

PERUVIAN ECONOMIC ASSOCIATION

Optimal Taxation and Life Cycle Labor Supply
Profile

Michael Kuklik

Nikita Cespedes

Working Paper No. 8, February 2014

The views expressed in this working paper are those of the author(s) and not those of the Peruvian Economic Association. The association itself takes no institutional policy positions.

Optimal Taxation and Life Cycle Labor Supply Profile*

Michael Kuklik & Nikita Cespedes†

December 2013

Abstract

The optimal capital income tax rate is 36 percent as reported by Conesa, Kitao, and Krueger (2009). This result is mainly driven by the market incompleteness as well as the endogenous labor supply in a life-cycle framework. We show that this model fails to account for the basic life-cycle features of the labor supply observed in the U.S. data. In this paper, we introduce into this model non-linear wages and inter-vivos transfers into this model in order to account for the life-cycle features of labor supply. The former makes hours of work highly persistent and helps to account for labor choices at the extensive margin over the life cycle. The latter allows us to account for labor choices early in life. The suggested model delivers an optimal capital income tax rate of 7.4 percent, which is significantly lower than what Conesa, Kitao, and Krueger (2009) found.

Resumen

La tasa óptima de impuesto a los ingresos de capital en Estados Unidos es 36% según Conesa y otros (2009). Este resultado se deriva de un modelo de ciclo de vida y se debe a la existencia de mercado incompletos y a la oferta laboral endógena. Se muestra que este modelo tiene problemas en explicar algunos aspectos básicos de la oferta de trabajo a lo largo de ciclo de vida de los trabajadores. En este trabajo, introducimos no linealidad en los salarios y transferencias entre personas y logramos reproducir las características de ciclo de vida de la oferta de trabajo. El primer supuesto induce a que las horas de trabajo sean altamente persistentes y ayuda a considerar las decisiones de trabajo en el margen extensivo a lo largo del ciclo de vida. El segundo supuesto permite modelar las decisiones de trabajo a temprana edad. El modelo propuesto sugiere que la tasa de impuesto óptima a los ingresos de capital es de 7,4%.

Keywords: Labor supply, optimal taxation, capital taxation, non-linear wage, inter-vivos transfer.

JEL Classification: E13, H21, H24, H25.

*We would like to thank Mark Aguiar, Mark Bilal, Yongsung Chang, Jay Hong, Ryan Michaels, Ronni Pavan, and the participants at the University of Rochester Macro/Applied Student Workshop for helpful comments and suggestions. All remaining errors are the authors.

†Michael Kuklik: Department of Economics, Long Island University, 1 University Plaza, Brooklyn, New York 11201, michal.kuklik@liu.edu; Nikita Cespedes: Central Bank of Peru (Jr. Antonio Miroquesada 441, Lima 1, Perú) and Pontificia Universidad Católica del Perú, ncespedes@puclp.edu.

1 Introduction

The optimal capital income tax rate is as high as 36 percent as it was suggested by Conesa, Kitao, and Krueger (2009).¹ Before CKK, the literature suggested that the optimal capital income tax rate should be zero or close to zero.² CKK argues that the striking high optimal capital income tax is driven by market incompleteness, the life-cycle profile of workers income and labor supply. Both arguments seem to have empirical support.

CKK introduces both features in a canonical OLG model. In this set up, a government wants the most productive households to work more hours in order to maximize welfare. Since productivity is age dependant, optimally a government would like to condition labor income tax rates to age. In particular, it would like to lower labor income taxes for older, more productive households. While the government cannot tax households based on age, it can use a capital income tax to imitate an age-dependent labor income tax. In a canonical model the most productive households work fewer hours. Therefore, in CKK it is optimal for a government to levy a high tax on capital income to induce the most productive households to work more.

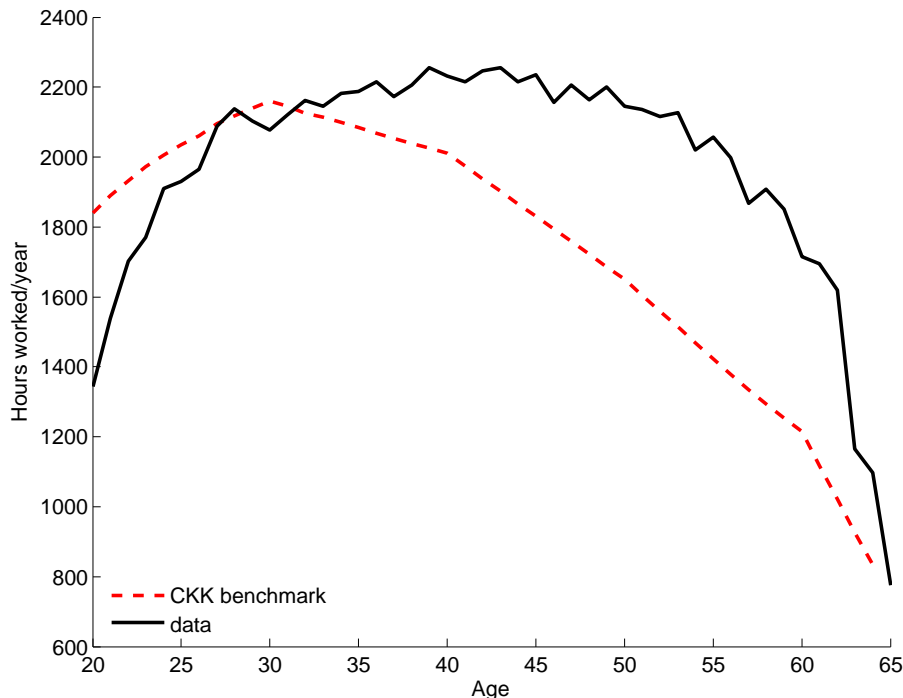
Labor supply dynamics is one of the main factors behind the CKK result. However, this model calibrated to the U.S. economy is unable to represent the life-cycle labor supply profile in the U.S. Figure 1 depicts the mean annual hours worked among men in the U.S. and in the CKK benchmark model. In the data mean annual hours exhibit an inverted U-shape over the life cycle. Hours worked increase in early life during the 20s and stay constant up to age 55, when they start to decrease. In the CKK benchmark model labor supply is monotonically decreasing from the age of 30 until retirement. A counterfactual labor supply in the CKK model therefore makes their optimality results inappropriate to the U.S. tax policy debate.

In this paper we show that a counterfactual labor supply in the CKK benchmark model is a major reason for the high optimal capital income tax rate. Furthermore, we extend the CKK model in order to account for the life-cycle features of labor supply in the U.S. Recall that the canonical life-cycle CKK model features Social Security, incomplete markets, and heterogeneity in individual productivity. In order to address labor supply issues we incorporate two additional elements into the model. The first is a non-linear mapping between hours worked and wages as in Hornstein and Prescott (1993) and Erosa, Fuster, and Kambourov (2010). The convexity of the wage schedule makes hours of work highly persistent and helps to account for labor choices at the extensive margin over the life cycle. The second

¹Hereafter referred to as CKK.

²Judd (1985), Chamley (1986), Chari and Kehoe (1999b), among other studies support the claim that optimally capital should not be taxed at all.

Figure 1: Labor supply in a canonical life-cycle model



is the inclusion of inter-vivos transfers from living parents to adult children, which allows us to model labor choices in early life up to age 30. We use this calibrated model to evaluate an optimal tax system consisting of a progressive labor income tax rate and a proportional capital income tax rate.

We find that the optimal U.S. capital income tax rate is 7.4 percent, which is significantly lower than what CKK report. The corresponding labor income tax is slightly progressive with an upper bound of 36 percent. The result is driven by the high persistence of hours of work in the second half of the working life. Due to the non-linear wage, households over the age of 40 work full time until the age of 55. This is in stark contrast to CKK, where households monotonically reduce their hours of work as they age. Thus, in our model the planner has no need to use the capital income tax to create an incentive for older households to work, while in CKK the planner resorts to imposing a high capital income tax, which exerts an upward pressure on the implicit price of leisure and thus stimulating the older, highly productive households to work for more hours.

