & + \lambda_t (c_{1,t}^{1-\sigma_1} + \beta c_{2,t+1}^{1-\sigma_1} - \chi(1-h_{1,t})^{-\sigma_2} h_{1,t} - \beta \chi(1-h_{2,t+1})^{-\sigma_2} h_{2,t+1}) \\ & \eta_t (\epsilon_2 c_{2,t+1}^{-\sigma_1} (1-h_{1,t})^{-\sigma_2} - c_{1,t}^{-\sigma_1} (1-h_{2,t+1})^{-\sigma_2}). \end{aligned} \quad (2.41)$$

The first order conditions with respect to $h_{1,t}$, $h_{2,t+1}$, K_{t+1} , $c_{1,t}$ and $c_{2,t+1}$ are

$$\rho_t = \chi(1-h_{1,t})^{-\sigma_2} \left[1 - \eta_t \epsilon_2 \sigma_2 \frac{c_{2,t+1}^{-\sigma_1}}{(1-h_{1,t})} + \lambda_t \left(1 + \frac{\sigma_2 h_{1,t}}{(1-h_{1,t})} \right) \right], \quad (2.42)$$

$$\rho_{t+1} \theta \epsilon_2 = \chi(1-h_{2,t+1})^{-\sigma_2} \beta \left[1 + \eta_t \sigma_2 \frac{c_{1,t}^{-\sigma_1}}{(1-h_{2,t+1})} + \lambda_t \left(1 + \frac{\sigma_2 h_{2,t+1}}{(1-h_{2,t+1})} \right) \right], \quad (2.43)$$

$$\rho_t = \theta(1+r)\rho_{t+1} \quad (2.44)$$

$$\rho_t = c_{1,t}^{-\sigma_1} \left[1 + \lambda_t(1-\sigma_1) + \frac{\eta_t \sigma_1 (1-h_{2,t+1})^{-\sigma_2}}{c_{1,t}} \right], \quad (2.45)$$

and

$$\theta \rho_{t+1} = \beta c_{2,t+1}^{-\sigma_1} \left[1 + \lambda_t(1-\sigma_1) - \frac{\eta_t \epsilon_2 \sigma_1 (1-h_{1,t})^{-\sigma_2}}{c_{2,t+1}} \right]. \quad (2.46)$$

Combining equations 2.44, 2.45, and 2.46 yields,

$$\left(\frac{c_{1,t}}{c_{2,t+1}} \right)^{-\sigma_1} = \beta(1+r) \left(\frac{1 + \lambda_t(1-\sigma_1) - \frac{\eta_t \epsilon_2 \sigma_1 (1-h_{1,t})^{-\sigma_2}}{c_{2,t+1}}}{1 + \lambda_t(1-\sigma_1) + \frac{\eta_t \sigma_1 (1-h_{2,t+1})^{-\sigma_2}}{c_{1,t}}} \right). \quad (2.47)$$

Constant Frisch Utility Function

The Lagrangian for this specification is

$$\begin{aligned}
\mathcal{L} = & \frac{c_{1,t}^{1-\varsigma_1}}{1-\varsigma_1} - \chi \frac{h_{1,t}^{1+\frac{1}{\varsigma_2}}}{1+\frac{1}{\varsigma_2}} + \beta \frac{c_{2,t+1}^{1-\varsigma_1}}{1-\varsigma_1} - \chi \frac{h_{2,t+1}^{1+\frac{1}{\varsigma_2}}}{1+\frac{1}{\varsigma_2}} \\
& - \rho_t (c_{1,t} + c_{2,t} + K_{t+1} - K_t + G_t - rK_t - w(h_{1,t} + h_{2,t}\epsilon_2)) \\
& - \rho_{t+1} \theta (c_{1,t+1} + c_{2,t+1} + K_{t+2} - K_{t+1} + G_{t+1} - rK_{t+1} - w(h_{1,t+1} + h_{2,t+1}\epsilon_2)) \\
& + \lambda_t (c_{1,t}^{1-\varsigma_1} + \beta c_{2,t+1}^{1-\varsigma_1} + \chi h_{1,t}^{1+\frac{1}{\varsigma_2}} + \beta \chi h_{2,t+1}^{1+\frac{1}{\varsigma_2}}).
\end{aligned} \tag{2.48}$$

The first order conditions with respect to labor, capital and consumption are

$$\rho_t = \chi h_{1,t}^{\frac{1}{\varsigma_2}} \left[1 + \lambda_t \left(1 + \frac{1}{\varsigma_2} \right) \right] \tag{2.49}$$

$$\rho_{t+1} \theta \epsilon_2 = \beta \chi h_{2,t+1}^{\frac{1}{\varsigma_2}} \left[1 + \lambda_t \left(1 + \frac{1}{\varsigma_2} \right) \right], \tag{2.50}$$

$$\rho_t = \theta (1+r) \rho_{t+1}, \tag{2.51}$$

$$\rho_t = c_{1,t}^{-\varsigma_1} + \lambda_t (1-\varsigma_1) c_{1,t}^{-\varsigma_1}, \tag{2.52}$$

and

$$\theta \rho_{t+1} = \beta c_{2,t+1}^{-\varsigma_1} + \beta \lambda_t (1-\varsigma_1) c_{2,t+1}^{-\varsigma_1}. \tag{2.53}$$

Combining the first order equations for the governments problem with consumption (equations 2.52 and 2.53) yields

$$\left(\frac{c_{2,t+1}}{c_{1,t}} \right)^{\varsigma_1} = \frac{\beta \rho_t}{\rho_{t+1} \theta}. \tag{2.54}$$

Taking the ratio of the agent's first order conditions, equations 2.5 and 2.6, under the constant Frisch utility specification gives

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \frac{1}{\epsilon_2} \left(\frac{c_{1,t}}{c_{2,t+1}} \right)^{-\varsigma_1} \left(\frac{h_{2,t+1}}{h_{1,t}} \right)^{\frac{1}{\varsigma_2}}. \tag{2.55}$$

Combining equation 2.54 and 2.55 yields

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \frac{\beta \rho_t}{\epsilon_2 \rho_{t+1} \theta} \left(\frac{h_{2,t+1}}{h_{1,t}} \right)^{\frac{1}{\varsigma_2}}. \quad (2.56)$$

The ratio of first order equations for the government with respect to young and old hours is

$$\frac{\rho_t \beta}{\epsilon_2 \rho_{t+1} \theta} \left(\frac{h_{2,t+1}}{h_{1,t}} \right)^{\frac{1}{\varsigma_2}} = \frac{1 + \lambda_t \left(1 + \frac{1}{\varsigma_2}\right)}{1 + \lambda_t \left(1 + \frac{1}{\varsigma_2}\right)}. \quad (2.57)$$

Combining equation 2.57 and 2.56 generates the following expression for the optimal labor taxes

$$\frac{1 - \tau_{h,2}}{1 - \tau_{h,1}} = \frac{1 + \lambda_t \left(1 + \frac{1}{\varsigma_2}\right)}{1 + \lambda_t \left(1 + \frac{1}{\varsigma_2}\right)} = 1. \quad (2.58)$$

No Government Savings or Borrowing

The inclusion of the no government borrowing or savings restriction means that equation 2.13 must be included as an additional constraint in the Lagrangian. The Lagrange is,

$$\begin{aligned} \mathcal{L} = & \frac{c_{1,t}^{1-\sigma_1}}{1-\sigma_1} + \chi \frac{(1-h_{1,t})^{1-\sigma_2}}{1-\sigma_2} + \beta \frac{c_{2,t+1}^{1-\sigma_1}}{1-\sigma_1} + \chi \frac{(1-h_{2,t+1})^{1-\sigma_2}}{1-\sigma_2} \\ & - \rho_t (c_{1,t} + c_{2,t} + K_{t+1} - K_t + G_t - rK_t - w(h_{1,t} + h_{2,t}\epsilon_2)) \\ & - \rho_{t+1} \theta (c_{1,t+1} + c_{2,t+1} + K_{t+2} - K_{t+1} + G_{t+1} - rK_{t+1} - w(h_{1,t+1} + h_{2,t+1}\epsilon_2)) \\ & + \lambda_t (c_{1,t}^{1-\sigma_1} + \beta c_{2,t+1}^{1-\sigma_1} - \chi (1-h_{1,t})^{-\sigma_2} h_{1,t} - \beta \chi (1-h_{2,t+1})^{-\sigma_2} h_{2,t+1}) \\ & \varphi_t (K_{t+1} - c_{1,t} - \chi c_{1,t}^{\sigma_1} (1-h_{1,t})^{-\sigma_2}). \end{aligned} \quad (2.59)$$

