Optimal Progressivity with Age-Dependent Taxation

Jonathan Heathcote

Federal Reserve Bank of Minneapolis

Kjetil Storesletten

University of Oslo

Gianluca Violante

Princeton University

Harvard University

How progressive should labor income taxation be?

- Arguments against progressivity: distortions
  - Labor supply choice
  - Human capital investment

How progressive should labor income taxation be?

- Arguments against progressivity: distortions
  - Labor supply choice
  - Human capital investment
- Arguments in favor of progressivity: missing markets
  - Unequal initial conditions
  - Labor market shocks
  - Increasing age-productivity profile

How progressive should labor income taxation be?

- Arguments against progressivity: distortions
  - Labor supply choice
  - Human capital investment
- Arguments in favor of progressivity: missing markets
  - Unequal initial conditions
  - Labor market shocks
  - Increasing age-productivity profile
- Q: Tagging $\rightarrow$ should optimal progressivity vary with age?

This paper

- OLG equilibrium model with:
  - × flexible labor supply [static choice]
  - × skill investment [dynamic choice]
  - ✓ differential disutility of work & learning ability [ex-ante heter.]
  - ✓ partial insurance against wage risk [ex-post uncertainty]
  - ✓ age profile for productivity and disutility of work [life cycle]

This paper

- OLG equilibrium model with:
  - × flexible labor supply [static choice]
  - × skill investment [dynamic choice]
  - ✓ differential disutility of work & learning ability [ex-ante heter.]
  - ✓ partial insurance against wage risk [ex-post uncertainty]
  - ✓ age profile for productivity and disutility of work [life cycle]
- Baseline: analytical model to isolate forces at work
- Extension: numerically solved model with borrowing and saving

TAX FUNCTION

Tax Function

$$T(y) = y - \lambda y^{1-\tau}$$

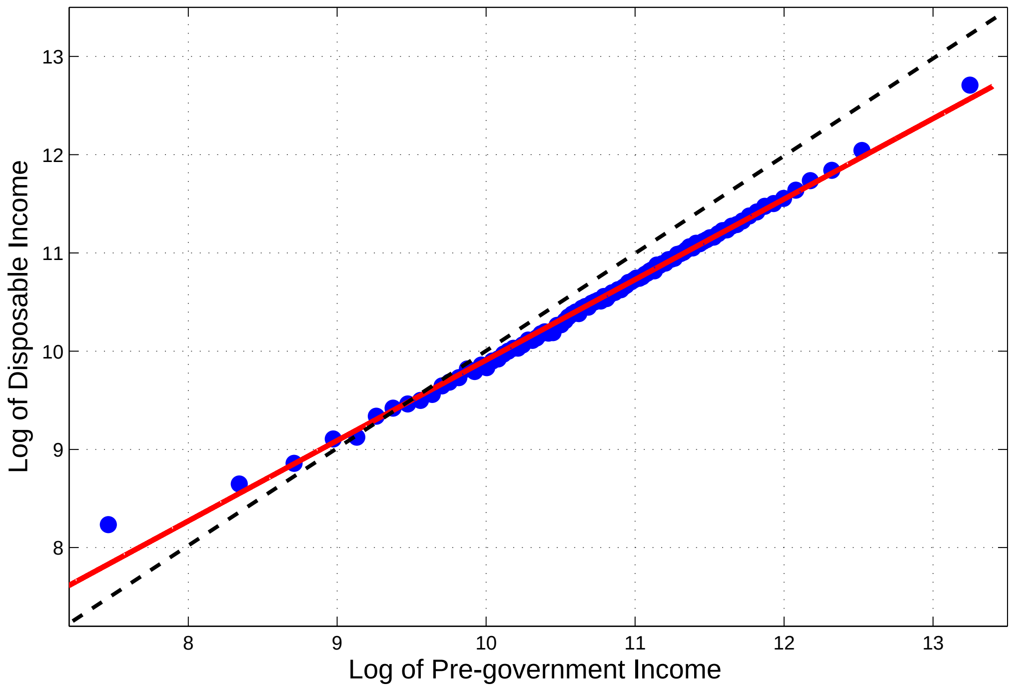
Tax Function

$$\log(y - T(y)) = \log \lambda + (1 - \tau) \log y$$

Tax Function

$$\log(y - T(y)) = \log \lambda + (1 - \tau) \log y$$

- It preserves analytical tractability
- It closely approximates U.S. tax/transfer system ($\tau^{US} = 0.181$)



Generalized Tax Function

- We generalize tax/transfer system to allow for age variation:

$$T_a(y) = y - \lambda_a y^{1-\tau_a}$$

Generalized Tax Function

- We generalize tax/transfer system to allow for age variation:

$$T_a(y) = y - \lambda_a y^{1-\tau_a}$$

- Does the US tax/transfer system display age dependence?

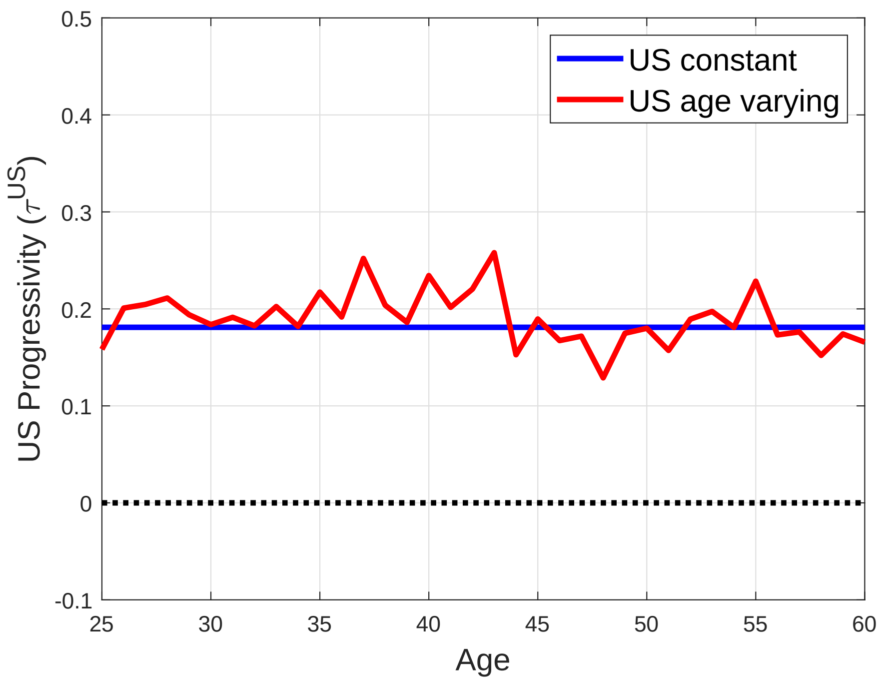
Generalized Tax Function

- We generalize tax/transfer system to allow for age variation:

$$T_a(y) = y - \lambda_a y^{1-\tau_a}$$

- Does the US tax/transfer system display age dependence?

- Estimate $\{\tau_a\}$ by household age



Related Literature

- Human capital: Best and Kleven (2013), Guvenen, Kuruscu, and Ozkan (2014), Kapicka and Neira (2016), Stantcheva (2017)

- Labor supply: Erosa and Gervais (2002), Karabarbounis (2016), Ndiaye (2017)

- Efficiency profile: Weinzierl (2009), Gorry and Oberfield (2012)

- Uninsurable risk: Farhi and Werning (2013), Golosov, Troshkin, and Tsyvinski (2016)

HSV: Transparency + GE + Transition + Quantitative

Environment

Preferences

- Preferences over consumption (c), hours (h), publicly-provided goods (G), and skill-investment (s) effort:

$$U_i = -v_i(s_i) + \mathbb{E}_0 \sum_{a=0}^{A} \beta^a u_i(c_{ia}, h_{ia}, G)$$

Preferences

- Preferences over consumption (c), hours (h), publicly-provided goods (G), and skill-investment (s) effort:

$$U_i = -v_i(s_i) + \mathbb{E}_0 \sum_{a=0}^{A} \beta^a u_i(c_{ia}, h_{ia}, G)$$

$$v_i(s_i) = \frac{1}{(\kappa_i)^{1/\psi}} \cdot \frac{s_i^{1+1/\psi}}{1+1/\psi}$$

$$\kappa_i \sim Exp\,(1)$$

Preferences

- Preferences over consumption (c), hours (h), publicly-provided goods (G), and skill-investment (s) effort:

$$U_i = -v_i(s_i) + \mathbb{E}_0 \sum_{a=0}^{A} \beta^a u_i(c_{ia}, h_{ia}, G)$$

$$v_i(s_i) \quad = \quad \frac{1}{(\kappa_i)^{1/\psi}} \cdot \frac{s_i^{1+1/\psi}}{1+1/\psi}$$

$$\kappa_i \quad \sim \quad Exp\,(1)$$

$$u_i\left(c_{ia}, h_{ia}, G\right) \quad = \quad \log c_{ia} - \frac{\exp\left[(1+\sigma)\left(\varphi_i + \bar{\varphi}_a\right)\right]}{1+\sigma}\left(h_{ia}\right)^{1+\sigma} + \chi \log G$$

$$\varphi_i \quad \sim \quad \mathcal{N}\left(\frac{v_\varphi}{2}, v_\varphi\right)$$

Technology

- Output is a CES aggregator over continuum of skill types s:

$$Y = \left[\int_0^\infty N(s)^{\frac{\theta-1}{\theta}} \, ds\right]^{\frac{\theta}{\theta-1}}, \quad \theta \in [1, \infty)$$

  - $N(s)$: effective hours of type s

- Aggregate resource constraint:

$$Y = \sum_{a=0}^{A} \int_{i=0}^{1} c_{i,a} \, di + G$$

  - WLOG: $G = gY$

Individual Wages and Earnings

- Hourly wages:

$$\log w_{ia} = \log p(s_i) + x_a + \alpha_{ia} + \varepsilon_{ia}$$

  - $p(s_i)$: skill price = marginal product of labor of type s
  - x_a: deterministic age-productivity profile
  - $\alpha_{ia} = \alpha_{i,a-1} + \omega_{ia}, \quad \omega_{ia} \sim \mathcal{N}\left(-\frac{v_\omega}{2}, v_\omega\right)$ [uninsurable]
  - $\varepsilon_{ia} \overset{iid}{\sim} \mathcal{N}\left(-\frac{v_{\varepsilon a}}{2}, v_{\varepsilon a}\right)$ [privately insurable]