Our results are consistent with Peterma's (2010) findings on the relationship between the Frisch elasticity of the labor supply and the optimal capital income tax. Our model generates a flat labor elasticity profile even though the utility specification we use implies a non-constant Frisch elasticity. Finally, we show that the inter-vivos transfer has a very small impact on the optimal capital income tax. Younger households are less productive than older ones. Thus, the cost of creating an incentive for

young households to work outweighs the benefit of the extra output they could produce.

Section 2 shows the importance of the labor supply profile on the optimal tax in the CKK framework. Section 3 describes the model and elaborates in detail on the inter-vivos transfers and non-linear wage. Section 4 presents a calibration of the model. Section 5 reports the results of an optimal tax experiment and provides some intuition. Finally, Section 6 offers concluding remarks.

2 The importance of labor supply in the optimal tax analysis

We perform a simple experiment that motivates the crucial role of the labor supply profile in the optimal income tax rate into the CKK framework. We recalibrate the CKK benchmark model to match the life-cycle average hours of work in the U.S.; we call it the modified model. In our experiment we find a new and arbitrary productivity profile such that the average hours of work for households older than 26 years of age match the U.S. data.³ To make the results of our tax optimization comparable to CKK, we adjust government consumption to keep its ratio to output unchanged. The remaining parameters are fixed at the CKK benchmark levels. In particular, the parameters that have a significant impact on the optimal capital tax, including the risk-aversion coefficient in the CRRA Cobb-Douglas utility function, $\sigma = 4$, and the discount factor, $\beta = 1.001$, are the same as in the CKK benchmark.

The right panel of Figure 2 presents the labor productivity profile of the original and modified models. The productivity profile in the modified model is much steeper, with the highest productivity shifted toward the older ages.⁴ Then, we compute the optimal capital income tax rate of the modified model as in CKK.

We find that the labor supply induced by the productivity profile has a significant impact on the optimal tax system. In the modified model, the optimal capital income tax is 30 percent, which is 12 percentage points lower than in CKK (Table 1). To compensate for the lower revenue from the capital tax, the marginal tax rate on labor income is 4 percentage points higher than in CKK. The labor income tax deduction and the degree of labor tax progressivity are approximately the same in both models.

The intuition behind this result is as follows. The government needs to levy taxes on income to pay for public spending. The optimal government policy is to tax the most productive households the least (Stiglitz 1987). Ideally, the labor income tax rate should decrease with age, since households' productivity increases with age (Erosa and Gervais 2002). When labor income tax is not allowed to depend on age, the government can affect the way in which households substitute labor intertemporally by taxing

³Average hours of work by age were estimated from PSID. The first 6 years of the representative labor profile were not targeted because productivity is not the most important factor in low employment among young households; we address this problem in the later sections.

⁴Recall that in the CKK model, the assumption driving labor supply is the average labor productivity profile.

Figure 2: Modified model: Hours and productivity profiles

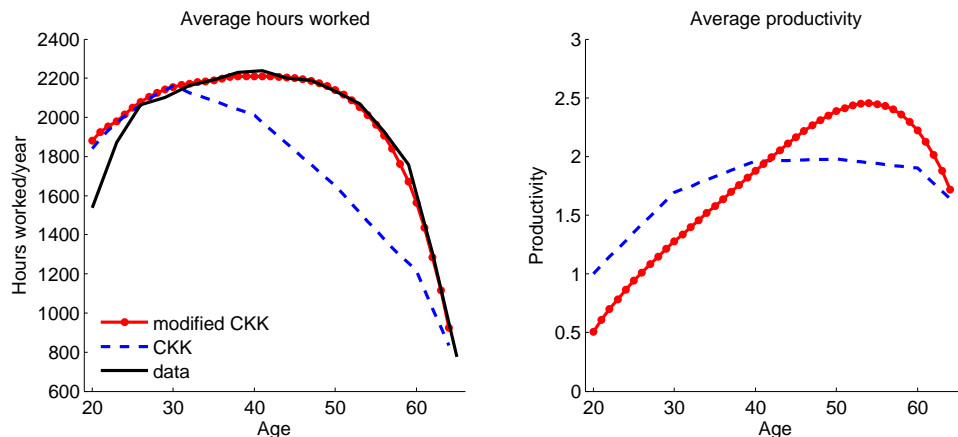


Table 1: Simulating impact of labor supply on the optimal tax

Optimal tax system	CKK	Modified CKK
Capital income tax rate (%)	42	30
Labor income tax rate (%)	21	25

Note: A labor income tax rate refers to an upper bound on marginal rate in the Gouveia and Strauss (1994) progressive tax function. In CKK the optimal capital tax rate is 36 percent. In replicating their experiment we obtained 42 percent. We attribute this discrepancy to the technical differences in our implementation of their model (continuous optimization over labor choice in the household's problem and our usage of the Nelder-Mead algorithm for the optimal tax search).

capital income. The higher the capital income tax rate, the stronger the incentive for households to work more.

In the CKK model, hours of work for households over 30 decline steadily until retirement, while households' productivity increases up to age 40 and then stays constant until age 60. The optimal tax is as high as 42 percent in order to stimulate households over 30 years of age to work more. In the modified model, households start to work fewer hours only in their mid 50s; this coincides with a peak in average productivity (Figure 2). The government only targets households over 50 years old, and exerts pressure with the capital income tax in order to induce them to work more, while in CKK all households over 30 are targeted by the government policy. The benefits of inducing households to work more are much smaller than in the CKK model, but the cost of distorting the capital margin is unchanged. Thus, in the modified model the optimal capital income tax is 12 percentage points lower than in the CKK model.

This simple exercise shows that the modeling of the labor supply life-cycle profile is crucial for the

quantitative analysis of optimal taxation. But imposing our ad hoc productivity profile does not address the source of the problem. In the next section, we develop a more comprehensive general equilibrium model that successfully captures the life-cycle features of the labor supply observed in the data.

3 The Model

The model is based on the standard overlapping generations model as in CKK. This model is extended in two dimensions. First, we follow Hornstein and Prescott (1993) in modeling production technology that leads to non-linear wages in competitive equilibrium. Second, we introduce inter-vivos transfers motivated by Gale and Scholz (1994) findings.

3.1 Demographics

The economy is populated with J overlapping generations. Households face life-span uncertainty and survive from age j to $j + 1$ with a conditional probability ψ_j . At age J death is certain, $\psi_J = 0$. The population grows at an annual rate of ϕ . Households cannot be insured against idiosyncratic mortality risk. Therefore, upon death, households leave an accidental bequest, Tr , which is distributed equally among all living households. Households work up to an exogenous retirement age j^R . Households are heterogeneous in education level and labor productivity. We divide households into two educational groups: college and non-college. Each period, working households face idiosyncratic productivity shocks. The young households randomly receive an exogenous inter-vivos transfer from their parents.

3.2 Endowments and preferences

During working age, households receive a stochastic endowment of effective units of labor, $e_{i,j}$, which depends on education type i and age j . We adopt the following process for this endowment:

$$\log e_{i,j} = \bar{e}_{i,j} + f_i + z_{i,j} + \epsilon_j \quad (1)$$

where $\bar{e}_{i,j}$ is an age-dependent average log-income, $f_i \sim N(0, \sigma_{i,f}^2)$ is a fixed effect, $z_{i,j}$ is a persistent component following an AR(1) process,

$$z_{i,j} = \rho_{i,z} z_{i,j-1} + \eta_{t,j}, \quad \eta_{i,j} \sim N(0, \sigma_{i,\eta}^2), \quad (2)$$

and ϵ_j is a transitory component distributed according to $N(0, \sigma_\epsilon^2)$.

Every period, households are endowed with one unit of time, which can be allocated between work or leisure. Preferences of a household with education type i are represented by a time separable utility function $u(c_j, 1 - h_j)$. Similarly, the discounted utility at the beginning of the working age is represented by

$$E \left[\sum_{j=1}^J \beta^{j-1} u(c_j, 1 - h_j) \right] \quad (3)$$

where c_j is consumption, h_j is hours worked, and β is a time discount factor. Expectation is taken with respect to the idiosyncratic labor productivity shock, inter-vivos transfer, and the mortality risk.