The first order conditions with respect to K_{t+1} , $c_{1,t}$ and $c_{2,t+1}$ are

$$\rho_t = \theta(1+r)\rho_{t+1}, \quad (2.60)$$

$$\rho_t = c_{1,t}^{-\sigma_1} \left[1 + \lambda_t(1-\sigma_1) - \varphi_t \left(\frac{1}{c_{1,t}^{-\sigma_1}} + \frac{\chi \sigma_1 (1-h_{1,t})^{-\sigma_2}}{c_{1,t}} \right) \right], \quad (2.61)$$

and

$$\theta \rho_{t+1} = \beta c_{2,t+1}^{-\sigma_1} \left[1 + \lambda_t(1-\sigma_1) \right].. \quad (2.62)$$

Combining equations 2.60, 2.61, and 2.62 yields

$$\left(\frac{c_{1,t}}{c_{2,t+1}}\right)^{-\sigma_1} = \beta(1+r) \left(\frac{1 + \lambda_{t+1}(1 - \sigma_1)}{1 + \lambda_t(1 - \sigma_1) - \varphi_t\left(\frac{1}{c_{1,t}^{-\sigma_1}} + \frac{\chi\sigma_1(1-h_{1,t})^{-\sigma_2}}{c_{1,t}}\right)}\right). \quad (2.63)$$

Chapter 3

Measuring the Frisch Labor Supply Elasticity: Evidence from a Pseudo Panel

Abstract

The level of the Frisch labor supply elasticity, as well as the profile over the lifetime are essential for welfare analysis of many policy changes in general equilibrium models. This paper uses a pseudo panel to examine two open questions with regards to the Frisch elasticity. First, can the difference between the macroeconomic calibration values and microeconomic estimates of the Frisch elasticity be explained by the extensive margin? Second, what is the shape of the Frisch labor supply elasticity profile? With regard to the first question, including only the intensive margin, I find estimates of the Frisch elasticity between 0.6-0.64. Including both the intensive and extensive margins, the elasticity estimates increase to 1.86-2.11. The difference between the two ranges indicates that the impact of the extensive margin is large enough to explain the difference between previous micro estimates and macro calibration values. With regard to the second question, the Frisch elasticity profile including just the intensive margin is flat over the working life. Including both the intensive and extensive margin, the Frisch elasticity profile is flat until it increases a statistically significant amount at the age of 55.

3.1 Introduction

Both the level and the shape over the lifetime of the Frisch labor supply elasticity are important for conducting Macroeconomic and Public Finance analysis in general equilibrium models. The average Frisch elasticity is essential in order to calibrate macro models such that the volatility of hours over the business cycle matches the economy. Additionally, the shape of the Frisch labor supply elasticity profile is important in determining the welfare impact of policy changes and determining optimal fiscal policies. For example, chapter 2, and Conesa et al. (2009) demonstrate that an upward sloping Frisch labor supply elasticity profile in an overlapping generation model is a strong motive for a positive tax on capital. This paper uses a pseudo panel of men in order to estimate the parameter value that identifies the Frisch elasticity for a representative agent model and demonstrates how this value changes over the life cycle.

The original microeconomic estimates of the Frisch elasticity are determined from changes in an individual's labor supply over his life cycle. These studies, including MaCurdy (1981) and Altonji (1986), find that the Frisch elasticity is between 0 - 0.54 and that the values are generally not statistically significant.¹ Subsequent works, such as Domeij and Flodén (2006), Imai and Keane (2004), Pistaferri (2003), Contreras and Sinclair (2008), and Chetty (2009), demonstrate that these original estimates may be biased downwards because they do not account for borrowing constraints, they disregard the life cycle impact of endogenous-age specific human capital, they omit correlated variables such as wage uncertainty, they use instruments that are weak, and they do not account for labor market frictions.

In contrast to the small microeconomic estimates, the values typically used in macro models are much larger. These values are calibrated such that the model's fluctuations in aggregate hours over the business cycle match the data. These macro calibration values of the Frisch elasticity are usually set between 2 - 4.² A widely accepted explanation for the difference in the microeconomic

¹See Mulligan (1995) for a survey of the original estimate.

²See Chetty et al. (2011) for a discussion of the values used.

estimates and the macro calibration values of the Frisch elasticity is that the microeconomic estimates exclude decisions on the extensive margin (see Chetty et al. (2011) for a detailed discussion).³

Chetty et al. (2011) examine whether the inclusion of the extensive margin can explain the larger macro calibration values compared to the microeconomic estimates. The authors use a meta analysis of separate quasi-experimental studies to determine the intensive and extensive parts of the macro elasticity. When calculating the size of the extensive margin, the authors focus on estimates of the participation rate elasticity. They find that the estimates of the participation rate elasticity are not as large as the difference between the microeconomic estimates and the macro calibration values. Therefore, they conclude that the inclusion of the extensive margin cannot explain the larger macro elasticity values.

I take an alternative approach to determine whether the extensive margin can explain the difference between the microeconomic estimates and the macro calibration values. Using a pseudo panel, I estimate the micro elasticity as the percentage change in the cohort's average hours worked on the intensive margin given a one percent change in the wage. Additionally, I estimate the macro elasticity as the percentage change in the cohort's average hours worked from both the intensive and extensive margin given a one percent change in the wage. Using a pseudo panel allows me to estimate the macro elasticity directly from the data as opposed to separately estimating the impacts of the intensive and extensive margin. Thus, instead of relying on the participation elasticity, I include the extensive margin's impact on the cohort's average hours worked in order to estimate the macro elasticity.⁴ I estimate that the micro Frisch elasticity is between 0.6 - 0.64 in my preferred specifications. This is on the upper end of the original microeconomic estimates.⁵ I estimate that the macro Frisch elasticity is between

³In order for the extensive margin to cause the difference individuals need to face some indivisibility of labor.

⁴Appendix 3.6.1 demonstrates that in order for the impact of the participation rate elasticity to be consistent with the impact of the extensive margin elasticity calculated in the pseudo panel, then individuals on the margin of working must work, on average, the same number hours as the rest of the workers.

⁵My estimates are different from the microeconomic estimates because I utilize a pseudo panel which implies that my estimates are fundamentally different than the estimate from a

1.86 - 2.11.⁶ These values are in line with the values typically used in calibrated macro models. The micro and macro estimates in this study demonstrate that the difference between the microeconomic estimates and the macro calibration values can plausibly be explained by the inclusion or exclusion of the extensive margin.

In addition to estimating the average Frisch elasticity, I also identify how the Frisch elasticity changes over the lifetime. I find that the micro Frisch elasticity is generally stable for the whole working life with some evidence that it is higher early in the working life. I find that the macro Frisch elasticity is stable from the ages of 23 to 54 and increases in the last eight years of the working life.

This paper is organized as follows: Section 2 derives the estimation equations from a simple labor supply model. Section 3 describes the data and discusses how I construct the pseudo panel. Section 4 presents the estimates of the average Frisch elasticity and the Frisch elasticity profile on both the micro and macro level. Section 5 concludes.

3.2 Labor Supply Model

In this section, I introduce the typical maximization problem for an individual. Next, I solve for the first order conditions. Manipulating these first order conditions, I create two different specifications that can be used to estimate the Frisch elasticity in a reduced form setting. I finish this section by demonstrating why estimates based from these specifications may be susceptible to omitted variable bias.

traditional panel data. The pseudo panel identifies the Frisch based on the change over the average life cycle and is not affected by idiosyncratic shocks. These idiosyncratic shocks have a similar effect to measurement error and cause the estimate to be lower.