- Gross earnings:

$$y_{ia} = \underbrace{p(s_i)}_{\text{skill investment}} \times \underbrace{\exp(x_a)}_{\text{life-cycle}} \times \underbrace{\exp(\alpha_{ia} + \varepsilon_{ia})}_{\text{shocks}} \times \underbrace{h_{ia}}_{\text{labor supply}}$$

Government

- Government budget constraint (no government debt):

$$gY = \sum_{a=0}^{A} \int_0^1 \underbrace{\left[y_i - \lambda_a y_i^{1-\tau_a}\right]}_{T_a(y_i)} di$$

- Government chooses vector $\{\lambda_a^*, \tau_a^*\}_{a=0}^{A}$ and g^*

Government

- Government budget constraint (no government debt):

$$gY = \sum_{a=0}^{A} \int_0^1 \underbrace{\left[y_i - \lambda_a y_i^{1-\tau_a}\right]}_{T_a(y_i)} di$$

- Government chooses vector $\{\lambda_a^*, \tau_a^*\}_{a=0}^{A}$ and g^*
  - Optimal public good provision: $g^* = \frac{\chi}{1+\chi}$
  - Samuelson condition: $MRS_{C,G} = MRT_{C,G} = 1$

EQUILIBRIUM ALLOCATIONS

Skill Prices and Skill Investment

- Skill price has the Mincerian form:

$$\log p(s) = \pi_0 + \pi_1 s(\kappa; \bar{\tau})$$

- Closed form expressions for equilibrium π_0 and π_1

- Optimal skill investment is linear in κ:

$$s\left(\kappa; \bar{\tau}\right) = \left[\left(1 - \bar{\tau}\right)\pi_1\right]^{\psi} \cdot \kappa$$

where: $\bar{\tau} = \frac{1-\beta}{1-\beta^{A+1}} \sum_{a=0}^{A} \beta^a \tau_a$

- Distribution of $p(s)$ is Pareto with parameter θ

Consumption and Hours

$$\log c_a \quad = \quad \log \lambda_a + (1-\tau_a)\left[\frac{\log(1-\tau_a)}{1+\sigma} - (\varphi + \bar{\varphi}_a) + \log p(s) + x_a + \alpha\right] + \mathcal{C}_a$$

- Progressivity determines the pass-through of shocks/inequality

Consumption and Hours

$$\log c_a = \log \lambda_a + (1-\tau_a)\left[\frac{\log(1-\tau_a)}{1+\sigma} - (\varphi + \bar{\varphi}_a) + \log p(s) + x_a + \alpha\right] + \mathcal{C}_a$$

- Progressivity determines the pass-through of shocks/inequality

$$\log h_a = \frac{\log(1-\tau_a)}{1+\sigma} - (\varphi + \bar{\varphi}_a) + \left(\frac{1-\tau_a}{\sigma+\tau_a}\right)\varepsilon - \mathcal{H}_a$$

- Log-utility $\rightarrow$ hours unaffected by $\{\lambda_a, p(s), x_a, \alpha\}$

Consumption and Hours

$$\log c_a = \log \lambda_a + (1-\tau_a)\left[\frac{\log(1-\tau_a)}{1+\sigma} - (\varphi + \bar{\varphi}_a) + \log p(s) + x_a + \alpha\right] + \mathcal{C}_a$$

- Progressivity determines the pass-through of shocks/inequality

$$\log h_a = \frac{\log(1-\tau_a)}{1+\sigma} - (\varphi + \bar{\varphi}_a) + \left(\frac{1-\tau_a}{\sigma+\tau_a}\right)\varepsilon - \mathcal{H}_a$$

- Log-utility $\rightarrow$ hours unaffected by $\{\lambda_a, p(s), x_a, \alpha\}$
- Note: insurable productivity shocks enters h but not c

Social Welfare

Social Welfare Function

- Utilitarian: equal weight on welfare of all currently alive agents, discounts welfare of future cohorts at rate β
- $\beta = 1$: SWF equals steady-state welfare
- $\beta < 1$: SWF embeds transition as planner cares for past cohorts
  - Transition driven by irreversible skill choice of past cohorts
  - Allow $\{\lambda_{a,t}\}$, $\{\tau_{a,t}\}$, g_t to vary freely by age and time
  - Initial condition: steady-state under τ^{US}
- Feasible to optimize over large vector of policy parameters because social welfare has a closed-form

STEADY-STATE ANALYSIS

Social Welfare Function $(\beta = 1)$

$$
\begin{aligned}
\mathcal{W}^{ss}(\{\tau_a\}) \quad = \quad & -\frac{1}{A}\sum_{a=0}^{A-1} \underbrace{\frac{1-\tau_a}{1+\sigma}}_{\text{Disutility of labor}} \\
+ \quad & (1+\chi)\log \underbrace{\left[\sum_{a=0}^{A-1}(1-\tau_a)^{\frac{1}{1+\sigma}} \cdot \exp\left(x_a - \bar{\varphi}_a\right) + \left(\frac{\tau_a\left(1+\hat{\sigma}_a\right)}{\hat{\sigma}_a^2} + \frac{1}{\hat{\sigma}_a}\right)\frac{v_{\varepsilon a}}{2}\right]}_{\text{Gain from labor supply: effective hours } N} \\
+ \quad & (1+\chi)\underbrace{\frac{1}{(1+\psi)(\theta-1)}\left[\psi\log\left(1-\bar{\tau}\right) + \log\left(\frac{1}{\eta\theta^{\psi}}\left(\frac{\theta}{\theta-1}\right)^{\theta(1+\psi)}\right)\right]}_{\text{Gain from skill investment: productivity: } \log(E[p(s)])} \\
- \quad & \underbrace{\frac{\psi}{1+\psi}\frac{1-\bar{\tau}}{\theta}}_{\text{Avg. skill inv. cost}} + \frac{1}{A}\sum_{a=0}^{A-1}\underbrace{\left[\log\left(1-\left(\frac{1-\tau_a}{\theta}\right)\right) + \left(\frac{1-\tau_a}{\theta}\right)\right]}_{\text{Cost of consumption dispersion across skills}} \\
- \quad & \frac{1}{A}\sum_{a=0}^{A-1}\underbrace{(1-\tau_a)^2\left(\frac{v_{\varphi}}{2} + a\frac{v_{\omega}}{2}\right)}_{\text{Cons. dispersion due to unins. risk and pref. heter.}}
\end{aligned}
$$

Optimal Policy: Theoretical results for $\beta = 1$

1. Optimal $\{\tau_a^*, \lambda_a^*\}$ are age-invariant if:
   (a) $v_\omega = 0$: flat profile of uninsurable productivity dispersion
   (b) $v_{\varepsilon a} = v_\varepsilon$: flat profile of insurable productivity dispersion
   (c) $\{x_a - \bar{\varphi}_a\}$ constant: flat profile of efficiency net of disutility

Optimal Policy: Theoretical results for $\beta = 1$

1. Optimal $\{\tau_a^*, \lambda_a^*\}$ are age-invariant if:

   (a) $v_\omega = 0$: flat profile of uninsurable productivity dispersion

   (b) $v_{\varepsilon a} = v_\varepsilon$: flat profile of insurable productivity dispersion

   (c) $\{x_a - \bar{\varphi}_a\}$ constant: flat profile of efficiency net of disutility

2. If, in addition, $\theta = \infty$ and $v_\varphi = 0$, the economy $\rightarrow$ RA and $\tau_a^* = -\chi$

   - Regressivity corrects the externality linked to valued G

Optimal Policy: Theoretical results for $\beta = 1$

1. Optimal $\{\tau_a^*, \lambda_a^*\}$ are age-invariant if:
   (a) $v_\omega = 0$: flat profile of uninsurable productivity dispersion
   (b) $v_{\varepsilon a} = v_\varepsilon$: flat profile of insurable productivity dispersion
   (c) $\{x_a - \bar{\varphi}_a\}$ constant: flat profile of efficiency net of disutility

2. If, in addition, $\theta = \infty$ and $v_\varphi = 0$, the economy $\rightarrow$ RA and $\tau_a^* = -\chi$
   - Regressivity corrects the externality linked to valued G

3. Given any profile for $\{\tau_a\}$, the optimal profile for $\{\lambda_a^*\}$ equates average consumption (i.e., the MUC_a) by age

Determinants of age profile of progressivity $(\beta = 1)$

(a) Uninsurable Risk channel
Permanent uninsurable risk $(v_\omega > 0)$ *implies that* $\{\tau_a^*\}$ *is increasing in age*

(b) Insurable Risk channel
Starting from $\tau_a > 0$, *rising insurable risk* $(v_{\varepsilon,a+1} > v_{\varepsilon,a})$ *implies that* $\tau_{a+1}^* < \tau_a^*$

(c) Life-Cycle channel
Age profile in $\{x_a - \bar{\varphi}_a\}$ *implies* $\{\tau_a^*\}$ *which is its mirror image*