3.3 Inter-vivos transfer

Data inspection shows that young people work less hours. Figure 1 illustrates how U.S. males of age 20 work on average 2000 hours a year. In the first 10 years of their working lives, they increase their working hours, to reach an average of 2200 hours a year, which stays constant up to age 55. The reason for the initial low intensity of work can be attributed to a lack of experience (low productivity) or attending college. But in the standard OLG model young workers with low productivity work the maximum hours as they have no savings and are borrowing constrained. Furthermore, college attendance can partially explain the low intensity of work in the early 20s. In fact, the difference in working hours between those attending college and those with a high school degree is very small. We propose inter-vivos transfers (transfers from living parents to children) as a likely explanation, which will help us to account for the low intensity of labor among workers below the age of 25.

Parents may contribute toward a young adult person's investment in education, make cash transfers at various points throughout the persons early adulthood, or provide support in the form of common goods like housing, food, and health care. This allows an adult child to maintain a constant or acceptable living standard without working full time. Previous studies have found inter-vivos transfers to be quite large (Gale and Scholz 1994), and thus, they likely play an important role in the adult child's labor decisions.⁵

We model an inter-vivos transfer as an exogenous shock iv following a Markov process $\pi^{iv}(iv', iv)$. The transfer from parents to children takes two values $\{0, \nu\}$. Households can receive a shock up to age $j^{iv} = 55$, then we can properly distinguish the transfer values according to: $iv \in \{0, \nu\}$ if $j < j^{iv}$ and $iv = 0$ if $j \geq j^{iv}$. This also implies that once a worker reaches age j^{iv} , the transition probability of the

⁵Gale and Scholz (1994) estimated the annual flows of inter-vivos transfers at \$2489 billion, college transfers at \$1,141 billion, and bequests as \$3,708 billion.

the inter-vivos transfer collapses to a zero transfer state.

3.4 Technology

Production takes place in a large number of plants. A new plant can be opened at zero cost. All workers at a plant have the same effective labor productivity e and work the same number of hours h . The production function of a plant that hires n workers and is defined as:

$$f(z, k, h, e) = zh^\xi k^\alpha e^{1-\alpha} \quad (4)$$

where h denotes hours worked per worker, k is an amount of capital in a plant, and e is the effective units of labor of each worker. Total factor productivity is denoted by z . Since there is no aggregate shock in the economy, z is normalized to one. Plants are subject to constant returns to scale for a given workweek h . Capital depreciates at rate δ .

Parameter ξ is the elasticity of production with respect to hours. If ξ is set to $1 - \alpha$, the production function is reduced to the standard Cobb-Douglas technology; this makes the decomposition of labor into hours and effective productivity irrelevant as only the total hours of effective labor matters (Osuna and Rios-Rull 2003). When $\xi \neq 1 - \alpha$, hours and effective labor are no longer perfect substitutes. Also notice that for $\xi = 1$, output is proportional to hours; this case is often interpreted as reflecting a lack of worker fatigue.

Note that papers on workweek hours and non-linear wages, such as Prescott and Townsend (1984), Hansen (1985), Rogerson (1988), Prescott and Rios-Rull (1992), Hornstein and Prescott (1993), Osuna and Rios-Rull (2003), and Fitzgerald (1998), use lotteries to get around non-convexities in general equilibrium macroeconomic models with a representative agent. Our economy is populated by heterogeneous agents. Thus the use of lotteries is redundant.

3.4.1 Plant's problem

To illustrate how non-linear wages are generated, consider a plant that hires workers with productivity e , who work h hours in a week. We assume that the plant observes the productivity of workers, taking an interest rate r and the salary $\hat{w}(h, e)$ paid to a worker as given. Since a production exhibits constant returns to scale, we can normalize the size of the plant to one worker. Plants choose capital k to maximize profits:

$$\pi = \max_k h^\xi k^\alpha e^{1-\alpha} - (r + \delta)k - \hat{w}(h, e). \quad (5)$$

The solution to the plant's optimization problem is an optimal capital k^* . We define the optimal level of capital per effective labor as:

$$\kappa(h) \equiv \frac{k^*}{e} = \left[\frac{\alpha h^\xi}{r + \delta} \right]^{1/(1-\alpha)}. \quad (6)$$

Since there is no entry cost, the zero-profit condition must hold:

$$0 = \pi = h^\xi \kappa(h)^\alpha e - (r + \delta) \kappa(h) e - \hat{w}(h, e). \quad (7)$$

From equation 7 we obtain the equilibrium salary $\hat{w}(h, e)$ for working hours h and effective labor e . Salary $\hat{w}(h, e)$ can be expressed as a product of effective labor e and the wage schedule $w(h)$:

$$\hat{w}(h, e) = w(h) e$$

where

$$w(h) = (r + \delta) \left[\frac{1 - \alpha}{\alpha} \right] \left[\frac{\alpha h^\xi}{r + \delta} \right]^{1/(1-\alpha)}. \quad (8)$$

The wage schedule $w(h)$ is non-linear in hours of work unless $\xi = 1 - \alpha$, and it augments in hours at an increasing rate for $\xi > 1 - \alpha$. Plants can be indexed by workweek hours h , workers' effective units of labor e , and capital per worker k . Since each plant employs one worker, we use a measure of households $\Phi(i, j, s, a)$ to aggregate over firms. Aggregate output can thus be written as

$$Y = \int h(i, j, s, a)^\xi k^\alpha e^{1-\alpha} \Phi(di, dj, ds, da). \quad (9)$$

3.5 Government

Government runs a self-financed Social Security System. Retired households receive benefits ss , which are independent of household income history. Social security is financed through a proportional social security tax τ_{ss} , which is levied up to a maximum labor income \bar{y} . Households that work h hours with effective labor productivity $e_{i,j}$, $e_{i,j} = e$, has a gross income $w(h) e_{i,j} h$. Gross income is the base for social security contributions, which is denoted by $\tau_{ss} \min\{w(h) e_{i,j} h, \bar{y}\}$.

In this economy the government funds its exogenously given consumption G by levying taxes. The tax system consists of three parts. First is a proportional tax τ_c on consumption expenditures. Second is a tax $T^k [y_k]$ on households' capital income, $y_k = r(a + tr^b)$, where r is the interest rate, a denotes

assets at the beginning of a period, and tr^b is an accidental bequest. Third is a tax $T^l[y_l]$ on households' taxable labor income y_l . Firms pay half of the Social Security; the other half is paid by households from gross income and is tax deductible. Thus, a household's taxable labor income is gross income adjusted for the other half of the contribution,

$$y_l = \begin{cases} w(h) e_{i,j} h - 0.5\tau_{ss} \min\{w(h) e_{i,j} h, \bar{y}\} & j < j^R \\ 0 & j \geq j^R \end{cases}. \quad (10)$$

We assume that capital income is taxed at a constant marginal tax rate τ_k ; thus $T^k[y_k] = \tau_k y_k$. The labor income tax is progressive and represented by the function $T^l(y)$. Thus, the tax system is denoted by $\{\tau_c, T^l[y_l], T^k[y_k]\}$. The total income tax is given by the function $T[y_l; y_k] = T^l[y_l] + T^k[y_k]$.

3.6 Markets

Markets are incomplete and only the purchase of a one-period risk-free bond is permitted. Thus, households cannot buy an insurance contract against idiosyncratic labor income shocks, but they can get insurance by buying risk-free bonds. Households are not allowed to borrow. There are no annuity markets; thus households cannot insure a mortality risk. Upon death, a household's remaining wealth is distributed equally among all households in the form of an accidental bequest tr^b .