⁶Mulligan (1995) does a similar analysis in a traditional panel however he does not attempt to control for the selection bias of individuals on the extensive margin

3.2.1 Individual's Decisions

Employing a typical utility function that is homothetic and separable in consumption and labor, an individual at age s solves the following problem,

$$\max E_s \sum_{j=s}^J \beta^{j-1} \left(\chi_{i,j}^c \frac{\mu}{1+\mu} c_{i,j}^{1+\frac{1}{\mu}} - \chi_{i,j}^h \frac{\gamma}{1+\gamma} h_{i,j}^{1+\frac{1}{\gamma}} \right) \quad (3.1)$$

subject to,

$$c_{i,j} + a_{i,j+1} = w_{i,j} h_{i,j} + (1+r_t) a_{i,j} \quad (3.2)$$

where E_s represents the expectation operator at age s , J is the age of death, $c_{i,j}$ is consumption of individual i at age j , h is hours worked, χ^c is a parameter that controls taste for consumption, χ^h is a parameter that controls tastes for work, β is the discount rate, a_j is savings, and r_t is the after tax return to savings. The first order conditions for the individual are,

$$\lambda_j = \chi_{i,j}^c c_{i,j}^{\frac{1}{\mu}} \quad (3.3)$$

$$\lambda_j w_j = \chi_{i,j}^h h_{i,j}^{\frac{1}{\gamma}} \quad (3.4)$$

$$\lambda_j = E_j \beta \Psi_{j,j+1} (1+r) \lambda_{j+1}^7 \quad (3.5)$$

where λ is the marginal utility of consumption. The parameter of interest, γ , is the Frisch labor supply elasticity. I derive two different specifications which I use to determine γ . By taking the logs and combining equations 3.3 and 3.4, I derive the first specification which relates hours to consumption, tastes, and wages,

$$\ln h_{i,j} = \gamma \left[\frac{1}{\mu} \ln c_{i,j} + \ln \chi_{i,j}^c - \ln \chi_{i,j}^h + \ln w_{i,j} \right] \quad (3.6)$$

⁷This is the intertemporal euler equation for an individual at the age j . If the individual is solving at a different age then the expectation operator should adjusted accordingly.

I derive the second specification by taking difference between two ages of the log of equation 3.4,

$$\Delta \ln h_{i,j+1} = \gamma[\Delta \ln \lambda_{i,j+1} + \Delta \ln w_{i,j+1} - \Delta \ln \chi_{i,j+1}^h] \quad (3.7)$$

where Δ represents the change over one year. Defining $\xi_{i,j+1} \equiv \lambda_{i,j+1} - E\lambda_{i,j+1}$, and combining equations 3.5 and 3.7 the second specification used to determine γ can be written as,

$$\Delta \ln h_{i,j+1} = \gamma[-\ln \beta - \ln(1 + r_t) + \xi_{i,j+1} + \Delta \ln w_{i,j+1} - \Delta \ln \chi_{i,j+1}^h] \quad (3.8)$$

This second specification relates the change in hours to the change in wages and preference parameters. I refer to equation 3.6 as the level specification and equation 3.8 as the change specification. Both specifications have been used in the past to estimate the Frisch elasticity.

Typically reduced forms of equations 3.6 and 3.8 are estimated to determine the Frisch elasticity. When estimating the Frisch elasticity, there is a general concern about omitted variables. In the level specification, marginal utility and taste preferences are typically unobserved and could be correlated with wages. In the difference specification, the unexpected changes in marginal utility and the changes in taste preferences are unobserved and could be correlated with wages. When estimating these equations, it is important to either use instruments for wages or control for these unobserved variables. Section 3.4.1 describes how I account for these omitted variables.

3.3 Data

I estimate the Frisch elasticity from a pseudo panel using reduced forms of equations 3.6 and 3.8. When estimating parameter values for macroeconomic calibration, it is natural to identify the parameter values from aggregate movements in the representative cohort's value as opposed to the individual's idiosyncratic change. Ideally, one would use the population averages by age and time as these

representative values. In the absence of these values, Deaton (1985) proposes creating a pseudo panel from repeated cross sectional data. The pseudo panel is created by taking the average values of the variables of interest by age and time from a repeated cross section. These time series of synthetic cohort averages are treated as approximations of the whole cohort's averages. Generally, treating a pseudo panel as genuine panel can cause a bias equivalent to measurement error since the economist only observes the average from the sample of the cohort and not the average from the whole cohort. However, Verbeek et al. (1992) demonstrates that with a sufficient number of individuals a pseudo panel can be treated as a genuine panel without introducing a significant bias.⁸

I create a pseudo panel by joining two different data sets. I use both the Consumption Expenditure Survey (CEX) and the March Current Population Survey (CPS). The CPS is an annual survey that contains data on an individual's characteristics, income, and hours worked.⁹ The CEX provides information on households expenditures, income, and characteristics.¹⁰ I rely on the CPS for hours, wage, and individual characteristics. I utilize the CEX for information on consumption.

I use a procedure similar to Pencavel (2002) in order to create a pseudo panel. I limit the sample to males neither in the armed forces nor self-employed and who are between the age of 22 and 62. In order to isolate the effect of the extensive margin, I create two panels, one that only includes individuals who are working and one that includes all individuals. Individuals who are not actively working, students, or retirees are considered not working. Additionally, if individ-

⁸In order to estimate equations 3.6 and 3.8 I use the natural log of the average value of the cohort as opposed to the using the average natural log of the cohort's value. Using the natural log of the average value corresponds to find the parameter value that corresponds to the representative cohort.

⁹The survey is done on a monthly basis but the more detailed March version of the survey is done on an annual basis. Data comes from Miriam King, Steven Ruggles, J. Trent Alexander, Sarah Flood, Katie Genadek, Matthew B. Schroeder, Brandon Trampe, and Rebecca Vick. Integrated Public Use Microdata Series, Current Population Survey: Version 3.0. [Machine-readable database]. Minneapolis: University of Minnesota, 2010.

¹⁰The CEX reports expenditures for a family and not for individuals. Expenditures are attributed to the head of the household, and in order to control for family size, expenditures are divided by family size. I use the NBER extraction initiated by Ed Harris and John Sabelhaus.

uals work less than 300 hours a year, they are considered to not be working.¹¹ I create an individual hourly wage from the CPS by dividing the sum of wage and salary income by the product of average hours worked a week and weeks worked in the year. Similar to French (2005), wages are dropped if an individual's hourly wage is less than 3 dollars or more than 1,000 dollars.¹² Past studies that have estimated the Frisch using the level specification typically use the Panel Study of Income Dynamic (PSID) (see Altonji (1986)). Since the PSID only has data on an individual's expenditures on food, these studies assume that expenditures on food are a proxy for total consumption and use food expenditures to control for marginal utility. In order to determine if this assumption affects the estimates of the Frisch elasticity, I estimate the level specification using both total expenditures and expenditures on food.

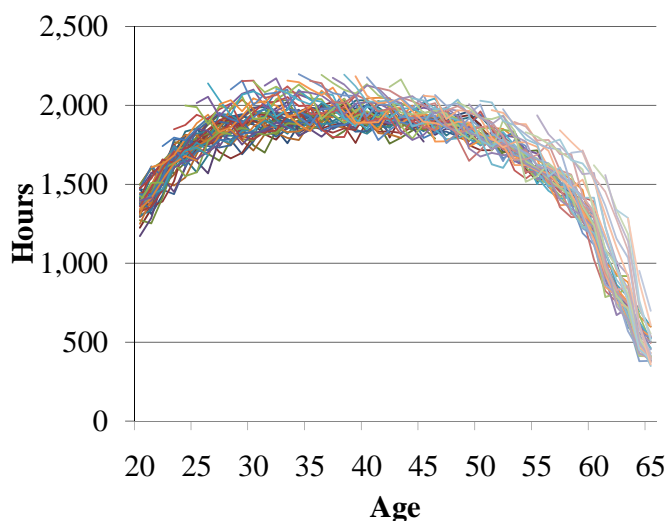


Figure 3.1: Hours (All)

Figures, 3.1, 3.2, 3.3, 3.4, and 3.5 plot the hours worked for all individuals, hours worked on the subset of the population who spend time working, wages for working individuals, total expenditures, expenditures on food for each cohort

¹¹I found that the estimates tended to not be sensitive to this restriction. However, many individuals who reported working less than 300 hours had an imputed wage that was unrealistically high. Therefore, I exclude them from the panel.

¹²All values are deflated to real dollars (2000) using the CPI.