- The optimal $\{\tau_a^*\}$ equates the labor wedge, $1 - MTR_a$, by age

$$1 - MTR_a = \lambda_a (1 - \tau_a) y_a^{-\tau_a} = 1$$

- It implements the first best

PARAMETERIZATION

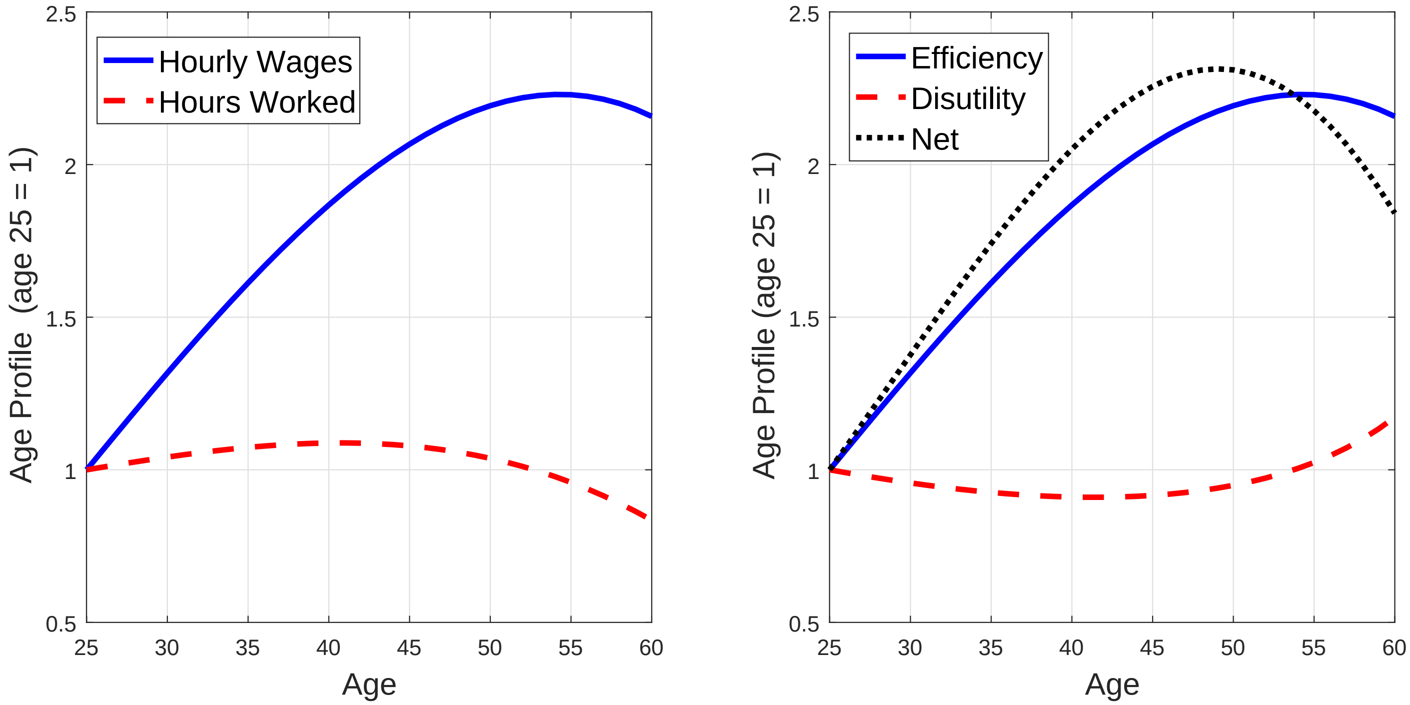
Parameterization

- Parameters: $\{\tau^{US}, \chi, \sigma, \psi, \theta, v_\varphi, v_\omega, v_{\varepsilon 0}, v_\eta\}$ and $\{x_a, \bar{\varphi}_a\}_{a=1}^{A}$
- US progressivity estimated on micro data $\rightarrow \tau^{US} = 0.181$
- Assume observed $G/Y = 0.19 = g^*$ $\rightarrow \chi = 0.233$
- Frisch elasticity (micro-evidence ~ 0.5) $\rightarrow \sigma = 2$
- Price-elasticity of skill investment $\rightarrow \psi = 0.65$

$$
\begin{aligned}
var_0(\log c) &\quad\rightarrow\quad \theta = 3.12 \\
var(\log h) &\quad\rightarrow\quad v_\varphi = 0.035 \\
cov(\log w, \log c) &\quad\rightarrow\quad v_\omega = 0.0058 \\
cov(\log w, \log h) &\quad\rightarrow\quad v_{\varepsilon,0} = 0.09, \Delta v_{\varepsilon,a} = 0.0044
\end{aligned}
$$

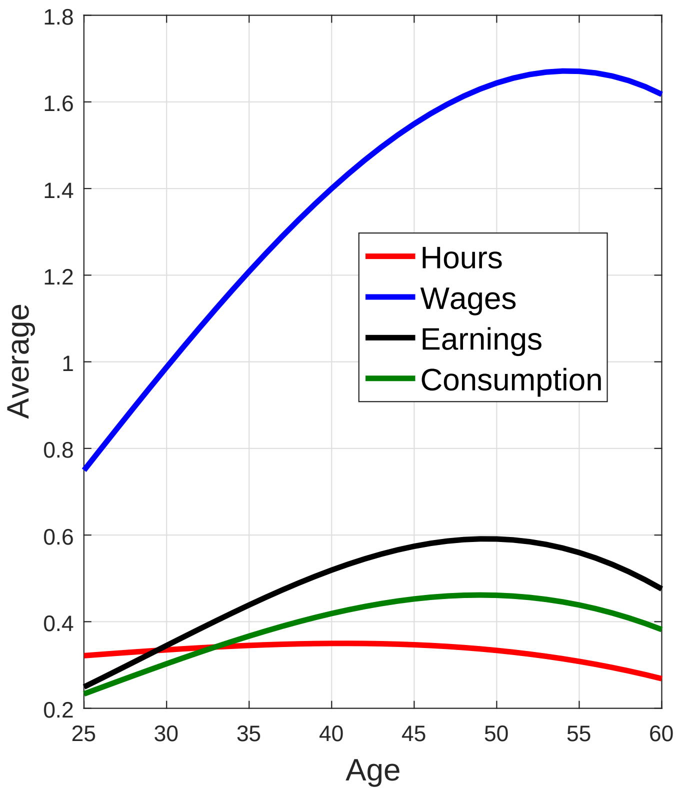
- $\{x_a, \bar{\varphi}_a\}_{a=1}^{A}$ estimated to match age profiles wages of and hours