3.7 Household problem

Households face two shocks: an inter-vivos transfer iv and a labor productivity shock $e_{i,j}$. We combine them into one random vector $s = (iv, e_{i,j})$ represented by a Markov probability matrix $\pi(s, s')$. Households are heterogeneous in four dimensions: education type i , age j , shocks s , and asset level a . Thus, a household can be characterized by a vector (i, j, s, a) . We denote a household's value function by $V_{i,j}(s, a)$ and the measure of households by $\Phi(i, j, s, a)$. Each period, a working household characterized by (i, j, s, a) learns the realizations of productivity shocks, then makes decisions about consumption c , hours of work h , and saving a' . All decisions are made under uncertainty about death at the end of the period, future inter-vivos transfers, and labor productivity shocks. Households receive a gross income $w(h) e_{i,j} h$ from h hours of work. During retirement, households face only a mortality risk as Social Security benefits ss are deterministic. Given prices, taxes, and transfers, a household solves the following problem:

$$V_{i,j}(s, a) = \max_{c, a', h} \{u(c, 1-h) + \beta \psi_j \mathbb{E}[V_{i,j+1}(s', a')]\} \quad (11)$$

subject to

$$(1 + \tau_c) c + a' = w(h) e_{i,j} h - \tau_{ss} \min \{w(h) e_{i,j} h, \bar{y}\} + tr^{iv} + (1 + r) (a + tr^b) - T[y_l; r(a + tr^b)] \quad \text{for } j < j^R \quad (12)$$

$$(1 + \tau_c) c + a' = ss + (1 + r) (a + tr^b) - T[0; r(a + tr^b)] \quad \text{for } j \geq j^R$$

$$y_l = w(h) e_{i,j} h - 0.5\tau_{ss} \min \{w(h) e_{i,j} h, \bar{y}\} \quad (13)$$

$$a' \geq 0, c \geq 0, 0 \leq h \leq 1 \quad (14)$$

3.8 Stationary equilibrium

Given a fiscal policy $\{G, \tau_c, T[y_l; y_k], \tau_{ss}, \bar{y}\}$, a recursive competitive equilibrium is a set comprising a value function V , household decision rules $\{c, h, a'\}$, a wage schedule function $w(h)$, a measure of households $\Phi(i, j, s, a)$, and a capital renting policy of a plant k such that:

1. The value function and decision rules solve the household problem (11, 12),
2. Plants make zero profit and choose capital optimally, (6, 7),
3. The government budget is balanced,

$$G = \int [\tau_c c(i, j, s, a) + T[y_l; y_k]] \Phi(di, dj, ds, da), \quad (15)$$

4. Social Security is self-financed:

$$\tau_{ss} \int_{j < j^R} \min \{w(h) e_{i,j} h, \bar{y}\} \Phi(di, dj, ds, da) = ss \int_{j \geq j^R} \Phi(di, dj, ds, da), \quad (16)$$

5. Accidental bequest transfers satisfy:

$$tr^b \int \Phi(di, dj, ds, da) = \int (1 - \psi_j) a'(i, j, s, a) \Phi(di, dj, ds, da), \quad (17)$$

6. Markets clear:

$$\int c'(i, j, s, a) \Phi(di, dj, ds, da) + K' + G = Y + (1 - \delta)K + Tr^{iv}, \quad (18)$$

where:

$$K = \int a'(i, j, s, a) \Phi(di, dj, ds, da) \quad (19)$$

$$Y = \int h(i, j, s, a)^\xi k^\alpha e^{1-\alpha} \Phi(di, dj, ds, da) \quad (20)$$

$$Tr^{iv} = \int tr^{iv} \Phi(di, dj, ds, da) \quad (21)$$

7. The distribution of agents follows the law of motion:

$$\Phi(i, j+1, s', a') = \begin{cases} \int \psi_j \pi(s, s') \Phi(i, j, ds, da) & \text{if } a' = a'(i, j, s, a) \\ 0 & \text{otherwise} \end{cases} \quad (22)$$

$$\Phi(i, 1, s, a) = \begin{cases} \bar{\pi}^i(i) & \text{if } a = 0, s = \bar{s} \\ 0 & \text{otherwise} \end{cases} \quad (23)$$

A steady state for this economy is a recursive competitive equilibrium where all prices, aggregate variables, and distribution are time invariant.

4 Calibration

In calibrating the model we follow CKK as closely as possible. To calibrate the productivity profile, we refer to Erosa, Fuster, and Kambourov (2010). Households enter the economy at age 20 and die with certainty at age 100. The retirement age is set to be 65. The population growth rate ϕ is set to 0.011, which is the long-run average in the United states. The conditional survival probabilities ψ_j are taken from Bell and Miller (2002).

The utility function takes the following form:

$$u(c, 1-h) = \ln c + \varphi \frac{(1-h)^{1-\sigma}}{1-\sigma}, \quad (24)$$

which allows for large differences in individual labor productivity with small differences in hours worked, as pointed out by Erosa and Gervais (2002). Following Osuna and Rios-Rull (2003) households are endowed with 5200 hours a year. The preference parameter φ governs the taste for leisure; it is chosen so that on average households work 42 percent of the available time. Following Erosa, Fuster, and Kambourov (2010), the coefficient σ of leisure is set to 3 so that the Frisch elasticity of labor when the household is working full time is 0.5. The discount factor β is chosen to match the capital-to-output ratio of 2.7 in the U.S.

Using the Time and Money Transfer Supplement of 1988 from the PSID, Altonji, Hayashi, and Kotlikoff (1996) estimate the annual probability of a transfer from parents to children to be 0.24 with an average transfer amount of \$1851. We set the probabilities in the inter-vivos transfer process to $\pi^{iv}(iv' = \nu, iv = 0) = 0.24$, $\pi^{iv}(iv' = 0, iv = \nu) = 1.0$. The transfer value ν is set to 0.1, which is in line with the Altonji, Hayashi, and Kotlikoff (1996) estimate of the average transfer value normalized by the median household income of \$42,000 in 2000.

The individual effective units of labor follow the process in equation (8). For each education type the parameters $(\sigma_{i,f}^2, \rho_{i,z}, \sigma_{i,\eta}^2)$ of the productivity process are taken from Erosa, Fuster, and Kambourov (2010), who estimate values for college and non-college males using the PSID data. The variance of the transitory component, $\sigma_{i,\epsilon}^2$, is calibrated to match the cross-sectional variance of household labor income at age 55, which Storesletten, Telmer, and Yaron (2004) report to be 0.8 for non-college⁶ and 0.7 for college households. All shocks are discretized using a Tauchen routine with a grid size of 15 for the autoregressive process, 4 for temporary shocks, and 2 for fixed effects.

Since our model has an active extensive margin in the labor supply decision, working households are not a random selection of the population. Therefore, we cannot follow a common methodology of using the age profile of mean wages as an age-dependent component of the productivity process. We adopt the methodology from Erosa, Fuster, and Kambourov (2010) to calibrate the average productivity profile. The average productivity for each type takes the form of a quartic polynomial in age. We calibrate each profile to match the average hours of work for the corresponding education type. The profiles are shown in Figure (3).

We use a standard value for the capital share, $\alpha = 0.36$. The depreciation rate δ is set to 0.072 to match an investment-to-output ratio of 25.5 percent observed in the data. The hours elasticity of wages was estimated by Aaronson and French (2004) to be 0.4. A corresponding elasticity for the wage function in equation 8 is $\frac{\xi}{1-\alpha} - 1$. Thus, we set $\xi = 1.4(1 - \alpha)$.

Government consumption G is calibrated to match the government spending-to-output ratio of 0.17

⁶ Average values for high school and non-high school individuals.

Table 2: Benchmark parameters

Parameters	Value	Target/Source
Demographics and Preferences		
Maximum age, J	81(100)	
Retirement age, j^R	46(65)	
Population growth, ϕ	0.011	U.S. data
Survival probabilities, ψ_j		Bell and Miller (2002)
Discount factor, β	0.962	$K/Y = 2.7$
Taste for leisure, φ	0.518	Avg. hours of work = 0.42
Leisure coefficient, σ	3.0	avg. labor Frisch elasticity of 0.5
Productivity process (non-college, college)		
Fixed effect, σ_f^2	(0.097, 0.072)	Erosa, Fuster, and Kambourov (2010)
AR(1) autocorrelation, ρ_z	(0.94, 0.97)	Erosa, Fuster, and Kambourov (2010)
AR(1) innovation, σ_η^2	(0.019, 0.021)	Erosa, Fuster, and Kambourov (2010)
Transitory shock, σ_ϵ^2	(0.0716, 0.0592)	$\sigma_{y_{i,55}} = (0.8, 0.7)$
Average productivity, $\bar{e}_{i,j}$		mean hours of work
Non-college	(0.219, -0.028 , $0.177e^{-2}$, $-0.245e^{-4}$, $0.064e^{-6}$)	
College	(5.296, -0.662 , $2.802e^{-2}$, $-4.347e^{-4}$, $2.206e^{-6}$)	
Technology		
Capital share, α	0.36	U.S. data
Wage schedule, ξ	$1.4(1 - \alpha)$	Aaronson and French (2004)
Depreciation rate, δ	0.072	$I/Y = 0.255$
TFP, z	1.0	normalization
Government		
Gov. consumption, G	1.86	$G/Y = 0.17$
Consumption tax, τ_c	0.05	Mendoza, Razin, and Tesar (1994)
Social security tax, τ_{ss}	0.124	U.S. Data
Max. SS tax, \bar{y}	0.586	$2.5 \cdot \text{avg. income}$

in the U.S. The consumption tax rate τ_c is set to 0.05 following Mendoza, Razin, and Tesar (1994). The Social Security tax τ_{ss} is set to 0.124. A cap on Social Security contributions, \bar{y} , is set to be 2.5 times the average income. The progressive labor income tax function takes the form of