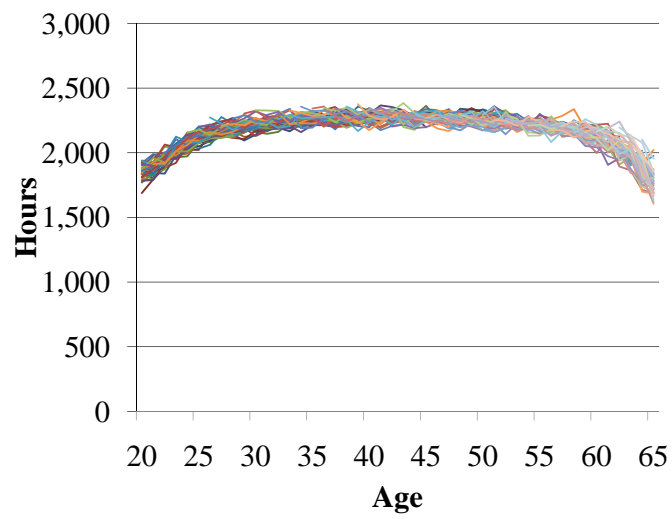


Figure 3.2: Hours (Workers)

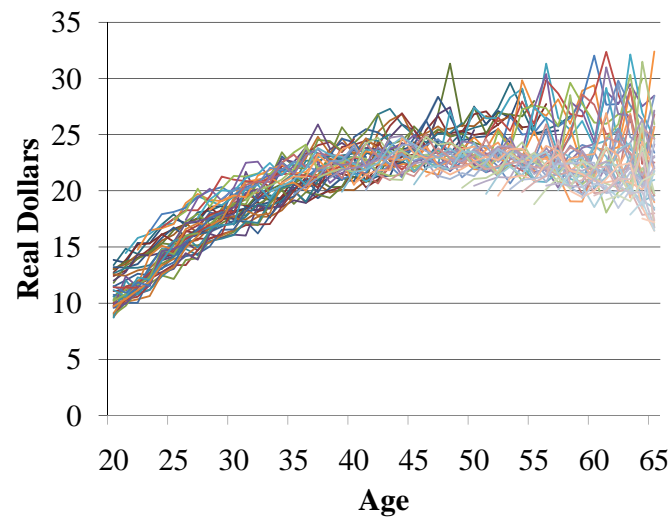


Figure 3.3: Wages (Workers)

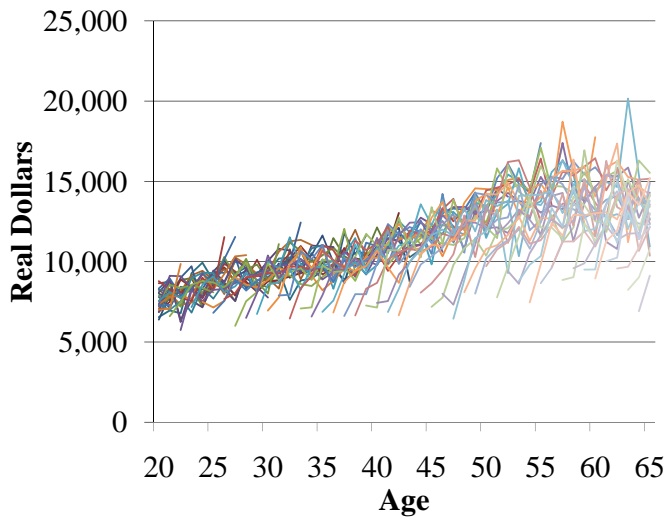


Figure 3.4: All Consumption

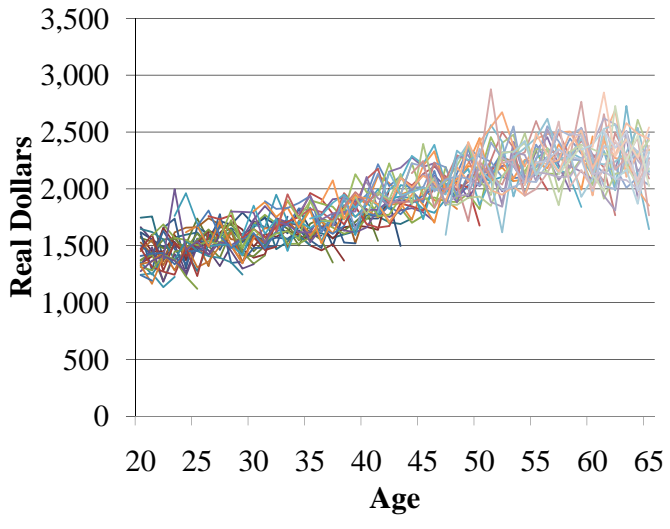


Figure 3.5: Food Consumption

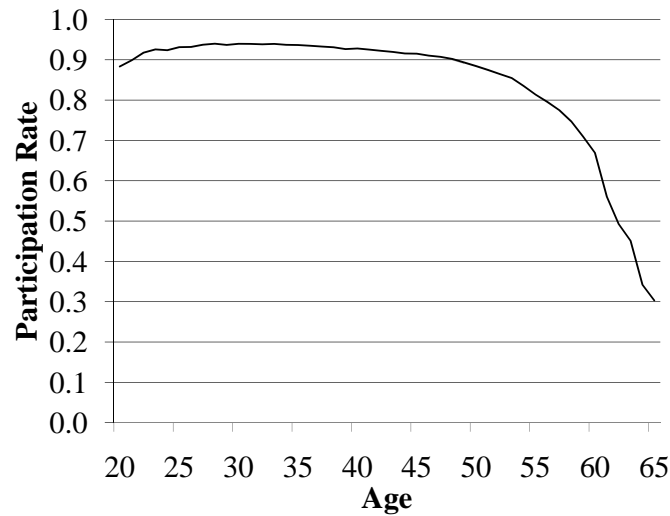


Figure 3.6: Participation Rate

by age, respectively.¹³ Figure 3.6 plots the average participation rate across the whole cohorts by age. Figures 3.1 and 3.2 plot the cohort's value of hours worked. Figure 3.2 depicts the value for the subset of the cohort that is working (the intensive margin), while figure 3.1 is the value for the whole cohort (intensive and extensive margin). The general shapes of the profiles in figures 3.1 and 3.2 look similar between the ages of 20 and 50. Starting around 50, the profiles in figure 3.1 decrease more quickly than the profile in figure 3.2. Focusing on figure 3.6, the average participation rate is steady across the ages 20 - 50, however it starts to fall rapidly after 50. The decrease in the participation rate is responsible for the different shapes in the hours profiles starting around 50.

Comparing figures 3.4 and 3.5, the general upward sloping shape of the total expenditures and food expenditures are similar. Although the general shape is the same, total consumption tends to be more volatile.

¹³The figures represent the average across all individuals in the cohort

3.4 Results

In this section I estimate both the aggregate Frisch elasticity and also the Frisch elasticity profile. First, I replicate previous studies' estimation of the aggregate micro Frisch elasticity. Second, I use my own specification to estimate the aggregate micro and macro Frisch elasticity. Finally, I end by estimating both the micro and macro Frisch elasticity profiles.

3.4.1 Aggregate Frisch

Replication

I begin by estimating the aggregate Frisch using the same specification as MaCurdy (1981), Altonji (1986), and Pencavel (2002). These microeconomic estimates are determined from data on working individuals and therefore only estimate the Frisch on the intensive margin. The first two studies estimate the Frisch elasticity using a traditional panel while Pencavel (2002) uses a pseudo panel. I replicate the author's different estimations in a consistent data set, my pseudo panel, in order to isolate the impact of the different specifications.

The previous studies estimate the Frisch elasticity using reduced form equations based on equations 3.6 and 3.8. The reduced form estimation equations are,

$$\ln h_{n,j} = \gamma \ln w_{n,j} + \beta \ln c_{n,j} + \zeta \text{TS}_{n,j} + e_{n,j} \quad (3.9)$$

$$\Delta \ln h_{n,j+1} = \gamma \Delta \ln w_{i,j+1} + \delta_{n+j} + \zeta \Delta \text{TS}_{n,j} + \epsilon_{n,j} \quad (3.10)$$

where $h_{n,j}$ is hours worked for the cohort born in year n at the age j , TS is a vector of taste shifters, and δ_{n+j} is a set of annual dummies.¹⁴ When estimating the Frisch elasticity, these previous studies not only differed in the estimation equations they used but they also included a different range of ages in their covered population and used different instruments in order to control for either measurement error,

¹⁴ δ_{n+j} is included in the change specification to control for annual changes in the after tax return to capital. MaCurdy (1981) and Altonji (1986) estimated the Frisch elasticity in a traditional panel. Therefore, in their estimation specification n would represent an individual not a cohort.

omitted variables, or both. When available, I replicate the previous studies using the authors' preferred specification, set of instruments, and set of controls in the pseudo panel.