Age Profile for Efficiency and Disutility of Work

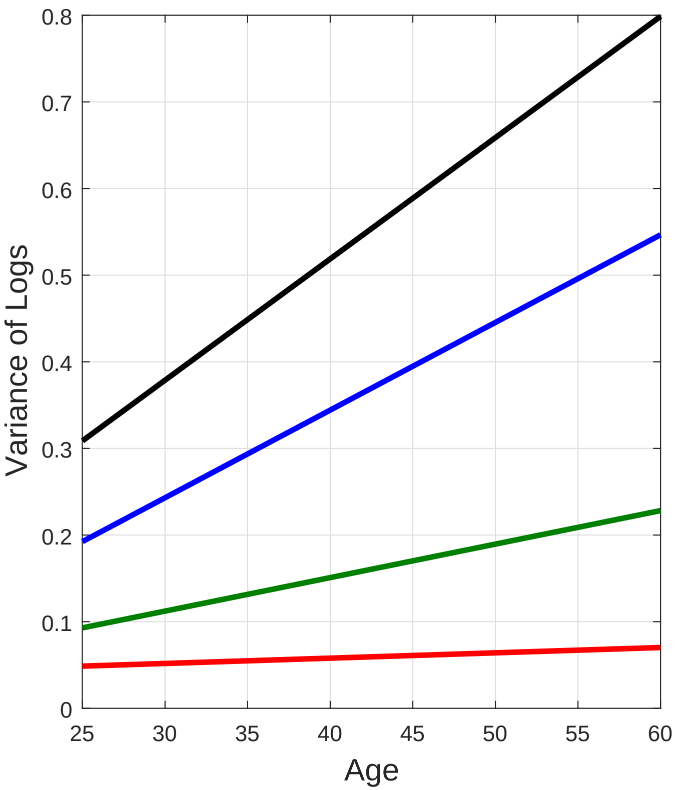


- Important: $\{x_a - \bar{\varphi}_a\}$ is hump-shaped

Life-cycle Means and Variances



Means

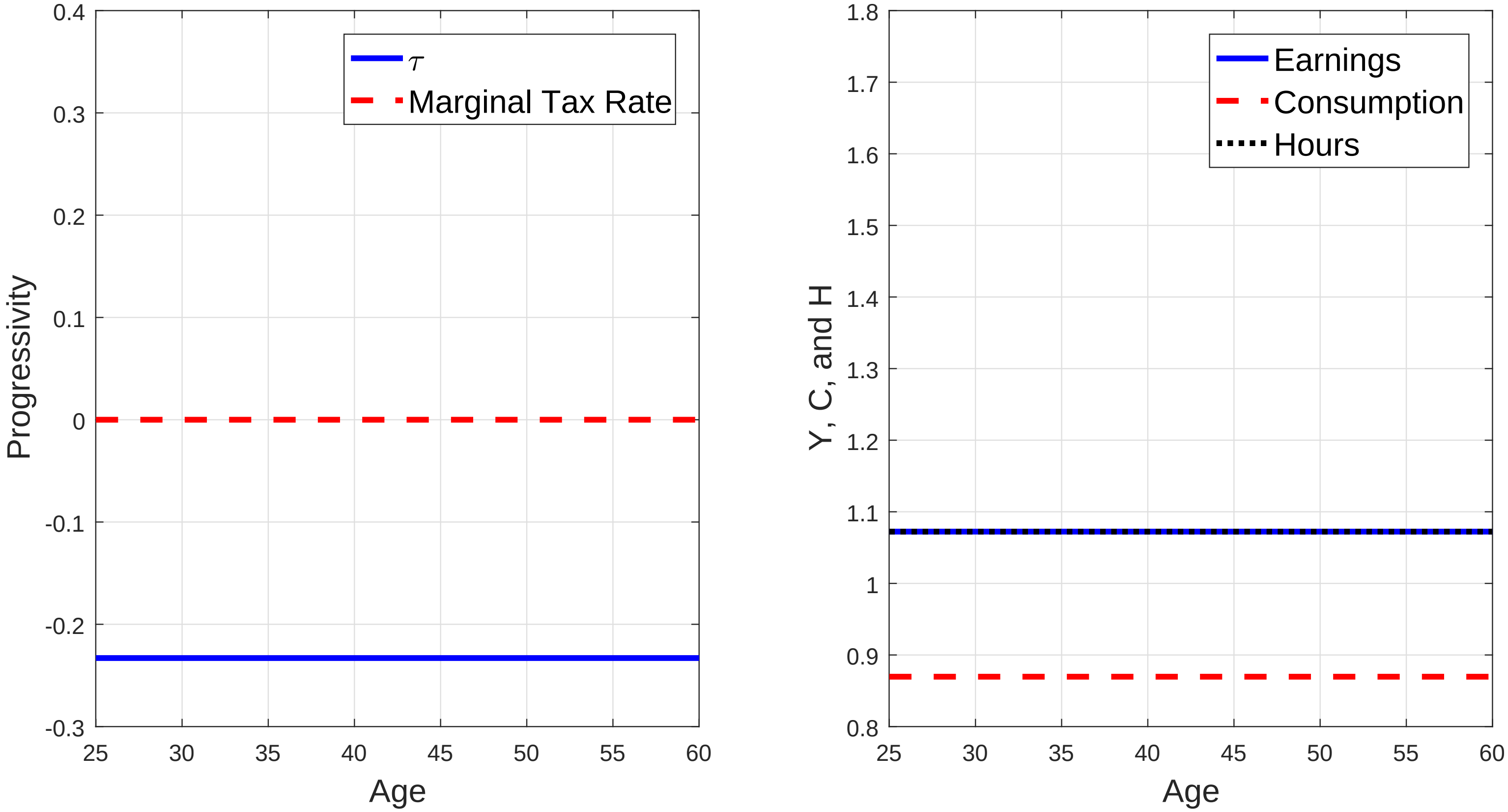


Variances of Logs

Quantitative Results

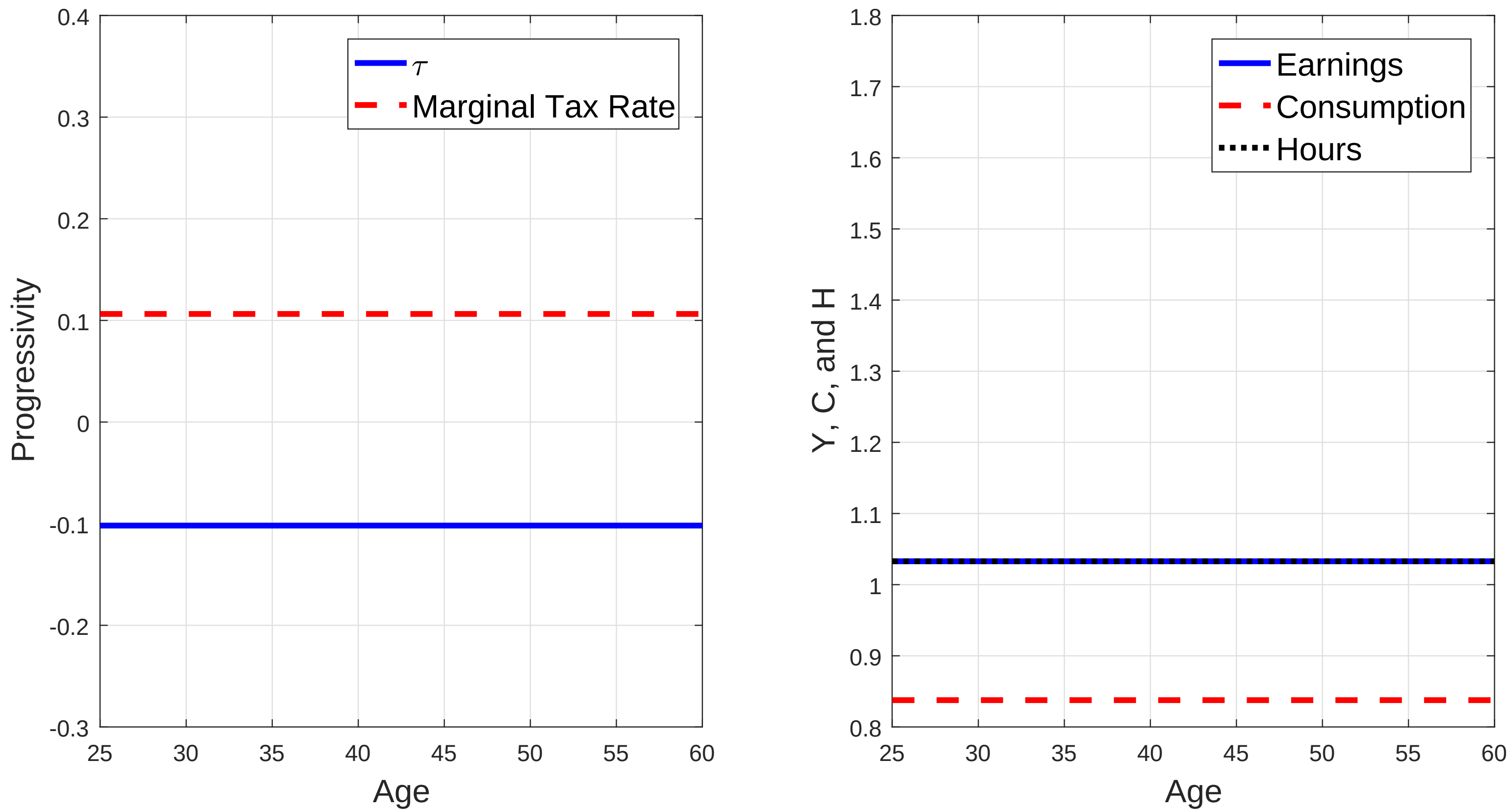
$$\beta = 1$$

Representative Agent



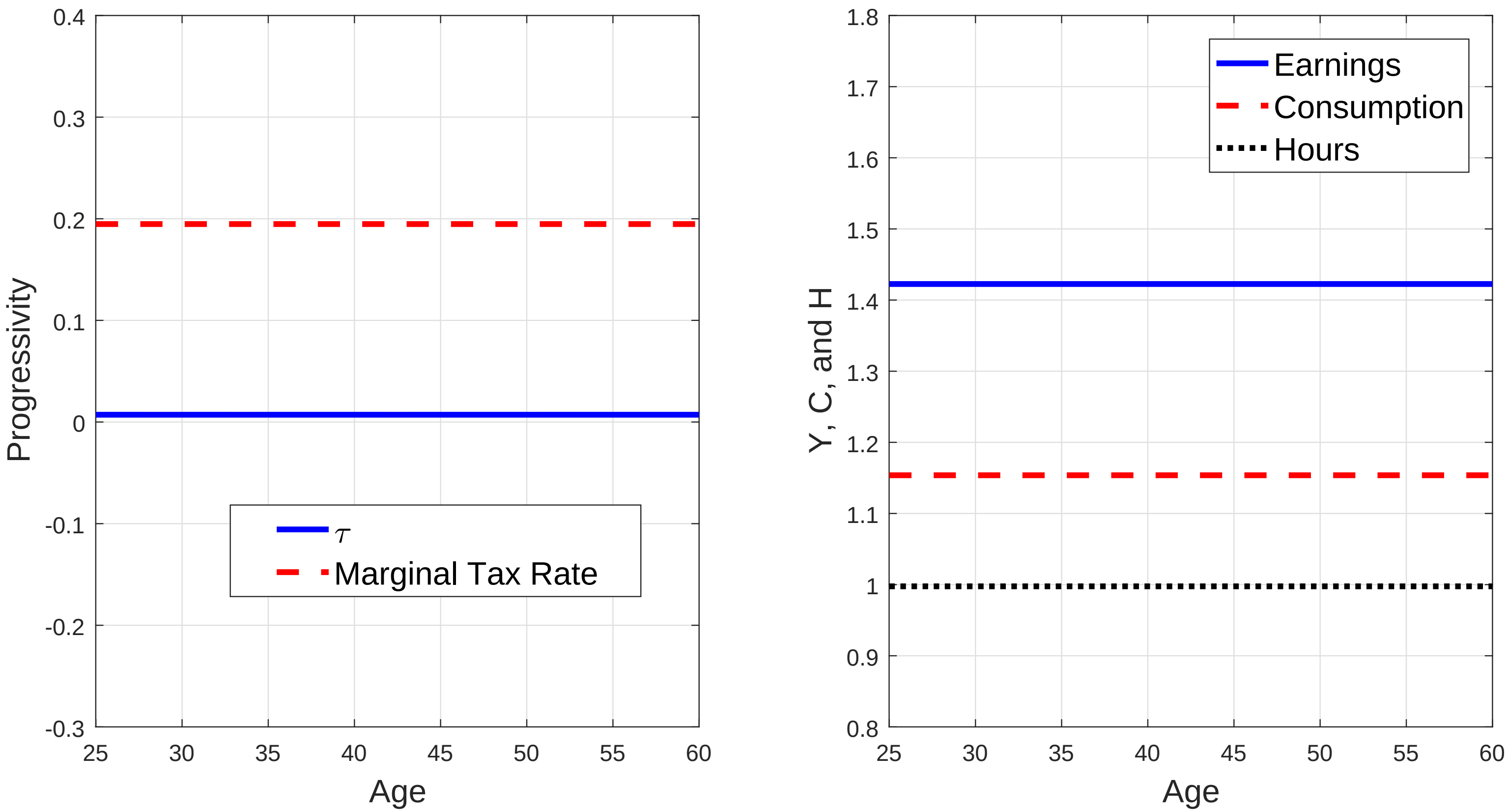
- Optimality: $\tau_a^* = -\chi$

Add Heterogeneity in Disutility of Work (φ)



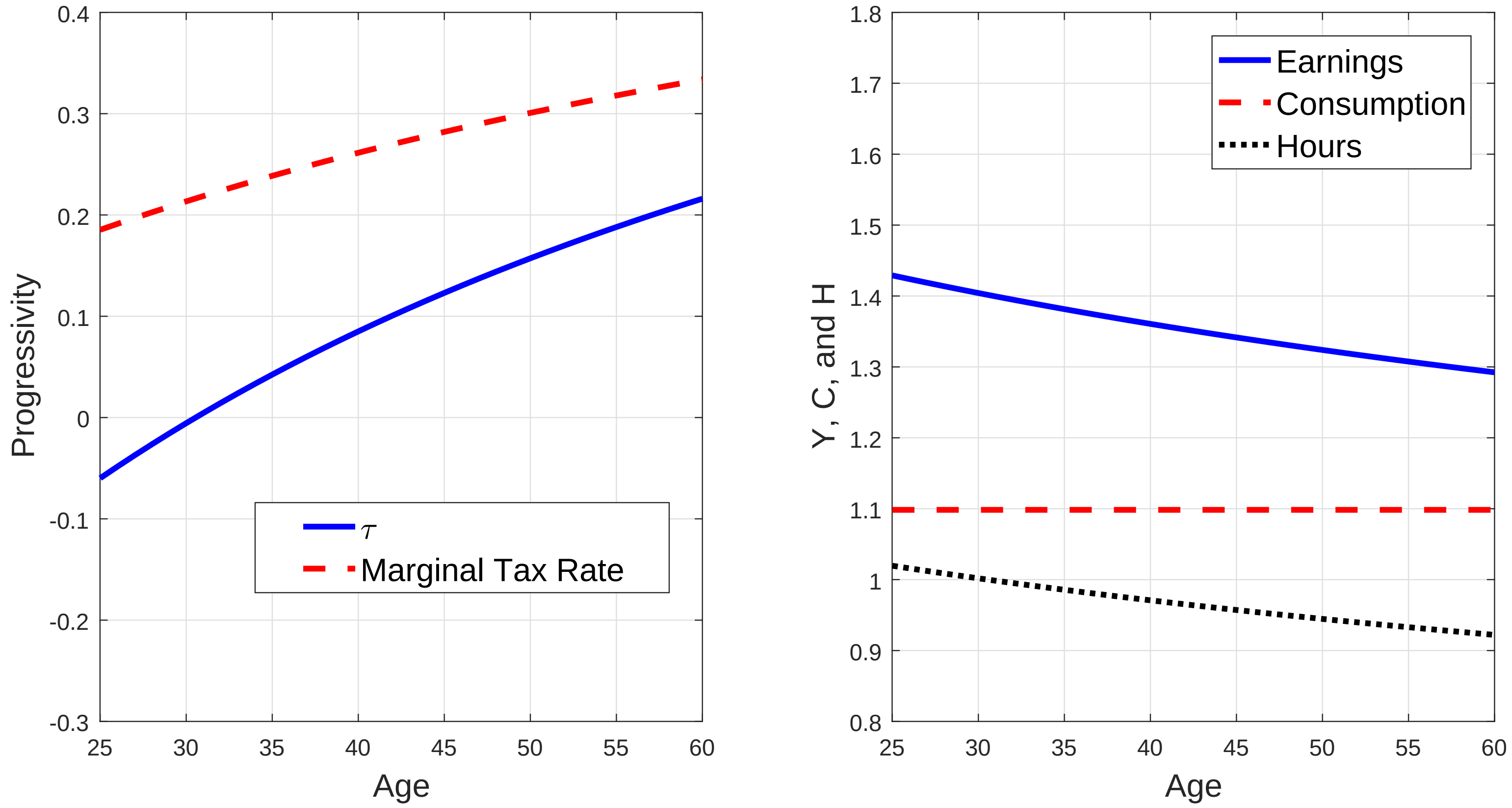
- τ_a^* still flat but shifted up (redistribution) $\Rightarrow$ lower labor supply