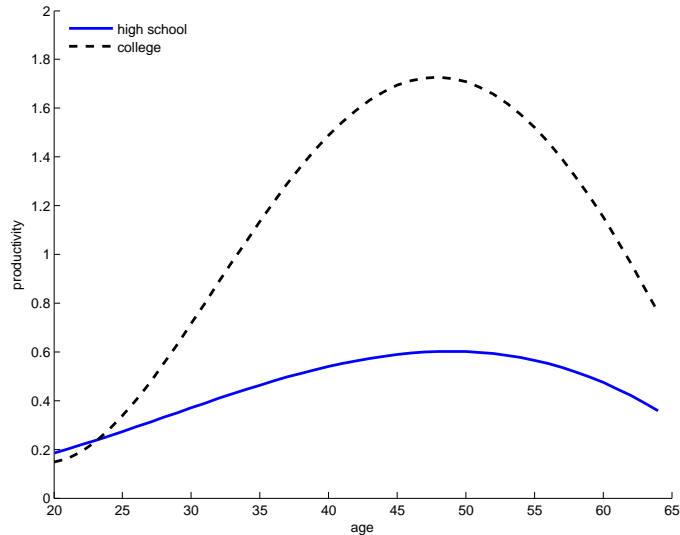
$$T^p(y; a_0, a_1, a_2) = a_0 \left(y - (y^{-a_1} + a_2)^{-1/a_1} \right) \quad (25)$$

where the parameter a_0 represents an upper limit on the marginal tax rate, a_1 determines the progressivity, and a_2 is a scaling parameter. Gouveia and Strauss (1994) introduced this functional form and estimated values of parameters that approximated the U.S. tax code as $a_0 = 0.258$ and $a_1 = 0.768$. The scaling parameter a_2 is used to clear the government budget constraint. Since Gouveia and Strauss (1994) combined labor and capital income in their estimations, we calibrate the model by progressively taxing combined labor and capital income,

$$T[y_l; y_k] = T^p(y_l + y_k; a_0, a_1, a_2). \quad (26)$$

Figure 4 presents the labor supply profiles of the benchmark model. The model succeeds in replicating the life-cycle features of labor supply. The benchmark model achieves a perfect fit of average hours to the U.S. data and the participation rate is very close to the U.S. data as well.

Figure 3: Benchmark: productivity profile



5 Experiments

In this exercise we look for the optimal tax system, $\{\tau_c, T^l[y_l] = T^p(y_l; a_0, a_1, a_2), T^k[y_k] = \tau_k y_k\}$, consisting of a consumption tax τ_c , a progressive tax on labor income, and a proportional tax on capital income. This implies a total income tax function of the form

$$T[y_l; y_k] = T^p(y_l; a_0, a_1, a_2) + \tau_k y_k. \quad (27)$$

The consumption marginal tax rate is the same as in the benchmark. The parameter a_2 , which controls the level of progressive income tax, is used to clear the government budget constraint. The interest rate and Social Security tax are used to clear the asset market and entitlement program, respectively. Parameters (a_0, a_1, τ_k) specify different versions of the tax system. All other parameters of the model are fixed at the benchmark level.

To rank the tax systems over the parameters (a_0, a_1, τ_k) , we adopt the same social welfare function as CKK. We assume that the government wants to maximize the ex ante lifetime utility of a household born in a steady state implied by the given tax system. A newborn household has zero assets and receives on average productivity shock $s = \bar{s}$. The government's objective function is

$$SWF(a_0, a_1, \tau_k) = \int V(t, j = 1, s, a | (a_0, a_1, \tau_k)) \Phi(dt, j = 1, ds, da), \quad (28)$$

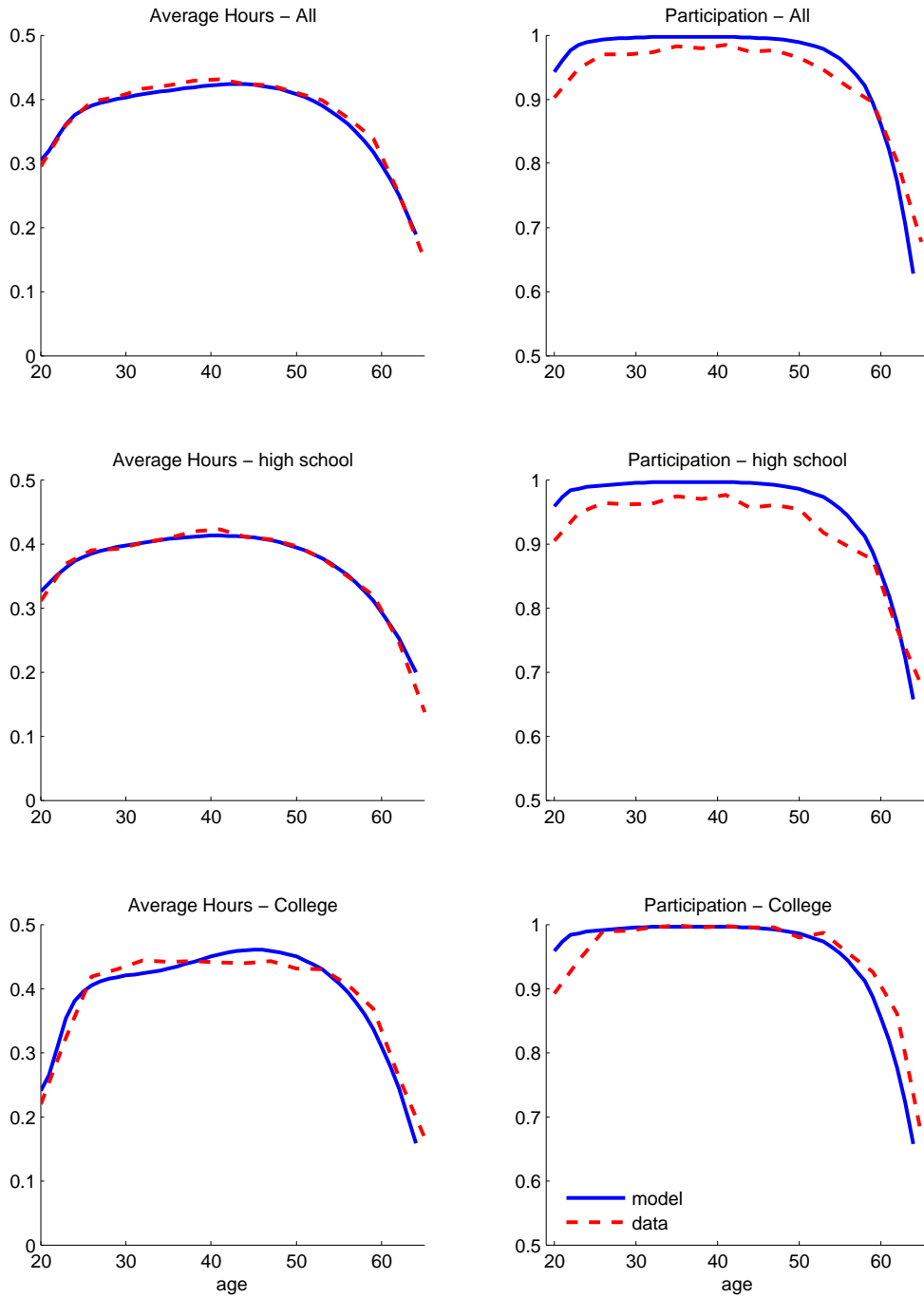
We use the Nelder-Mead algorithm to maximize the welfare function $SWF(a_0, a_1, \tau_k)$ with respect to three parameters (a_0, a_1, τ_k) .