Table 3.1 presents the estimates from the replication exercise. The first column replicates MaCurdy (1981), the second and third column replicate Altonji (1986), and the last column replicates Pencavel (2002). MaCurdy (1981) estimates equation 3.10, using polynomials of age and education as instruments for wages in order to account for the omitted variable of unexpected shocks to marginal utility. Altonji (1986) estimates both equation 3.9 and 3.10. Column I and II, are both estimates using the change specification. However there are three differences in the estimation strategies. First, column II (Altonji (1986)) includes older individuals in his population. Second, column II includes controls for taste shifters. Third, column II includes a larger set of instruments.¹⁵ These three differences cause the estimated Frisch to be higher in the Altonji change specification than the MaCurdy specification.

The Altonji level specification, column III, is based on equation 3.9. Comparing column II and column III, the difference is that column II uses the change specification and column III uses the level specification. Table 3.1 demonstrates that the estimates of the Frisch elasticity tends to be around half as large using the level specification compared to the estimates with the change specification. The order of magnitude of the decrease is consistent with the decrease observed in Altonji (1986) using a traditional panel.

Column IV, replicates the specification in Pencavel (2002). Different from MaCurdy (1981) and Altonji (1986), Pencavel (2002) uses a pseudo panel. Pencavel (2002) notes that he is mostly concerned with measurement error both on

¹⁵The controls that the Altonji change specification includes are number of children under the age of five, an indicator variable for children under the age of five, an indicator variable for whether the individual is married, an indicator variable for whether the individual lives with his spouse, an indicator variable for whether the individual lives in an urban area, and an indicator variable for whether the individual is a minority. In a pseudo panel the indicator variables are the average over all the individuals in a cohort so the cohort's value is not binary. When replicating Altonji I include year dummies even though they are not included in equation 3.9 since Altonji (1986) reports including them. I found that there was not an economically significant difference in the results when I did not include these dummies.

Table 3.1: Aggregate Micro Frisch Elasticity (Replications)

	I	II	III	IV
Author:	MaCurdy	Altonji	Altonji	Pencavel
Specification:	Change	Change	Level	Change
Age:	22-55	25-60	25-60	25-64
Original Estimates:	.1 - .45	.1 - .54	.09 - .17	0.15
lnwage			0.304 (0.0220)	
lnfood			-0.317 (0.0318)	
chlnwage	0.486 (0.0261)	0.692 (0.116)		0.307 (0.312)
constant	-0.0117 (0.00584)	0.0536 (0.0370)	9.054 (0.227)	-0.00613 (0.00384)
Summary Statistics For 1st Stage				
F-stat for excluded instruments:	165.7	19.29	527.3	1.513
Shea Partial R ² in 1st stage:	0.0812	0.0196	0.778	0.00235
Hansen J Stat for Overidentification of Instruments (P-Value):	0.169	0.277	0.0175	0.743
<u>Instruments:</u>				
age	x	x	x	
age2			x	
age3				
educ	x	x	x	
educ2		x	x	
age educ	x	x	x	
Past wage				
Pencavel				x
<u>Controls:</u>				
year	x	x	x	
children5		x	x	
children5d				
married				
marriedlive				
metro		x	x	
raced		x	x	

Notes: The F-stat for excluded instruments is for the wage regressions. The Shea Partial R² measures the relevance of the instruments taking into account intercorrelations among the instruments. Consistent with previous studies, the standard errors are clustered on cohort for column I, II and III.

the individual level and also measurement error created from using a pseudo panel. Therefore, instead of using instruments that are uncorrelated with unexpected shocks to marginal utility, he uses instruments that are aggregate measures correlated with labor demand. Specifically, he uses U.S. real trade balances on current accounts and the level of real merchandise imports. The estimates from replicating this specification (column IV) are lower compared to the estimates from the other change specifications (column I and II). The estimates in column IV are lower because not including controls for taste shifters nor using instruments that account for the correlation between unexpected shocks to marginal utility and wages cause a downward bias.

Overall, I find that when I compare the estimates of the Frisch elasticity from the pseudo panel to the previous estimates from a traditional panel (see original estimate row), the point estimates from the pseudo panel tend to be larger. Since the estimates from the pseudo panel are identified from the changes in the cohort's average over the life cycle as opposed to the idiosyncratic changes for each individual one would expect the estimates to be larger.¹⁶ Since this paper is focusing on estimating the parameter value for calibrating a general equilibrium model with a representative individual, using a pseudo panel is a natural framework.

New Estimates

In this section I estimate both the micro and macro Frisch elasticity using the pseudo panel. When estimating the macro Frisch elasticity, I observe an hours decision for everyone but I only observe a wage for individuals that work. As a first pass, when I estimate the macro Frisch elasticity, I calculate the cohort's wage as the average wage over only working individuals. One concern is that I may be introducing a selection bias into the estimates by only include the wages from the subset of the cohort that works since an individual's decision to work may be affected by his potential wage. Therefore, in addition to estimating the macro elasticity using the observed wages, I estimate the macro elasticity with two

¹⁶The changes to wages that come from idiosyncratic shocks act as noise causing the estimates to be smaller.

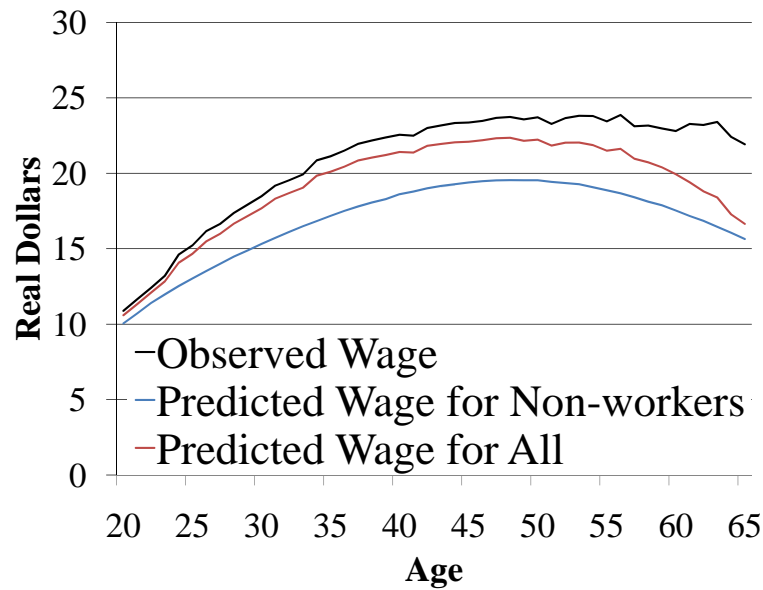


Figure 3.7: Average Wage by age for All Individuals

different wage series that attempt to account for selection bias. First, I include a predicted wage for the non-working individuals in the cohort's average. On an individual level, I regress the observed wages for workers on education, education squared, age, age squared, and the interaction of age and education. Using the coefficients from the individual level regression, I predict the potential wage for non-workers. The prediction is done on an individual level and the cohort's wage is calculated as the average over both the actual wages for workers and predicted wages for non-workers. As a second series, I use the coefficients from the individual level regression to predict an individual level wage for everyone (even for those for whom I observe a wage).

Figure 3.7 plots the average wage by age, across all cohorts, from the three different wage series. The average wage is lower for the two series that attempt to control for selection. The difference in these series indicates that the predicted wages are lower for non-working individuals compared to working individuals. Additionally, since a larger portion of the population chooses not to work later in life, the difference between those working and whole cohort's average wage is larger later in life.

Tables 3.2 presents the estimates of the micro aggregate Frisch elasticity, respectively. Columns I and II of table 3.2 are estimates using the change specification and columns III, IV, and V are estimates using the level specification. Columns III and IV use total consumption to control for marginal utility, while column V uses expenditures on food to control for marginal utility. The estimates in column II, IV and V include the controls for shifts in tastes: race, urban population, children under five, married, and living with spouse.

The Frisch elasticity estimates using the change specification are between .6 - .65 (see columns I and II). The difference between the two columns is that column II controls for shifts in tastes while column I does not include these controls. I include controls for tastes because I am concerned that my instruments are correlated with tastes for work and consumption. I find that including these controls changes the estimates. There is some concern that because all of the specifications fail the Hansen J-test for overidentification that I have not adequately controlled for these changes in tastes. Column III and IV are the estimates of the Frisch elasticity in the level specification using total expenditures to control for marginal utility. The range of these estimates is .25 - .64. Once again, the difference between these two columns is that column IV includes controls for shifts in tastes. The range between the estimates that do and do not include controls for shifts in tastes is much larger in the level specification compared to the change specification because in the level specification both the level and change in tastes over the life cycle impacts the results. In contrast, in the change specification, only changes in tastes over the life cycle impact the results. The large difference between the Frisch elasticity estimates in column IV and V indicates that there is an economically significant impact to using total expenditures versus expenditures on food to control for marginal utility. Some previous studies, such as Altonji (1986), use expenditures on food to control for marginal utility because of data limitations. However, expenditures on food are far less volatile than total expenditures (see figures 3.5 and 3.4). For the rest of the aggregate analysis, I focus on estimates that include controls for shifts in tastes and use total expenditures as opposed to expenditures on food to control for marginal utility (column II and IV).