Add Heterogeneity in Ability (θ finite)



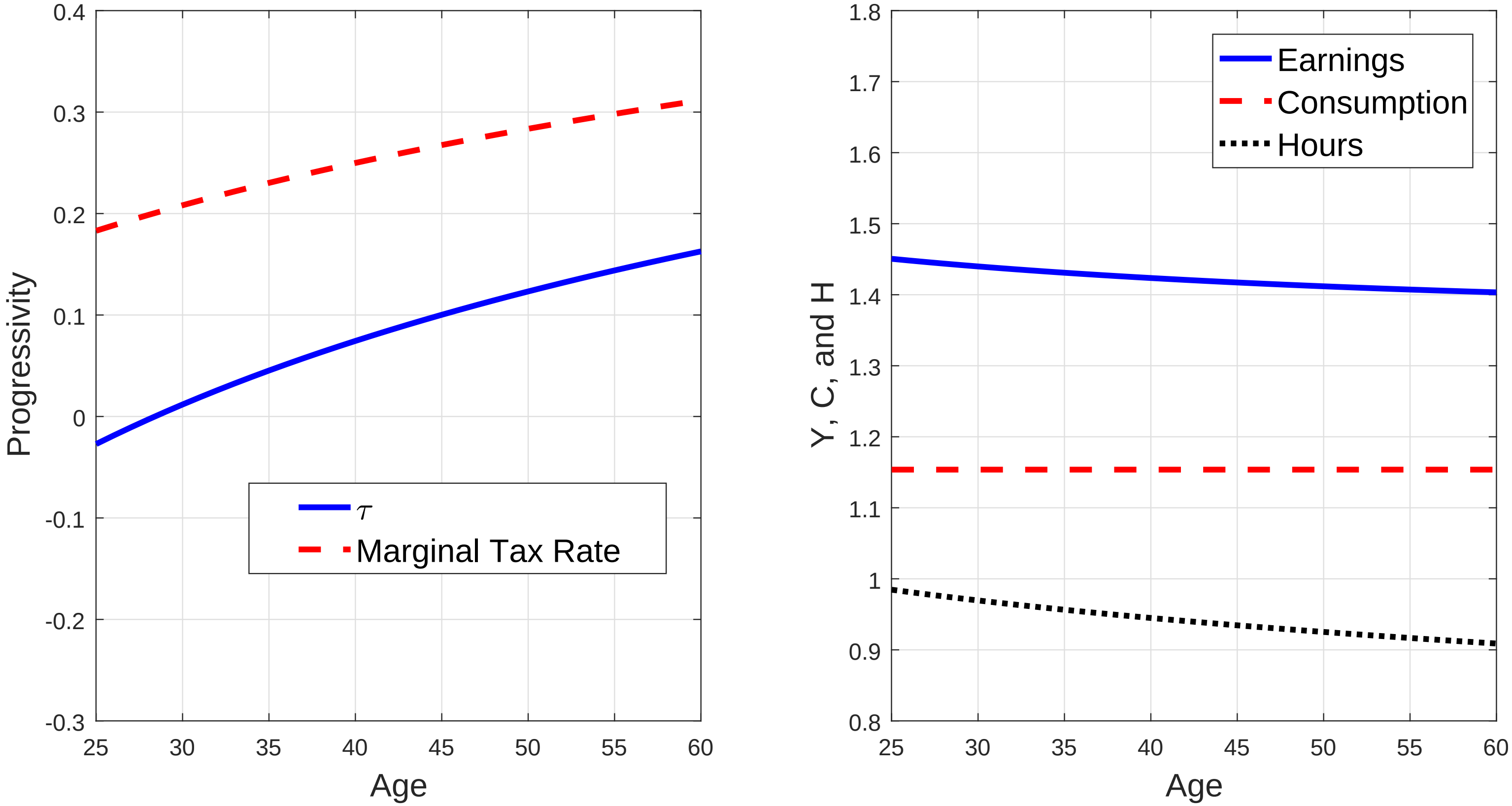
- τ_a^* still flat but shifted further up (redistribution > distortion)

Add Uninsurable Risk ($v_\omega > 0$)



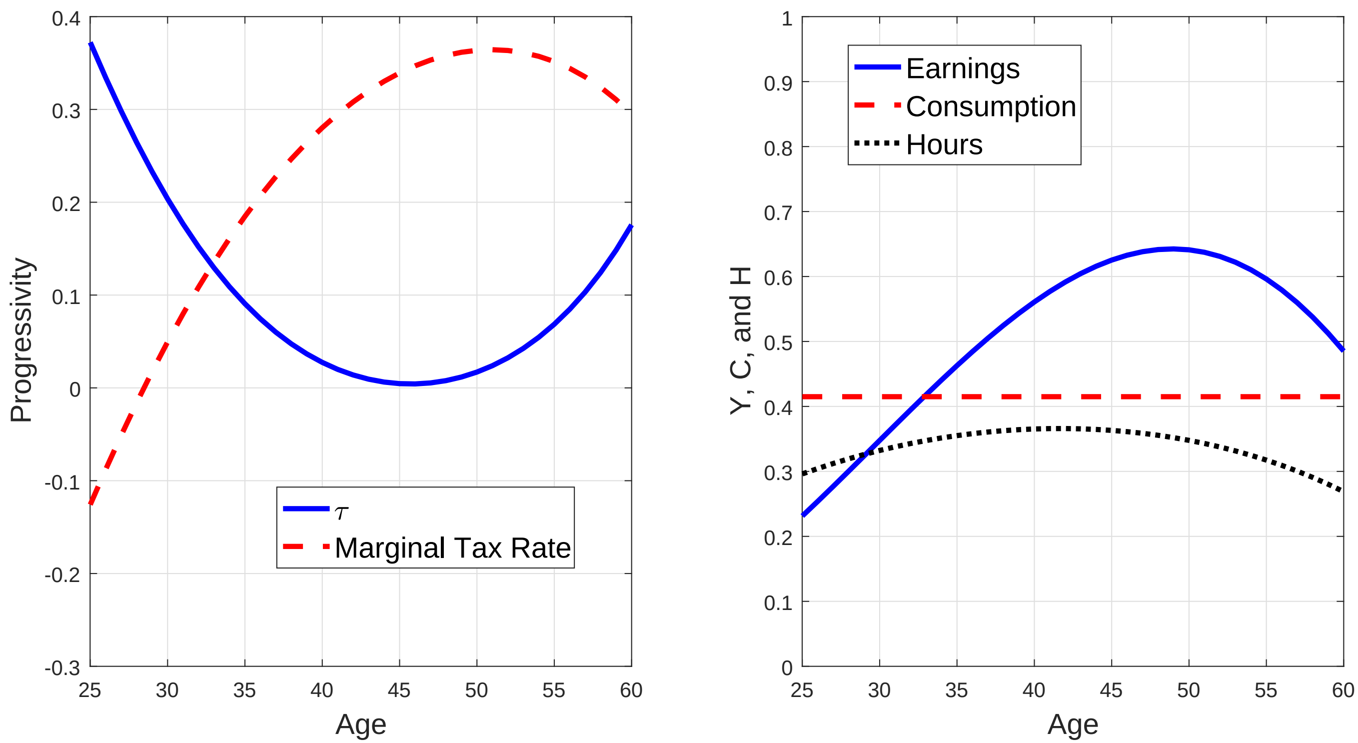
- Profile for τ_a^* steeper: more redistribution needed later in life since uninsurable risk cumulates

Add Insurable Risk ($v_\varepsilon > 0$)