5.1 Optimal tax system

The optimal capital income tax is 7.4 percent (36 percent in CKK). The optimal labor income tax is characterized by $a_0 = 0.355$, $a_1 = 4.23$, and $a_2 = 23903$. The optimal progressive labor tax has an upper bound on the marginal rate of 35.5 percent, which is significantly higher than 25.8, the current U.S. value (23 percent in CKK). In the benchmark everyone pays labor income tax regardless of income level, while in the optimal tax system, there is a deductible of \$2,100 (\$7,200 in CKK). The progressive tax function is presented in Figure 5.

The upper left panel of Figure 6 depicts the participation rate in the benchmark and optimal tax economies. In the benchmark, participation is high and persistent between age 25 and 49. Optimal tax policy increases participation among old households, who defer their exit from the labor market on average by 5 years. Since the labor income tax rate increases by 10 percentage points in the optimal

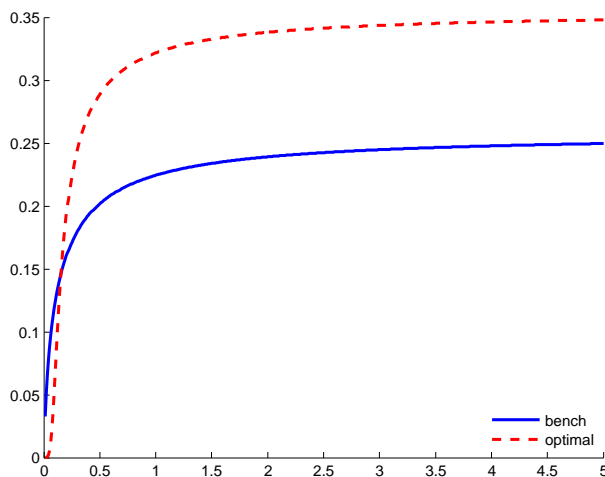
Figure 4: Benchmark: labor supply



tax system, the disposable incomes of older households with high productivity experience a considerable decrease. Consequently, it takes more than a year of work for households to accumulate enough assets to retire earlier.

Average positive hours are presented in the middle left panel of Figure 6. Young households, up to

Figure 5: Progressive labor income tax function, average tax rate



age 25, increase their working hours in comparison to the benchmark. At age 20, households work 71 hours a year more than in the benchmark. The optimal labor income tax is more favorable to low-income households, many of which are young due to the life-cycle average productivity profile. In the optimal tax system there is a deduction of \$2,100 and the average tax rate is lower for low-income households (Figure 5). Thus, the optimal tax system is less distortionary for this age group.

Households older than 27 work on average fewer hours than in the benchmark. The drop in hours worked increases progressively with age. Initially at age of 27, average hours decrease by 8 hours a year. Later on, at age 40, average positive hours decline by 18 hours a year, while at age 60, households work 71 hours a year less than in the benchmark. The optimal tax policy causes significant distortions to the labor margin due to the higher labor income tax, but hours decrease only slightly as the non-linearity in wages prevents a large response.

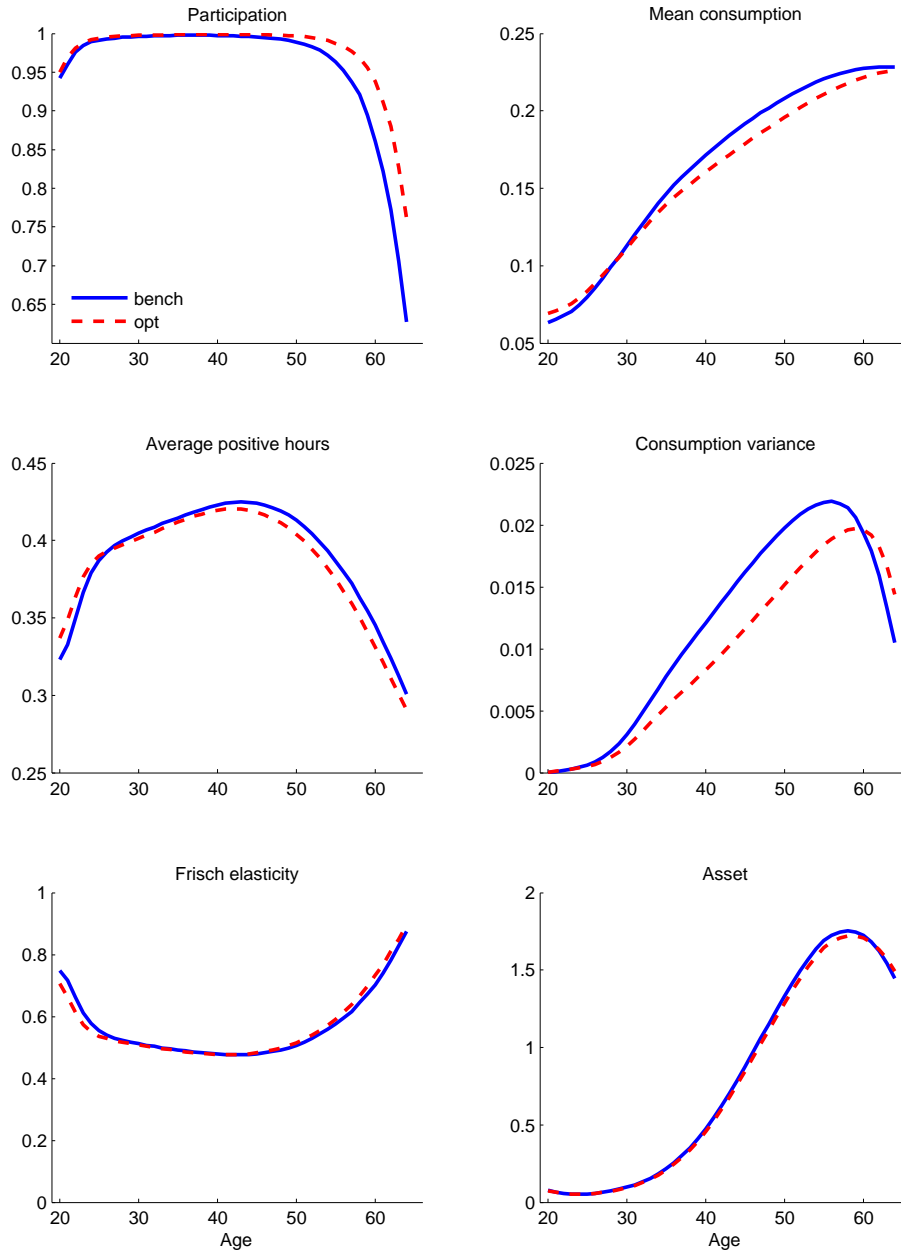
The Frisch elasticity, presented in the bottom left panel of Figure 6, is an inverted picture of the average positive hours. The labor elasticity is high among the newborn and declines up to age 25, to reach a plateau around 0.5, then starts to increase again at age 45 until retirement. The Frisch elasticity, as implied by the utility function in our calibration, takes the form of

$$f = \frac{1-h}{\sigma h}. \quad (29)$$

It is a function of hours worked. The optimal tax policy thus lowers the Frisch elasticity of households less than 27 years old and increases the elasticity of households over 50. This is a direct implication of the changes in hours worked among households in these age groups due to the optimal tax system.

The optimal tax system also has significant implications in terms of redistribution. The two top

Figure 6: Optimal tax system



right panels in Figure 6 show the mean and variance of consumption. The optimal tax policy increases consumption among households below age 27; this is the age group that benefits from the deduction and decline in the average tax rate at low labor income levels in the optimal tax system. The consumption of households over the age of 27 decreases. The optimal labor tax schedule is significantly higher for households with high incomes, even up to 10 percentage points higher (Figure 5). This impacts households' disposable incomes and, thus, their consumption. The optimal tax policy, thus, contributes

Table 3: Changes in aggregate variables of results with optimal tax system respect to benchmark economy

Aggregates	Percent change
Interest rate, r	-2.68
Labor supply, N	-1.33
Output, Y	-1.52
Capital, K	-0.29
Consumption, C	-2.27

to a decrease in inequality. The variance of consumption decreases significantly, especially among older households. This is due to the shifting of the tax burden toward older households with more assets. The same effect is observed in CKK.