Table 3.2: Aggregate Micro Frisch Elasticity

Specification: Taste Shifters	I Change No	II Change Yes	III Level No	IV Level Yes	V Level Yes
chlnwage	0.649 (0.0301)	0.596 (0.0558)			
lnwage			0.250 (0.0244)	0.642 (0.0357)	0.511 (0.0388)
lnconsumption			-0.160 (0.0226)	-0.135 (0.0188)	
lnfood					-0.365 (0.0534)
constant	-0.00676 (0.00804)	-0.00596 (0.00760)	8.334 (0.180)	7.475 (0.152)	9.188 (0.394)
Summary Statistics For 1st Stage					
F-stat for excluded					
instruments:	280.3	107.0	611.8	170.3	170.3
Shea Partial R ²					
in 1st stage:	0.117	0.0840	0.603	0.445	0.519
Hansen J-Stat for					
Overidentification of					
Instruments (P-Value):	8.34e-06	1.65e-05	3.07e-09	0.000674	0.00513

Notes: In the first stage of the change (level) specification education, education squared, age, age squared, age cubed and the age and education interaction are used as instruments for wage (and consumption). The F-stat for excluded instruments is for the wage regressions. The Shea Partial R² measures the relevance of the instruments taking into account intercorrelations among the instruments. The standard errors are clustered on cohort.

The estimates of the micro Frisch elasticity using these preferred specifications are between .6 - .64.

Table 3.3 presents the macro estimates of the aggregate Frisch elasticity. The main difference between table 3.2 and table 3.3 is that the macro Frisch elasticity (table 3.3) is estimated from cohort averages that include all individuals as opposed to the micro Frisch elasticity (table 3.2) which is estimated from cohort averages that only use individuals who work. The estimates of the macro Frisch elasticity range from 1.26 - 2.11. Column I, III, and V are the estimates using the change specification. Column II, IV, and VI are the estimates using the level specification. Column I and II are the estimates that do not account for selection into the workforce. Column III and IV control for the selection bias by using cohort averages that include predicted wages for individuals who do not work. Column V and VI control for selection bias by using cohort averages of predicted wages for all individuals. In columns II - VI, when I control for selection, the range narrows to 1.86 - 2.11. I prefer to rely on the specifications that control for selection bias.

Examining figures 3.6 and 3.7, one can get a sense of why there is a downward bias introduced by not controlling for selection into the labor force. Figure 3.7 demonstrates that the predicted average wage for individuals who do not work is lower than the predicted average wage for individuals that do work. Figure 3.6 indicates that the majority of the individuals who choose not to work are over fifty. Therefore, including predicted wages will cause a steeper downward slope to the wage profile after the age of fifty. Additionally, aggregate hours are generally decreasing during these later years. Therefore, including predicted wages for non-working individuals will tend to increase the correlation between wages and hours. Comparing column I to columns III and V (column II to columns IV and VI), the estimates for the Frisch elasticity tend to be lower when I do not control for selection in the level (change) specification. However, the effect of controlling for selection is much smaller in the change specification than in the level specification.

Overall, in my preferred specifications, I find that the micro elasticities are between .6 - .64 and the macro elasticities are between 1.86 - 2.11. The extensive

Table 3.3: Aggregate Macro Frisch Elasticity

	I	II	III	IV	V	VI
Spec.:	Change	Level	Change	Level	Change	Level
Selection						
Control:	None	None	Predict Missing	Predict Missing	Predict All	Predict All
lnwage		2.044 (0.266)		2.107 (0.163)		2.112 (0.176)
lncon.		-1.383 (0.125)		-1.119 (0.0829)		-1.063 (0.0885)
chlnwage	1.256 (0.133)		1.863 (0.131)		1.948 (0.0771)	
constant	-0.0362 (0.0136)	16.31 (0.899)	-0.0379 (0.0157)	13.81 (0.652)	-0.0587 (0.00758)	12.13 (0.716)
Summary Statistics for 1st Stage						
F-stat	41.22	126.0	147.4	277.5	1379	21791
R ²	0.0586	0.283	0.153	0.612	0.731	0.921
J-stat	1.09e-09	3.76e-06	9.20e-11	1.19e-06	0	2.22e-07

Notes: In the first stage of the change (level) specification education, education squared, age, age squared, age cubed and the age and education interaction are used as instruments for wage (and consumption). The F-stat is the test of the null hypothesis that the excluded instruments are irrelevant. The test is for the wage regressions. The R² statistic is the Shea partial r-squared. The Shea Partial R² measures the relevance of the instruments taking into account intercorrelations among the instruments. The J-stat is the Hansen J-stat for overidentification of instruments. The standard errors are clustered on cohort.

margin is responsible for the difference between these two ranges. The magnitude of the difference between the two ranges indicates that the impact of the extensive margin is large enough to explain the difference between the previous microeconomic estimates of the Frisch elasticity and the calibrated values of the Frisch elasticity used in macro models.¹⁷

3.4.2 Frisch Elasticity Profile

Specification

In order to estimate the Frisch labor supply elasticity profile, I change equations 3.9 and 3.10 by interacting wages with a set of dummy variable for five age bins (23 - 30, 31 - 38, 39 - 46, 47 - 54, 55 - 62). This specification implies that I am estimating a separate γ value for each age bin.¹⁸ Consistent with my preferred specifications, I estimate the profiles using the controls for taste shifters and using total consumption to control for marginal utility in the level specification.

One concern with this specification is that it requires strong parametric assumptions. Implicitly, I am assuming that age affect hours worked through just two specific channels. First, since I use age as an instrument, age affects hours through wages only in the specified parametric form. Second, age affects hours worked through an age dependent Frisch elasticity parameter. Since these two assumptions impose strong parametric assumptions, the profile results should be interpreted with caution.

¹⁷If females tend to provide labor more elastically on the extensive margin then including females in the macro panels would increase the macro Frisch elasticity estimates. Therefore, if the previous values used in macro models are calibrated to fluctuations in total hours from both males and females and micro values are estimates just from males then these estimates are a lower bound of the impact of the extensive margin.

¹⁸In order to increase the power, I do not interact wage with the dummy variables in the first stage regression. Instead I regress wages on the instruments and use the same coefficients to predict wages for all ages. I then interact the predicted wages with the set of dummy variables in the second stage in order to allow the γ coefficient to vary by age.

Profiles

Table 3.4 and figure 3.8 presents the two estimates of the micro Frisch elasticity profiles using the level and the change specifications. Generally the profile is flat for both specifications. None of the points are statistically different from any other point. However, focusing just on the point estimates, the Frisch elasticity is economically higher for young individuals in the level specification.

Table 3.4: Micro Frisch Elasticity Profile

Specification:	I Change	II Level
23-30	0.598 (0.25)	0.814 (0.34)
31-38	0.535 (0.38)	0.48 (0.44)
39-46	0.582 (0.5)	0.402 (0.6)
47-54	0.51 (0.51)	0.45 (0.47)
55-62	0.506 (0.69)	0.567 (0.42)

Notes: In the first stage of the change (level) specification education, education squared, age, age squared, age cubed and the interaction between age and education are used as instruments for wage (and consumption). The standard errors are clustered on cohort.

Table 3.5 and figure 3.9 presents the six estimates of the macro Frisch elasticity profiles. I estimate the profiles under the change and level specification, without a control for selection bias and with the two controls for selection bias. I find that generally the Frisch elasticity profiles are flat until the end of the working life. I find that five of the six estimates of the profile show a strong upward slope to the profile in the last age bin.¹⁹ The difference between the rest of the age bins

¹⁹The only profile that does not show a strong increase is the change specification that does not control for selection bias; this is the same specification for which the aggregate Frisch estimate is economically different from the other estimates (see column I of table 3.3).