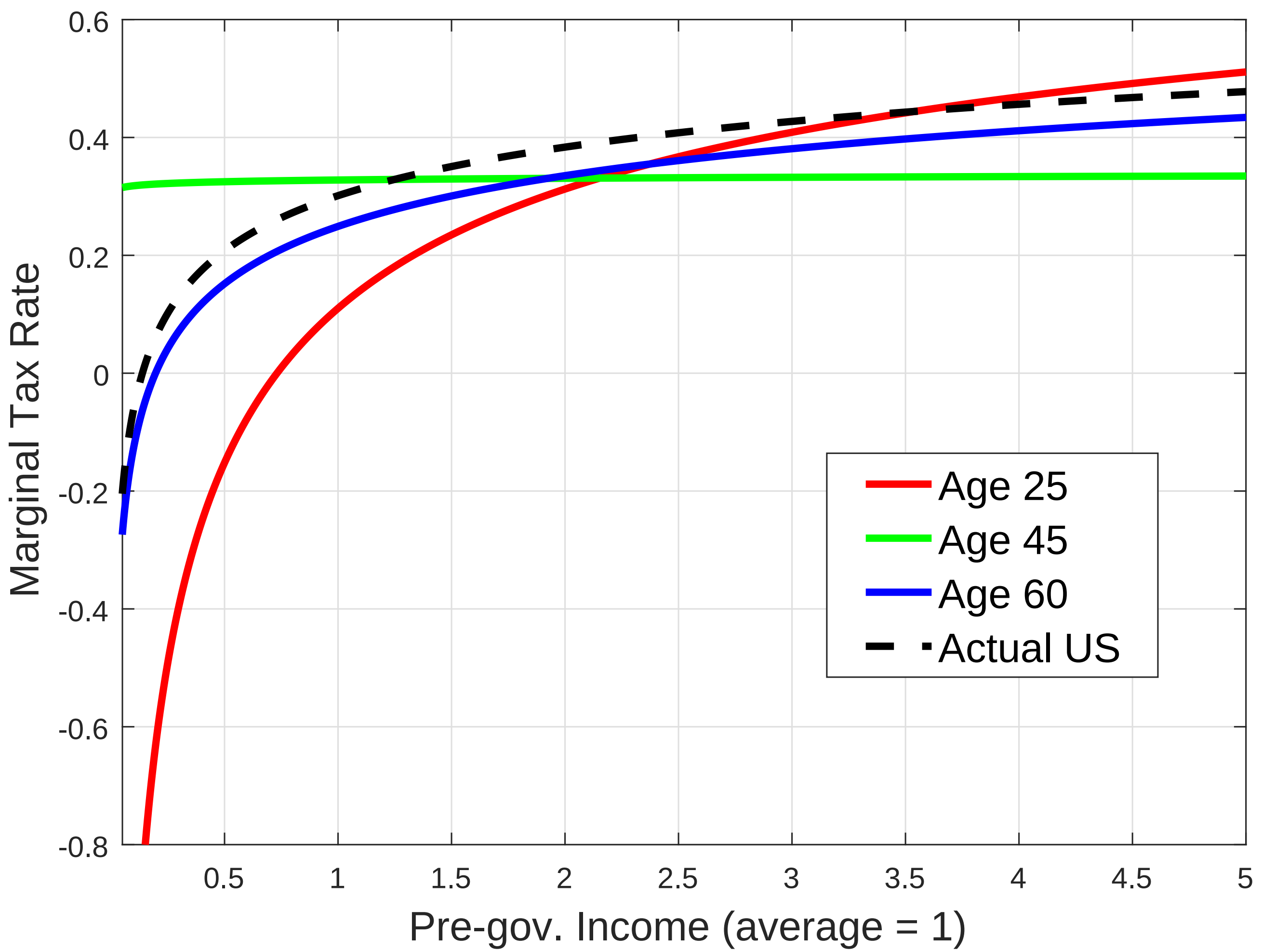
- Profile for τ_a^* is flattened but still upward sloping

Add Life Cycle $\{x_a, \bar{\varphi}_a\}$



- $x_a - \bar{\varphi}_a$ hump-shaped $\Rightarrow$ earnings are hump-shaped
- λ_a is U-shaped to equalize consumption across ages
- Smoothing $1 - MTR_a = \lambda_a(1 - \tau_a)y_a^{-\tau_a} \Rightarrow \tau_a$ is U-shaped as well

All Channels: Marginal Tax Rates by Age

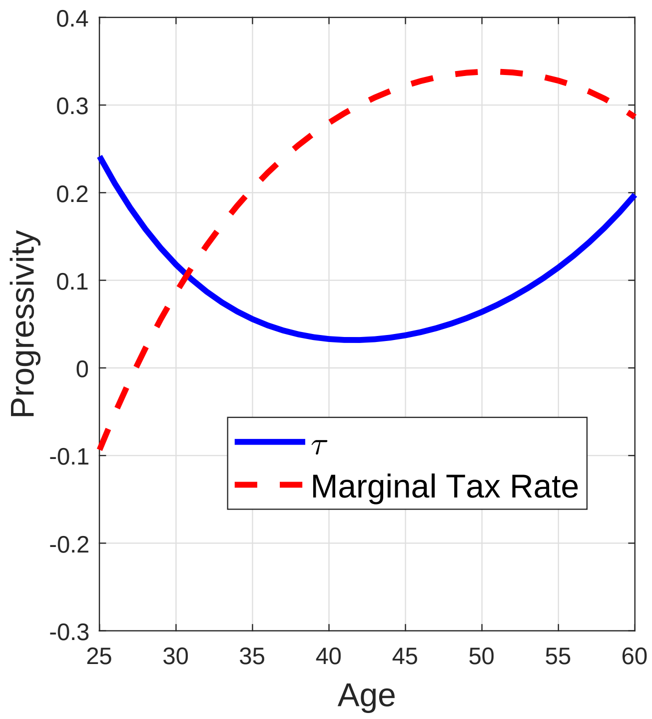


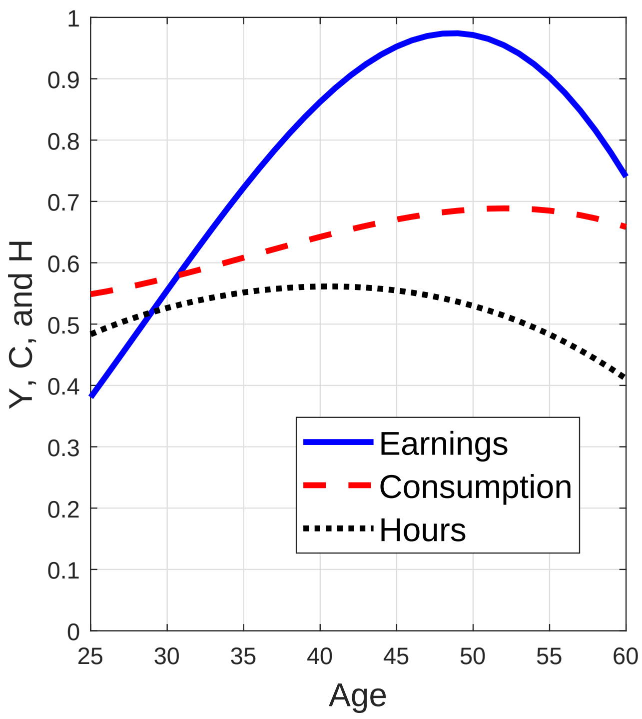
Age Varying Preferences for Consumption

- Use standard equivalence scale for household size to set desired consumption by age $\Rightarrow$ age path for $u(c_a)$ shifter

Age Varying Preferences for Consumption

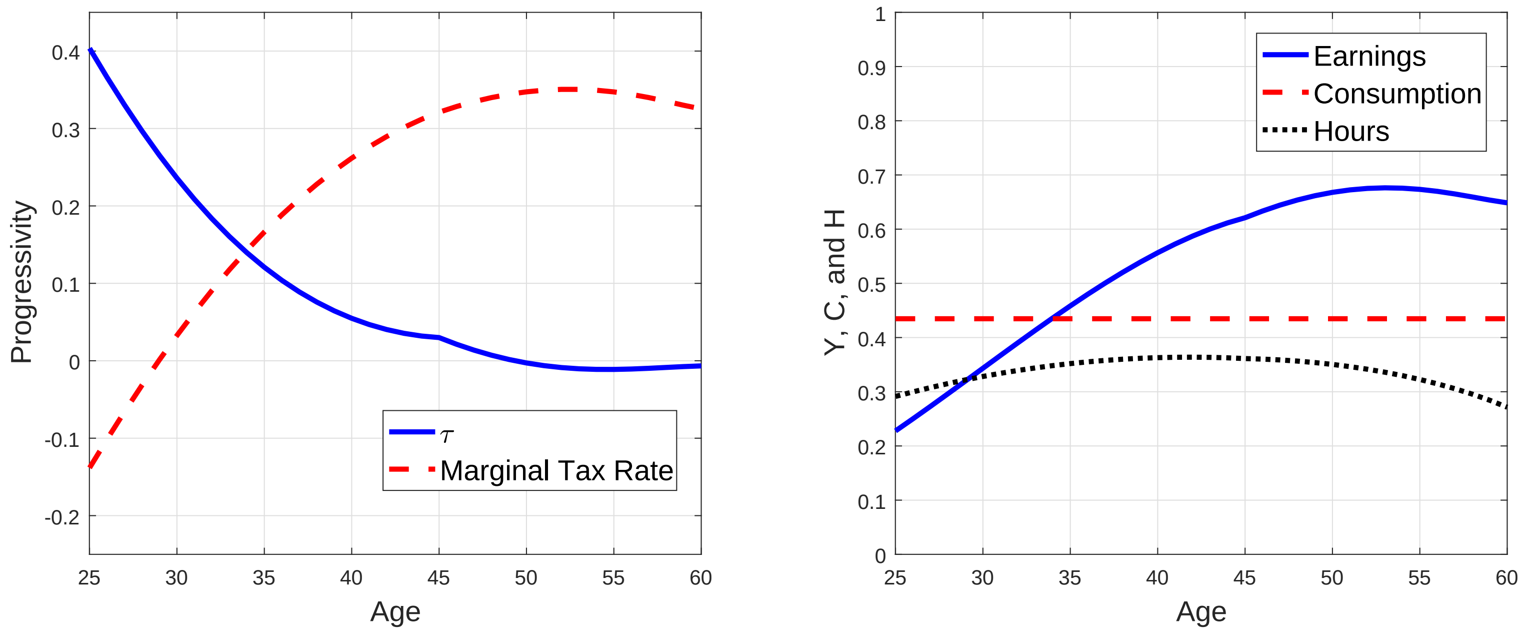
- Use standard equivalence scale for household size to set desired consumption by age $\Rightarrow$ age path for $u(c_a)$ shifter





- Some consumption inequality over the life cycle is efficient $\Rightarrow$ less redistribution through λ_a and flatter profile for τ_a

Age Varying Frisch Elasticity



- Frisch at age 60 three times larger than at age 45 (Blundell et al.)
- It pushes optimal progressivity down at older ages

Transitional Dynamics

$$\beta < 1$$

Optimal Policy with Transition

1. The optimal value for spending is $g_t = \frac{\chi}{1+\chi}$
2. Given any values for $\{\tau_{a,t}\}$, the optimal profiles $\{\lambda^*_{a,t}\}$ equate average consumption by age at each date t

Optimal Policy with Transition

1. The optimal value for spending is $g_t = \frac{\chi}{1+\chi}$

2. Given any values for $\{\tau_{a,t}\}$, the optimal profiles $\{\lambda^*_{a,t}\}$ equate average consumption by age at each date t