The impact of the optimal tax system on aggregates is shown in Table 3. Capital drops by 0.29 percent relative to the benchmark. As a consequence aggregate output and consumption fall by 1.5 and 2.3 percent, respectively. The interest rate decreases by 2.7 percent, while aggregate labor supply drops by 1.4 percent. Even though in our exercise the optimal capital income tax is relatively low, it distorts the saving margin more than a progressive tax on joined income in the benchmark. The drop in aggregate capital is an order of magnitude smaller than in CKK, where the optimal capital tax of 36 percent causes aggregate capital to decline by more than 6 percent.

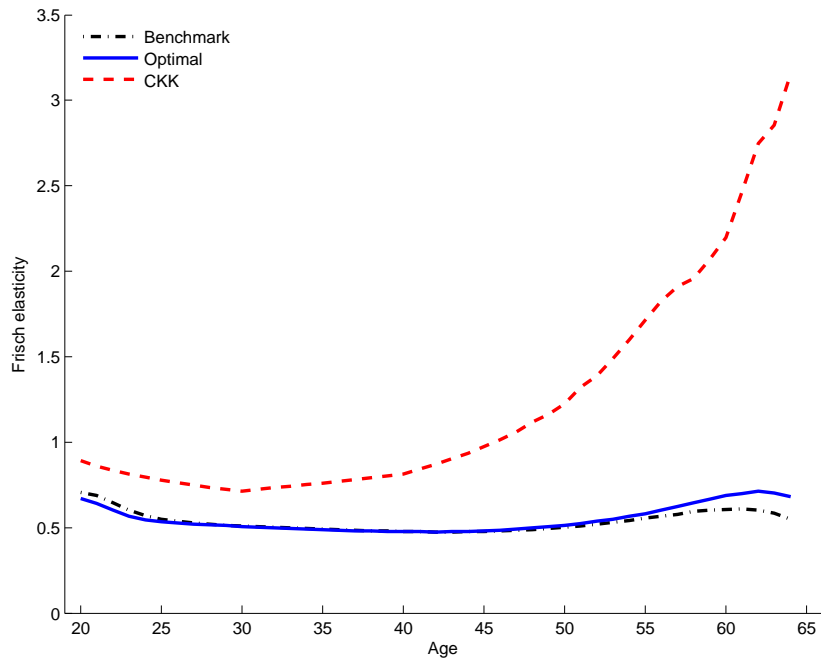
5.2 Why is the optimal capital income tax low?

In a life-cycle model, a household's consumption-labor plan varies over its life-time. It comes as no surprise that a social planner varies tax policy over the life-cycle to maximize welfare as well. Erosa and Gervais (2002) show that optimally a planner would like to condition a labor income tax on age. This age-dependence arises from the life-cycle component of an exogenous labor productivity process, the average of which depends on age. Productivity has a hump shape; it increases up to age 55 and then drops until retirement. The optimal age-dependent labor income tax is decreasing with age. A planner wants to make leisure relatively more expensive to highly productive households by lowering labor income tax and thus increasing their labor supply. Peterma (2010) shows that the magnitude of the labor income tax variation in age depends on the Frisch elasticity of labor Céspedes and Rendon (2012). For a constant elasticity of labor the optimal labor income tax rate is age invariant. If the Frisch elasticity is non-constant and depends on hours worked (equation 29), then an optimal labor income tax decreases as the household works less. In other words, the labor income tax rate decreases with age when the Frisch labor elasticity rises over a household's working life. In fact, the smaller the difference

in hours worked over time, the less the optimal labor tax varies with age.

In the U.S. age discrimination is not permitted within a tax system. Erosa and Gervais (2002) show that a planner can use the tax on capital income to mimic the incentive created by the age-dependent labor income tax. The capital income tax has an intertemporal effect on the implicit price of leisure. By taxing today's capital income the planner is effectively taxing next period's consumption, which is a complementary good with next period leisure. This makes next period's leisure relatively more expensive and creates the incentive to work more. So the capital income tax can imperfectly substitute for a time-varying labor income tax.

Figure 7: Frisch elasticity at the intensive margin



Therefore, the optimal capital income tax depends on the Frisch elasticity of labor supply. Facing a high elasticity of labor, a planner wants to decrease the labor income tax burden with age. Since a tax cannot be age-dependent, the planner can use a capital income tax to mimic a decreasing labor tax. If the labor tax rate drops steeply with age, the planner needs to resort to a high capital income tax.

Figure 7 depicts the Frisch elasticity of labor supply at the intensive margin. Our benchmark and optimal tax models use a separable utility function (eq. 24) with a corresponding Frisch elasticity of labor represented by equation 29. The CKK model uses a Cobb-Douglas utility function with a share of consumption $\gamma = 4.4$ and relative risk aversion $\sigma = 4$. The Frisch elasticity for Cobb-Douglas utility functions is:

$$f^{CD} = \frac{(1-h)[1-\gamma(1-\sigma)]}{\sigma h}. \quad (30)$$

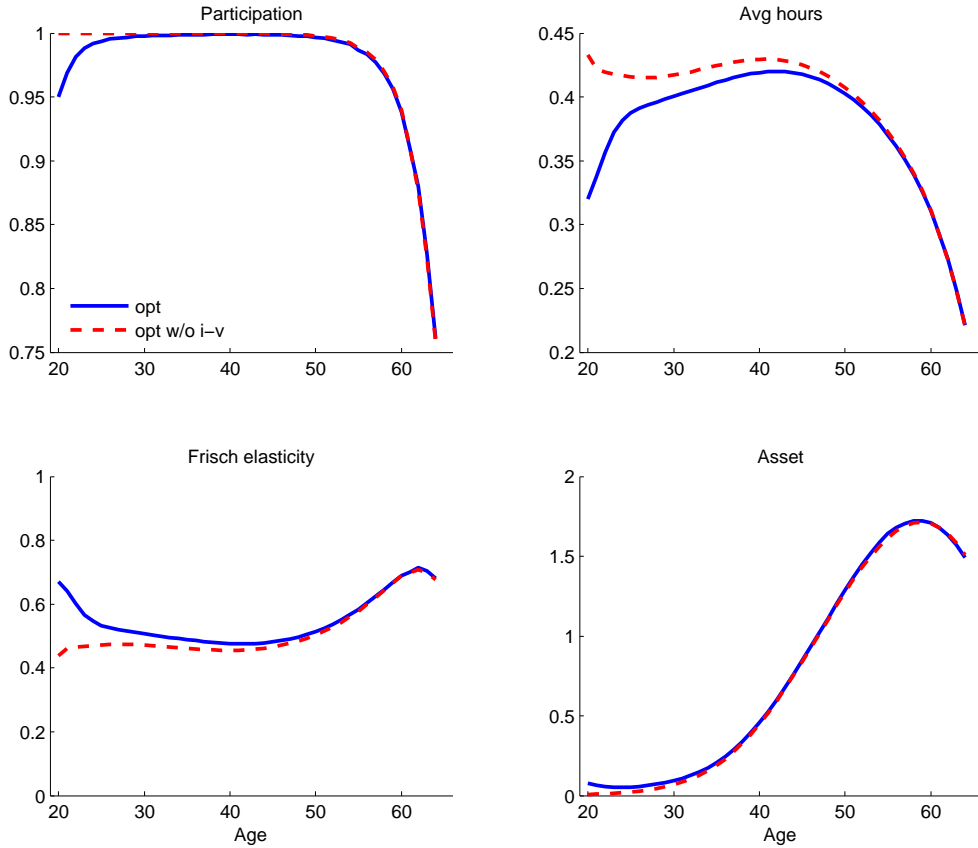
In the CKK model the elasticity of labor supply increases exponentially after age 30, while in our model it stays roughly flat over the life cycle. The maximum deviation from the mean elasticity is over 100 percent in CKK and less than 25 percent in our model. This striking result is a consequence of different representations of life-cycle hours worked in each model. From the planner's perspective, a steep growth in elasticity over the life cycle requires providing large incentives for older, more productive households to work. It can be achieved by either lowering the labor income tax or imposing a high capital income tax. Both methods make leisure relatively more expensive, thus stimulating households to work more. In our model and the U.S. data, households work full time for a majority of their working lives. Their hours of work and participation are persistent over the life cycle. This corresponds to a low and persistent Frisch elasticity of labor. The planner therefore does not need to create incentives for the older, most productive households to work. Thus, there is no need to excessively vary the labor income tax over time or to impose a high capital income tax.