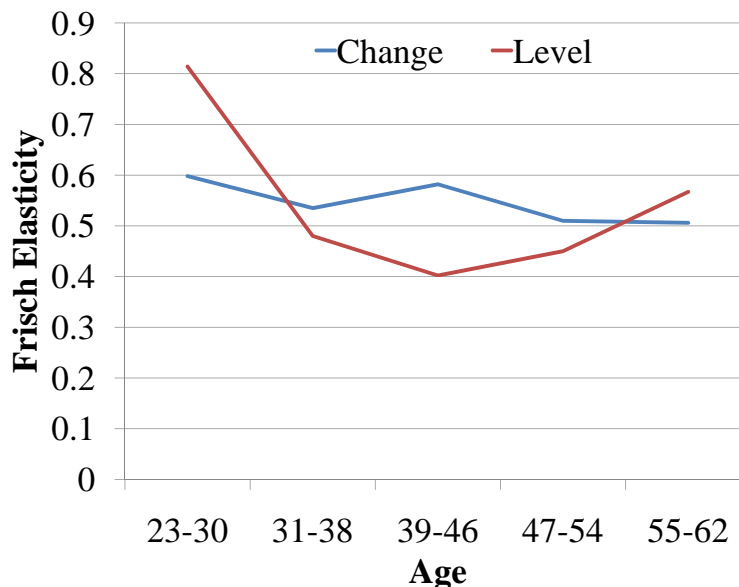


Figure 3.8: Micro Frisch Elasticity Profile

tend to be statistically insignificant.

Generally, I find that the micro Frisch elasticity profiles are flat and that the macro Frisch elasticity profiles are upward sloping after age 55. Comparing the micro and macro profiles indicates that the extensive margin causes agents to become more responsive to changes in their wage towards the end of their working life.

3.5 Conclusion

In this paper I estimate both the micro and macro aggregate Frisch elasticity and Frisch elasticity profile. I estimate that the aggregate micro Frisch elasticity is between 0.6 - 0.64 and the aggregate macro Frisch elasticity is between 1.86 - 2.11. The magnitude of the difference between these ranges is in line with the difference between the microeconomic estimates of the Frisch elasticity and the macro calibration values. Therefore, these estimates indicate that the effect of the extensive margin is large enough to explain the difference in the previous values considered as the micro and macro elasticities.

Table 3.5: Macro Frisch Elasticity Profiles

	I	II	III	IV	V	VI
Specification:	Change	Level	Change	Level	Change	Level
Selection						
Control:	None	None	Predict Missing	Predict Missing	Predict All	Predict All
23-30	1.227 (0.51)	1.069 (1.03)	1.483 (0.92)	1.552 (1.59)	0.936 (0.51)	0.688 (1.23)
31-38	1.196 (0.63)	1.058 (1.25)	1.963 (1.75)	1.497 (2.03)	1.176 (0.79)	0.495 (1.64)
39-46	1.255 (0.87)	1.338 (1.95)	2.276 (2.15)	2.236 (3.18)	1.874 (1.47)	1.232 (2.99)
47-54	1.203 (0.85)	1.177 (1.06)	2.252 (2.16)	1.604 (1.5)	3.189 (3.31)	1.636 (1.21)
55-62	0.763 (1.77)	2.161 (1.9)	3.304 (2.02)	3.679 (2.45)	7.459 (2.36)	5.044 (2.13)

Notes: In the first stage of the change (level) specification education, education squared, age, age squared, age cubed and the interaction between age and education are used as instruments for wage (and consumption). The standard errors are clustered on cohort.

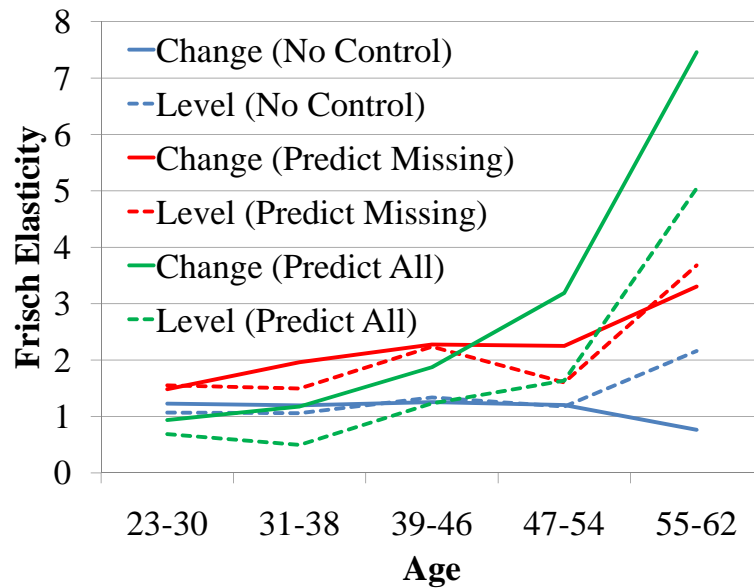


Figure 3.9: Macro Frisch Elasticity Profile

I find that the micro Frisch elasticity profile is generally flat. However, I find that the macro Frisch elasticity profile is upward sloping after the age of 55. These two profiles indicate that on the extensive margin, agents generally become more elastic as they age.

In order to estimate the Frisch elasticities using these specifications, I need to make strong parametric assumptions. Specifically, in order to estimate the Frisch elasticity profiles, I need to make strong assumptions about how age affects hours. I leave it for further work to examine the effect on the results of relaxing these assumptions.

3.6 Appendix

3.6.1 Implications of using participation rate elasticity

In order to calculate the macro elasticity, Chetty et al. (2011) adds the micro (intensive margin elasticity) and the extensive margin elasticity. The authors value for the extensive margin comes from a meta analysis that focuses on studies that

primarily estimate the labor force participation elasticity. The sum of the intensive margin elasticity and the labor force participation elasticity need not be the same as calculating the sum of the extensive and intensive margin Frisch elasticity. Let us consider an economy over two periods which experiences a temporary change in the after tax wage. Let P_e be the number of individuals who choose not to work, P_i be the number of individuals who choose to work, and h_i be the average hours worked by working individuals in the first period. In the second period let P'_e be the number of individuals who choose not to work in the second period, P'_i be the individuals who work in both periods, h'_i be the average hours worked by the individuals who worked in both periods, \tilde{P}'_i be the individuals who work only in the second period, and \tilde{h}'_i be the average hours these new entrants work. Chetty et al. (2011) uses the change in the participation rate elasticity as the extensive margin elasticity and therefore calculates the sum of the percentage change in hours as the expression $\frac{h'_i - h_i}{h_i} + \frac{\tilde{P}'_i}{P_i}$. In contrast, calculating the sum of the percentage change in hours using the actual change in average hours from both the extensive and intensive margin would be $\frac{P'_i h'_i + \tilde{P}'_i \tilde{h}'_i - P_i h_i}{P_i h_i}$. These two expressions are only equivalent if $h_i = \tilde{h}'_i$. If individuals entering the work force tend to work less hours in their first period working, then the estimates of the macro elasticity in Chetty et al. (2011) will be biased downward.

Bibliography

- Aiyagari, S. Rao**, “Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting,” *The Journal of Political Economy*, December 1995, *103* (6), 1158–1175.
- Altonji, Joseph G.**, “Intertemporal Substitution in Labor Supply: Evidence from Micro Data,” *The Journal of Political Economy*, 1986, *94* (3), S176–S215.
- Altuğ, Sumru and Robert Miller**, “The Effect of Work Experience on Female Wages and Labour Supply,” *Review of Economic Studies*, 1998, *65*.
- Armenter, Roc and Stefania Albanesi**, “Understanding Capital Taxation in Ramsey Models,” Working Paper September 2009.
- Atkeson, Andrew, Varadarajan Chari, and Patrick Kehoe**, “Taxing Capital Income: A Bad Idea,” *Federal Reserve Bank of Minneapolis Quarterly Review*, Summer 1999, *23* (3), 3–17.
- Bell, Felicitie and Michael Miller**, “Life Tables for the United States Social Security Area 1900 - 2100,” Actuarial Study 120, Office of the Chief Actuary, Social Security Administration 2002.
- Bénabou, Roland**, “Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?,” *Econometrica*, 2002, *70* (2), 481–517.
- Caucutt, Elizabeth M., Selahattin Imrohoroglu, and Krishna B. Kumar**, “Does the progressivity of taxes matter for economic growth?,” Technical Report 2000.
- Chamley, Christophe**, “Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives,” *Econometrica*, May 1986, *54* (3), 607–622.
- , “Capital income taxation, wealth distribution and borrowing constraints,” *Journal of Public Economics*, 2001, *79* (1), 55 – 69.
- Chang, Yongsun, Joao Gomes, and Frank Schorfheide**, “Learning-By-Doing as a Propagation Mechanism,” *American Economic Review*, December 2002, *92* (5), 1498–1520.