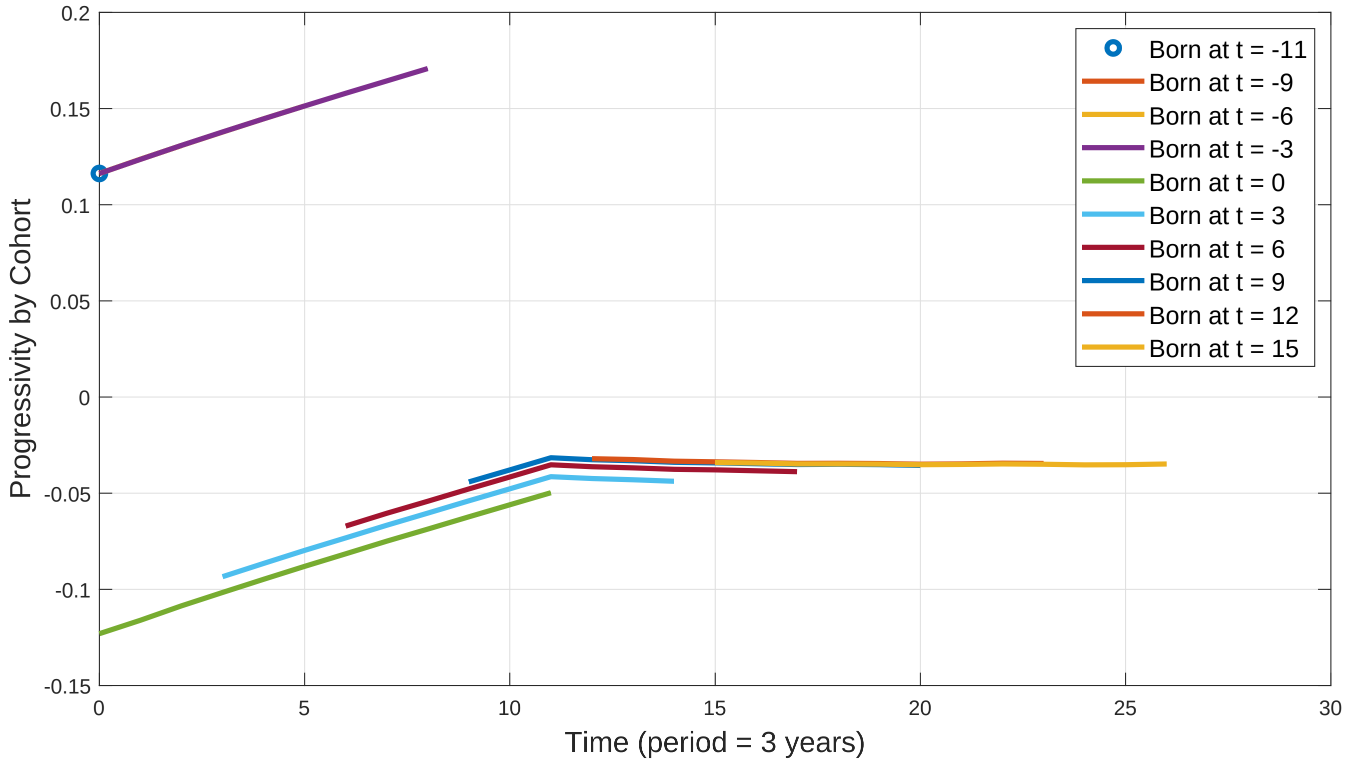
3. If (i) skill is the only source of heterogeneity and (ii) labor supply is inelastic, then optimal reform at $t = 0$ features:

   (a) $\tau^*_{a,t} = 1$ for all $a > t$ (max expropriation from existing cohorts)

   (b) $\tau^*_{0+j,t+j} = \tau^*_{0,t} < 1$ for all $j = 1, .., A - 1$ and for all $t \geq 0$ (flat τ_a profiles for the new cohorts)

   Reminiscent of capital taxation, but progressivity varies by cohort, not time since human capital is non-tradable

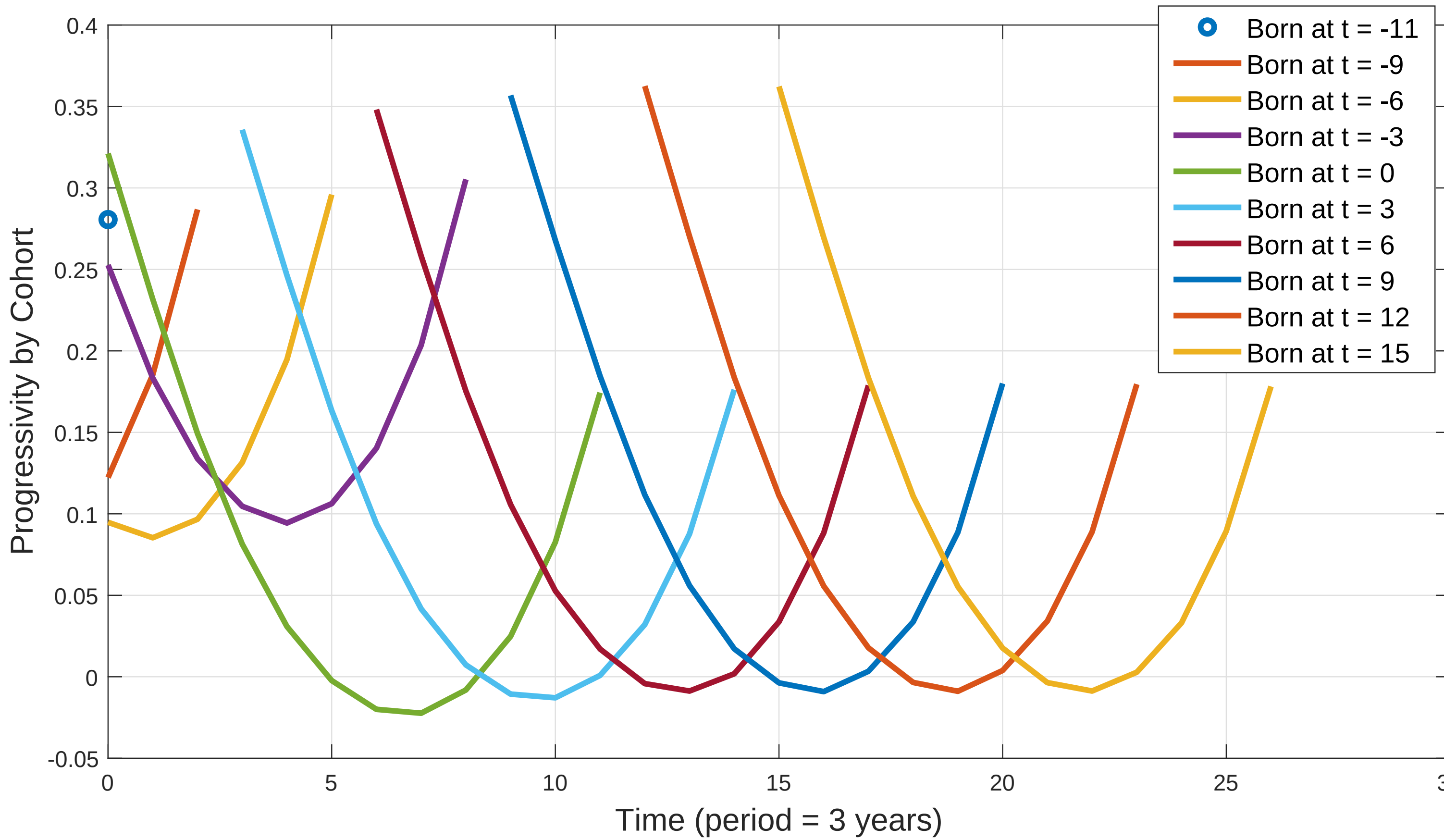
Transition: Skill Heterogeneity + Elastic Labor



1. τ_a higher for existing cohorts: no skill investment distortion

2. τ_a rises with age: output grows, planner can redistribute more

Optimal Policy with Transition: Baseline



Welfare Gains

- Equivalent variation: % of lifetime consumption
- Computed relative to the US tax/transfer system

	Benchmark	U.S. BL	Natural BL
(λ^*, τ^*) constant	0.10		
λ^* age-varying, τ^* constant	1.69		
λ^* constant, τ^* age-varying	2.10		
(λ^*, τ^*) age-varying	2.42		

INTERTEMPORAL TRADE

Introducing Borrowing and Lending

- Modification to baseline model:
  - Non-contingent bonds in zero net supply s.t. credit limit
  - No insurable productivity risk
  - Tax levied on y net of savings:

$$c_a = \lambda_a (wh + Rb - b')^{1-\tau_a}$$

Introducing Borrowing and Lending

- Modification to baseline model:
  - Non-contingent bonds in zero net supply s.t. credit limit
  - No insurable productivity risk
  - Tax levied on y net of savings:

$$c_a = \lambda_a (wh + Rb - b')^{1-\tau_a}$$

- Numerical solution:
  - Skill investment decision rules still in closed form
  - Solve numerically for hours worked, savings, interest rate
  - Search for optimal $\{\tau_a\}$ as 2nd order polynomial of age

Estimation of Consumer Credit Limit

- SCF 2013 data, households 25-60. We sum four components:

  (a) Limit on credit cards

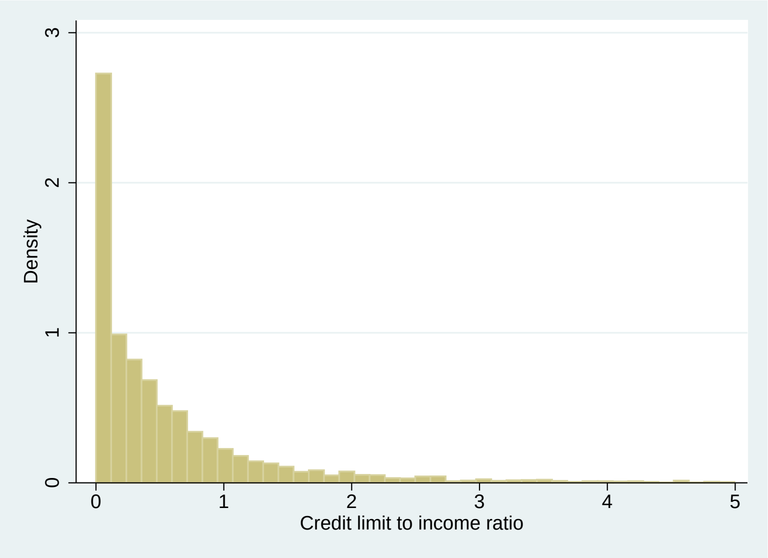
  (b) Limit on HELOCs

  (c) $2 \times$ installment loans for durables

  (d) $2 \times$ other debt (e.g., short-term loans from IRA)