5.3 Effect of inter-vivos transfers

To find out the effect of the inter-vivos transfer on the optimal taxation scheme, we shut down the transfers and look for the new optimal tax system. Without transfers, the parameters of the optimal labor income tax are $a_0 = 0.37$, $a_1 = 4.2$, and $a_2 = 51352$. The optimal capital tax rate, τ_k , is 6.8 percent; this is 0.6 percentage point lower than in the benchmark. The inter-vivos transfers have a small impact on the optimal tax system. In the optimal tax system without the inter-vivo transfers, the tax burden is slightly shifted from the capital income to the labor income tax.

The four panels in Figure (8) present the participation rate, average positive hours worked, Frisch elasticity, and asset holdings profiles for the optimal model with and without the transfers. In the optimal tax system with transfers, only 95 percent of households work at age 20 and participation increases with age, reaching 100 percent at age 27. Without the inter-vivos transfers, the participation of young households in the labor market increases to 100 percent. At the intensive margin, households below age 50 work significantly more. On average the 20-year-old households work 520 hours a year more than in the optimal tax system with the transfers. The impact of the inter-vivos transfer on hours worked diminishes with age as households rely more on their own income the older they get. Notice that without the transfers, average positive hours of work up to age 50 are roughly flat around 42 percent of the available time and are more persistent. In this age group, the Frisch elasticity of labor drops due

Figure 8: Effect of the inter-vivos transfers on the optimal tax system



to the increase in hours worked. Asset holdings drop among households up to age 35; the impact of the transfers on the asset holdings of older households is very small.

Without the inter-vivos transfer, households up to age 50 work more. In this age group the average positive hours reach the calibrated upper bound of 42 percent of available time. Thus, eliminating transfers favors the planner's policy to encourage highly productive households to work more, especially households who are close to their productivity peak at age 47. The need for a planner to use incentives for households to work more diminishes. Thus, the optimal capital tax rate can be lowered by 0.6 percent. Since a planner's policy targets the older, highly productive households, we can associate a relatively small decrease in the capital income tax to a relatively small increase in hours worked among households in their 30s and 40s. The households up to age 30 have relatively low productivity and thus a small impact on the planner's tax policy. Thus, the cost of creating an incentive for young households to work outweighs the benefit of the extra output they produce. From this exercise we conclude that the low optimal capital income tax in our model is driven by the labor supply of older households. The mechanism that drives the high participation and long hours of work in this part of the labor supply

profile is a non-linear wage schedule. The inter-vivos transfers and labor decisions of young households have a relatively small impact on the optimal tax policy.

6 Conclusion

In a canonical life-cycle model with heterogeneous agents the optimal capital income tax is as high as 36 percent, as reported by CKK. This is a striking result in light of other analysis, most prominently Chari and Kehoe (1999a), who argue that the optimal capital income tax is zero. Even though the endogenous labor supply drives the results in CKK, their model does not capture key life-cycle features of labor supply.

In this paper we develop a life-cycle model with a realistic labor supply profile, which we employ to analyze the optimal tax system with capital and labor income taxes. The key features of the model are non-linear wages and inter-vivos transfers. The former makes hours of work highly persistent and helps to account for labor choices at the extensive margin over the life cycle. The latter allows us to account for labor choices in early life.

We find that the optimal capital income tax is 7.4 percent, which is significantly lower than what CKK report. The corresponding labor income tax is slightly progressive with an upper bound of 36 percent. The result is driven by the high persistence of hours worked among households over age 40. Finally, we show that inter-vivos transfers have a relatively small impact on the optimal capital income tax. Without the transfers, the optimal capital income tax would be lower by 0.6 percentage point.

References

- AARONSON, D., AND E. FRENCH (2004): “The Effect of Part-Time Work on Wages: Evidence from the Social Security Rules,” *Journal of Labor Economics*, 22(2), 329–352.
- ALTONJI, J. G., F. HAYASHI, AND L. KOTLIKOFF (1996): “The Effects of Income and Wealth on Time and Money Transfers between Parents and Children,” NBER Working Papers 5522, National Bureau of Economic Research, Inc.
- BELL, F. C., AND M. L. MILLER (2002): *Life tables for the United States Social Security area, 1900-2100*. Social Security Administration, Office of the Chief Actuary.
- CHAMLEY, C. (1986): “Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives,” *Econometrica*, 54(3), 607–22.
- CHARI, V., AND P. J. KEHOE (1999a): “Chapter 26 Optimal fiscal and monetary policy,” vol. 1, Part 3 of *Handbook of Macroeconomics*, pp. 1671 – 1745. Elsevier.
- (1999b): “Optimal fiscal and monetary policy,” in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and M. Woodford, vol. 1 of *Handbook of Macroeconomics*, chap. 26, pp. 1671–1745. Elsevier.
- CONESA, J. C., S. KITAO, AND D. KRUEGER (2009): “Taxing Capital? Not a Bad Idea after All,” *American Economic Review*, 99(1), 25–48.
- CÉSPÉDES, N., AND S. RENDON (2012): “The Frisch Elasticity in Labor Markets with High Job Turnover,” *IZA Discussion Papers*, (6991).
- EROSA, A., L. FUSTER, AND G. KAMBOUROV (2010): “The Heterogeneity and Dynamics of Individual Labor Supply over the Life-Cycle: Facts and Theory,” Unpublished manuscript.
- EROSA, A., AND M. GERVAIS (2002): “Optimal Taxation in Life-Cycle Economies,” *Journal of Economic Theory*, 105(2), 338–369.
- FITZGERALD, T. J. (1998): “Reducing working hours: a general equilibrium analysis,” Working Paper 9801, Federal Reserve Bank of Cleveland.
- GALE, W. G., AND J. K. SCHOLZ (1994): “Intergenerational Transfers and the Accumulation of Wealth,” *Journal of Economic Perspectives*, 8(4), 145–60.
- GOUVEIA, M., AND R. P. STRAUSS (1994): “Effective Federal Individual Tax Functions: an Exploratory Empirical Analysis,” *National Tax Journal* 317-339 (June 1994), 47, 317–339.

- HANSEN, G. D. (1985): “Indivisible labor and the business cycle,” *Journal of Monetary Economics*, 16(3), 309–327.
- HORNSTEIN, A., AND E. C. PRESCOTT (1993): “The Firm and the Plant in General Equilibrium Theory,” in *General equilibrium, growth, and trade II: essays in honor of Lionel McKenzie*, ed. by R. Becker, M. Boldrin, R. Jones, and W. Thomson, Economic theory, econometrics, and mathematical economics. Academic Press.
- JUDD, K. L. (1985): “Redistributive taxation in a simple perfect foresight model,” *Journal of Public Economics*, 28(1), 59–83.
- MENDOZA, E. G., A. RAZIN, AND L. L. TESAR (1994): “Effective tax rates in macroeconomics: Cross-country estimates of tax rates on factor incomes and consumption,” *Journal of Monetary Economics*, 34(3), 297–323.
- OSUNA, V., AND J.-V. RIOS-RULL (2003): “Implementing the 35 Hour Workweek by Means of Overtime Taxation,” *Review of Economic Dynamics*, 6(1), 179–206.
- PETERMA, W. B. (2010): “Determining the Motives for a Positive Optimal Tax on Capital,” Unpublished manuscript.
- PRESCOTT, E. C., AND J.-V. RIOS-RULL (1992): “Classical competitive analysis of economies with Islands,” *Journal of Economic Theory*, 57(1), 73 – 98.
- PRESCOTT, E. C., AND R. M. TOWNSEND (1984): “Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard,” *Econometrica*, 52(1), 21–45.
- ROGERSON, R. (1988): “Indivisible labor, lotteries and equilibrium,” *Journal of Monetary Economics*, 21(1), 3–16.
- STIGLITZ, J. E. (1987): “Pareto efficient and optimal taxation and the new new welfare economics,” in *Handbook of Public Economics*, ed. by A. J. Auerbach, and M. Feldstein, vol. 2 of *Handbook of Public Economics*, chap. 15, pp. 991–1042. Elsevier.
- STORESLETTEN, K., C. I. TELMER, AND A. YARON (2004): “Consumption and risk sharing over the life cycle,” *Journal of Monetary Economics*, 51(3), 609–633.