- Chen, Been-Lon, Hung-Ju Chen, and Ping Wang**, “Taxing Capital is Not a Bad Idea Indeed: The Role of Human Capital and Labor-Market Frictions,” Working Paper May 2010.
- Chetty, Raj**, “Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply,” Working Paper 15616, NBER December 2009.
- , **Adam Guren, Day Manoli, and Andrea Weber**, “Does Indivisible Labor Explain the Difference between micro and Macro Elasticities? A Meta-Analysis of Extensive Margin Elasticities,” Working Paper, Harvard January 2011.
- Conesa, Juan Carlos and Dirk Krueger**, “On the Optimal Progressivity of the Income Tax Code,” *The Journal of Monetary Economics*, October 2006, 53 (7), 1425–1450.
- , **Sagiri Kitao, and Dirk Krueger**, “Taxing Capital? Not a Bad Idea After All!,” *American Economic Review*, 2009, 99 (1), 25–38.
- Contreras, Juan and Sven Sinclair**, “Labor supply response in macroeconomic models: Assessing the empirical validity of the intertemporal labor supply response from a stochastic overlapping generations model with incomplete markets,” MPRA Paper 10533, University Library of Munich, Germany September 2008.
- Correia, Isabel H.**, “Should capital income be taxed in the steady state?,” *Journal of Public Economics*, 1996, 60 (1), 147 – 151.
- Cossa, Ricardo, James Heckman, and Lance Lochner**, “Wage Subsidies and Skill Formation: A Study of the Earned Income Tax Credit,” Working Paper, University of Rochester 1999.
- Deaton, Angus**, “Panel Data From Time Series of Cross-sections,” *Journal of Econometrics*, 1985, 30, 109 – 126.
- Domeij, David and Martin Flodén**, “The labor-supply elasticity and borrowing constraints: Why estimates are biased,” *Review of Economic Dynamics*, 2006, 9 (2), 242 – 262.
- Erosa, Andrés and Martin Gervais**, “Optimal Taxation in Life Cycle Economies,” *Journal of Economic Theory*, 2002, 105, 338–369.
- Frazis, Harley and Mark Loewenstein**, “On-the-Job-Training,” 2006.
- French, Eric**, “The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour,” *The Review of Economic Studies*, 2005, 72 (2), 395–427.

- Fuster, Luisa, Ayse Imrohoroglu, and Selahattin Imrohoroglu**, “Altruism, incomplete markets, and tax reform,” *Journal of Monetary Economics*, 2008, 55 (1), 65 – 90. Carnegie-Rochester Conference Series on Public Policy: Advances in Dynamic Public Finance April 20-21, 2007.
- Garriga, Carlos**, “Optimal Fiscal Policy in Overlapping Generations Models,” Working Papers in Economics 66, Universitat de Barcelona. Espai de Recerca en Economia January 2001.
- Gervais, Martin**, Working Paper, University of Southampton February 2010.
- Gouveia, Miguel and Robert Strauss**, “Effective Federal Individual Income Tax Functions: An Exploratory Empirical Analysis,” *National Tax Journal*, June 1994, 47 (2), 317–339.
- Guvenen, Fatih, Burhanettin Kuruscu, and Serdar Ozkan**, “Taxation of Human Capital and Wage Inequality: A Cross-Country Analysis,” NBER Working Papers 15526, National Bureau of Economic Research, Inc 2009.
- Hansen, Gary D. and Selahattin Imrohoroglu**, “Business cycle fluctuations and the life cycle: How important is on-the-job skill accumulation?,” *Journal of Economic Theory*, 2009, 144 (6), 2293 – 2309. Dynamic General Equilibrium.
- Hansen, G.D.**, “The Cyclical and Secular Behaviour of the Labour Input: Comparing Efficiency Units and Hours Worked,” *Journal of Applied Econometrics*, January - March 1993, 8 (1), 71–80.
- Heckman, James J., Lance Lochner, and Christopher Taber**, “Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents,,” *Review of Economic Dynamics*, 1998, 1 (1), 1 – 58.
- Hubbard, Glenn and Kenneth Judd**, “Liquidity constraints, Fiscal Policy, and Consumption,” *Brookings Papers on Economic Activity*, 1986, 1986 (1), 1–59.
- Imai, Susumu and Michael Keane**, “Intertemporal Labor Supply and Human Capital Accumulation,” *International Economic Review*, May 2004, 45 (2), 601–641.
- İmrohorođlu
- İmrohorođlu, Selahattin**, “A Quantitative Analysis of Capital Income Taxation,” *International Economic Review*, May 1998, 39 (2), 307–328.
- Jacobs, Bas**, “Optimal Income Taxation with Endogenous Human Capital,” *Journal of Public Economic Theory*, 2005, 7 (2), 295–315.

- and **A. Lans Bovenberg**, “Human capital and optimal positive taxation of capital income,” *International Tax and Public Finance*, 2009, pp. 1–28. 10.1007/s10797-009-9120-5.
- Jones, Larry E. and Rodolfo E. Manuelli**, “On the Taxation of Human Capital,” Unpublished Manuscript, Northwestern University and University of Wisconsin November 1999.
- , – , and **Peter E. Rossi**, “On the Optimal Taxation of Capital Income,” *Journal of Economic Theory*, 1997, 73, 93 – 117.
- Judd, Kenneth**, “Redistributive Taxation in a Simple Perfect Foresight Model,” *Journal of Public Economics*, 1985, 28, 59–83.
- Judd, Kenneth L.**, “Optimal taxation and spending in general competitive growth models,” *Journal of Public Economics*, 1999, 71 (1), 1 – 26.
- Kapicka, Marek**, “Optimal income taxation with human capital accumulation and limited record keeping,” *Review of Economic Dynamics*, 2006, 9 (4), 612 – 639.
- , “The Dynamics of Optimal Taxation when Human Capital is Endogenous,” Working Paper August 2009.
- Kuruscu, Burhanettin**, “Training and Lifetime Income,” *The American Economic Review*, 2006, 96 (3), 832–846.
- Lucas, Robert E. and Nancy L. Stokey**, “Optimal fiscal and monetary policy in an economy without capital,” *Journal of Monetary Economics*, 1983, 12 (1), 55 – 93.
- MaCurdy, Thomas E.**, “An Empirical Model of Labor Supply in a Life-Cycle Setting,” *The Journal of Political Economy*, 1981, 89 (6), 1059–1085.
- Mincer, Jacob**, “Job Training: Costs, Returns, and Wage Profiles,” Working Paper 3208, NBER 1989.
- Mulligan, Casey**, “The Intertemporal Substitution of Work—What Does the Evidence Say?,” University of Chicago - Population Research Center, Chicago - Population Research Center June 1995.
- Nakajima, Makoto**, “Optimal Capital Income Taxation with Housing,” Working Paper 10-11, Federal Reserve Bank of Philadelphia April 2010.
- Panousi, Vasia**, “Capital Taxation with Entrepreneurial Risk,” Mimeo 2009.

- Pencavel, John**, “A Cohort Analysis of the Association between Work Hours and Wages among Men,” *The Journal of Human Resources*, 2002, 37 (2), pp. 251–274.
- Pistaferri, Luigi**, “Anticipated and Unanticipated Wage Changes, Wage Risk, and Intertemporal Labor Supply,” *Journal of Labor Economics*, 2003, 21 (3), 729–754.
- Rawls, John**, *A Theory of Justice*, Cambridge, Massachusetts: Harvard University Press, 1971.
- Reis, Catarina**, “Optimal Taxation with Unobservable Investment in Human Capital,” Mimeo, MIT 2007.
- Rogerson, Richard and Johanna Wallenius**, “Micro and macro elasticities in a life cycle model with taxes,” *Journal of Economic Theory*, 2009, 144 (6), 2277 – 2292. Dynamic General Equilibrium.
- Smyth, Seamus**, “A Balancing Act: Optimal Nonlinear Taxation in Overlapping Generations Models,” Mimeo, Harvard May 2006.
- Topel, Robert**, “Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority,” Technical Report 3294, NBER 1990.
- Verbeek, M., , and T. Nijman**, “Can Cohort Data be Treated as Genuine Panel Data?,” *Empirical Economics*, 1992, 17, 9–23.