Estimation of Consumer Credit Limit

- SCF 2013 data, households 25-60. We sum four components:
  (a) Limit on credit cards
  (b) Limit on HELOCs
  (c) $2 \times$ installment loans for durables
  (d) $2 \times$ other debt (e.g., short-term loans from IRA)



Estimation of Consumer Credit Limit

- SCF 2013 data, households 25-60. We sum four components:

  (a) Limit on credit cards

  (b) Limit on HELOCs

  (c) $2 \times$ installment loans for durables

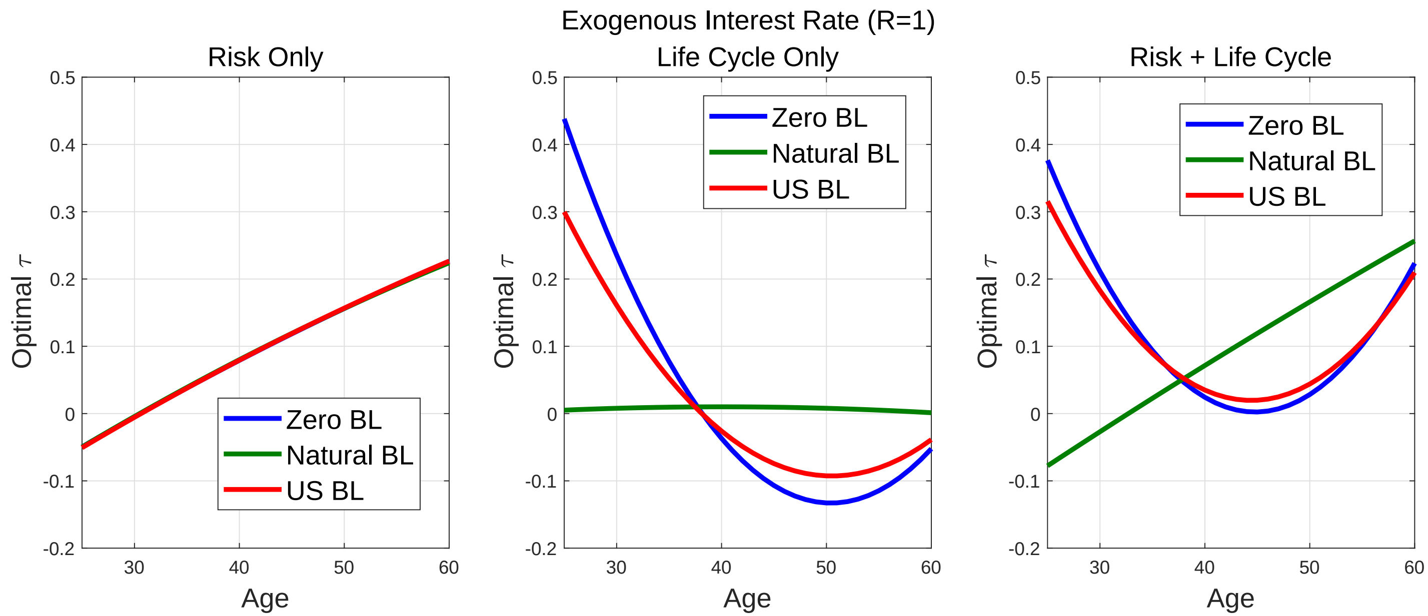
  (d) $2 \times$ other debt (e.g., short-term loans from IRA)

- We set it to $1.5 \times$ annual income (90th percentile)

- Zero BL (tightest) $\Rightarrow$ autarky

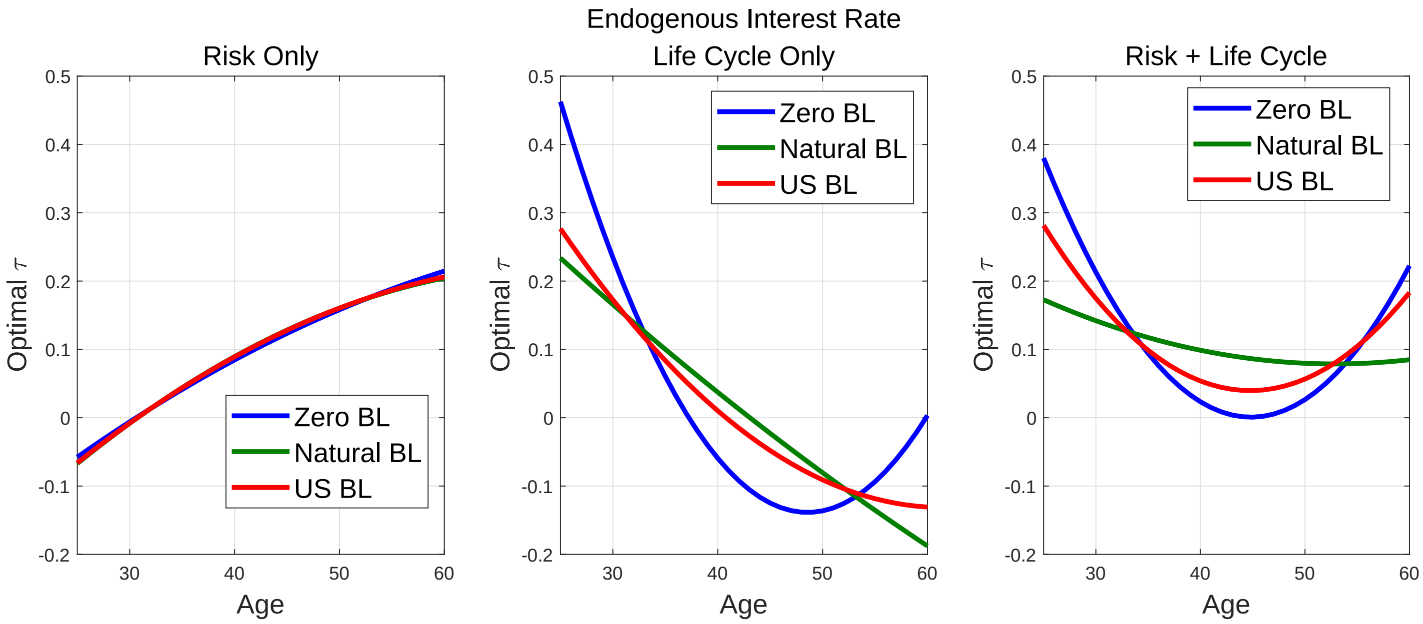
- Natural BL (loosest): max 30 times annual income

Optimal Progressivity with Borrowing/Saving: $\beta R = 1$



- Zero BL: $\{\tau_a^*\}$ almost identical to benchmark model
- Natural BL: $\{\tau_a^*\}$ close to a model with flat profile for $\{x_a - \bar{\varphi}_a\}$
- U.S. BL: $\{\tau_a^*\}$ very similar to autarky/benchmark case

Optimal Progressivity with Borrowing/Saving: R^*



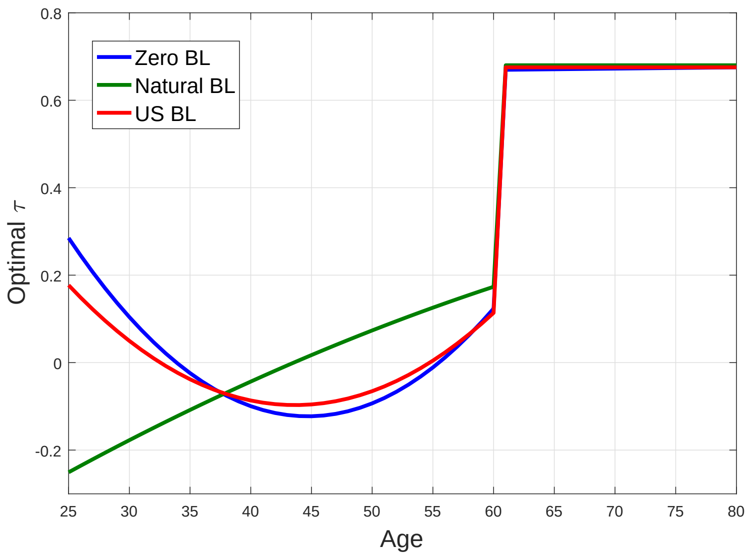
- Interest rate channel: $\{\tau_a^*\}$ more downward sloping
  - $\beta R^* > 1$, but planner wants to equate C_a across ages
  - λ_a decreasing so that after tax interest rate is 1 (EE wedge)
  - τ_a also decreasing to equate labor wedge

Extension with Retirement and Pensions

- Disposable income in retirement: $\lambda_a \left[p(s_i) \exp(\alpha_{i,A} - \varphi_i)\right]^{1-\tau_a}$

Extension with Retirement and Pensions

- Disposable income in retirement: $\lambda_a \left[p(s_i)\exp(\alpha_{i,A} - \varphi_i)\right]^{1-\tau_a}$



- Jump in τ_a: no labor supply distortion in retirement
- Flat profile in retirement: no motive for age dependence
- No full compression: it would distort too much dynamic skill choice
- Lower τ_a during working life: skill choice depends on $\bar{\tau}$

Welfare Gains

- Equivalent variation: % of lifetime consumption
- Computed relative to the US tax/transfer system

	Benchmark	U.S. BL	Natural BL
(λ^*, τ^*) constant	0.10	0.16	0.18
λ^* age-varying, τ^* constant	1.69	1.07	0.67
λ^* constant, τ^* age-varying	2.10	1.63	1.36
(λ^*, τ^*) age-varying	2.42	1.76	1.38

Lessons

- Distinct roles for λ_a and τ_a:
  - Tax level λ_a delivers redistribution across age groups
  - Progressivity τ_a is key for skill investment and labor supply distortions, and for redistribution / insurance within age groups
- Forces shaping how progressivity varies with age roughly offset:
  - Uninsurable risk + sunk skill investment $\Rightarrow \tau_a$ rises with age
  - Rising labor productivity and insurable risk $\Rightarrow \tau_a$ falls with age
- U-shape profile for progressivity is optimal, but dampened if:
  - borrowing limits are very loose
  - preferences for consumption display a strong hump

Thanks!