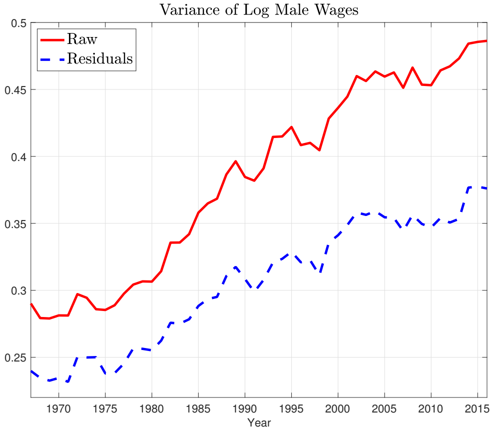
How Should Tax Progressivity Respond to Rising Income Inequality?

Jonathan Heathcote (Minneapolis Fed)
Kjetil Storesletten (U of Oslo)
Giovanni L. Violante (Princeton U)

Surrey

April 7, 2021

Rising wage dispersion in U.S:



1980-2005: variance of log wage increased by 0.15

- Between-group (college premium) inequality: up 0.05
- Within-group (residual) inequality: up 0.10

Government response

- How should government respond to rising inequality?
- Natural policy instrument: tax/transfer policies
- Other potential tools are less direct: education/training, labor market regulation (minimum wage, active labor policies), competition policies (trade, migration), ...
- Must be addressed within a macro model
  - ... need a theory of rising inequality,
    ... necessary to incorporate general-equilibrium effects of taxes, government budget balance
- Build and calibrate model that allows taxes to affect GDP, pre-tax income distribution, and consumption distribution

Paper in a nutshell

- Should the tax- and transfer system be amended to address rising inequality? How?
- Answer: it depends on **why** inequality increased
  - Within-group uninsurable risk $\Rightarrow \tau^* \uparrow$
  - Within-group insurable risk $\Rightarrow \tau^* \downarrow$
  - Skill-biased technical change (between-group) dispersion $\Rightarrow \tau^* \uparrow$
    if and only if human capital accumulation elasticity is low
- Calibrated model:
  - Opposing effect cancel Optimal progressivity is approximately unchanged.
  - Optimal to *reduce* τ^* in response to skill-biased technical change.
- ... **zero change** similar to empirical 1980-2015 outcome

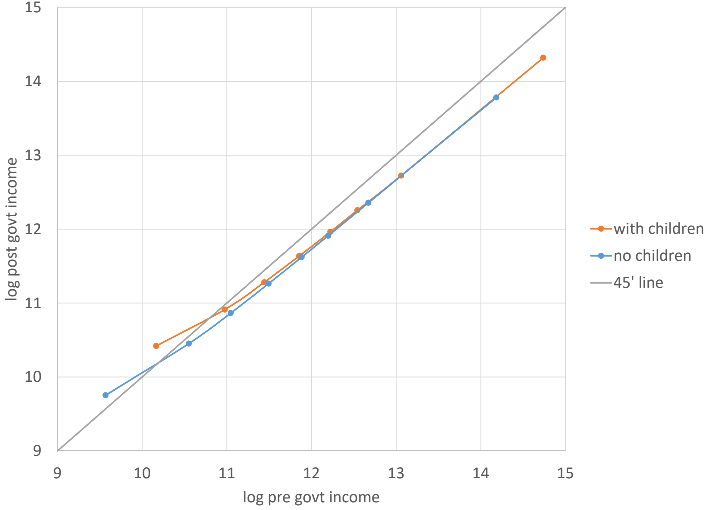
Plan

- Measure simplified tax/transfer system
- Lay out macro model that can address pros and cons of tax progressivity
- Use model to quantify optimal tax progressivity in response to rising inequality
  - optimal response hinges on drivers of inequality
  - estimate channels

U.S. Tax System - by Percentile

- How could complex tax systems be measured?
- Non-parametric: plot pre-tax income against disposable income by percentiles in logs
    - U.S.: use tax and transfer model from the Congressional Budget Office (CBO)
    - **Pre-gov. income**:
      Market income OR
      Market income + social security transfers
    - **Post-gov income**:
      Pre-government income minus taxes plus transfers
    - Use CBO tabulations for various quantiles
      (separately for households with children and non-elderly households without children)

U.S. Tax System - by Percentiles



Disposable income: log-linear function of y

- Log-linear tax/transfer function yields remarkably good fit:

$$\begin{aligned} \log\left[y - T(y)\right] &= \log(\lambda) + (1-\tau)\log\left[y\right] \\ &\Rightarrow \\ y - T(y) &= \lambda y^{1-\tau} \end{aligned}$$

  where $y =$ pre-tax/transfer earnings, $T(y) =$ tax

- log-linear tax/transfer fn.: long tradition in public econ.
  - Feldstein (1969), Jakobsson (1976), Kakwani (1977), Persson (1983), Benabou (2000), Heathcote, Storesletten, and Violante (2017), etc.

Taxes

- Assume the following tax system:

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

- Log-linear mapping between disposable (post-government) earnings $\tilde{y}_i$ and pre-government earnings y_i

$$\tilde{y}_i = \lambda y_i^{1-\tau}$$

- τ is an index of progressivity in two ways:
  1. Thus, $(1-\tau)$ measures the elasticity of post-tax to pre-tax income
  2. A tax scheme is "progressive" if the ratio of marginal to average tax rates is larger than one for every level of income y_i. Within our class, we have

$$\frac{T'(y_i)}{T(y_i)/y_i} = \frac{1-\lambda(1-\tau)\,y_i^{-\tau}}{1-\lambda y_i^{-\tau}}.$$

  - When $\tau > 0$, the ratio is > 1.

Disposable income: log-linear function of y

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

- The parameter τ measures the degree of progressivity:
  - $\tau = 1:$ full redistribution $\rightarrow$ $T(y) = y - \lambda$
  - $0 < \tau < 1$: progressivity $\rightarrow$ $T'(y) > \frac{T(y)}{y}$
  - $\tau = 0:$ no redistribution $\rightarrow$ $T'(y) = \frac{T(y)}{y} = 1 - \lambda$
  - $\tau < 0:$ regressivity $\rightarrow$ $T'(y) < \frac{T(y)}{y}$
- Break-even income level: $y^0 = \lambda^{\frac{1}{\tau}}$

Restrictions: (i) no lump-sum transfer $\&$ (ii) $T'(y)$ monotone

Alternative estimates for progressivity τ (CBO data)

Specification	Estimates of τ		
	1979-83	2012-16	Change
Baseline			
With children	0.218	0.207	-0.011
Without children	0.155	0.164	+0.009
Average	0.186	0.186	-0.001
Alternative income measures			
Pre-govt inc. $\equiv$ market income	0.236	0.216	-0.020
Pre-govt inc. $\equiv$ post-govt + taxes	0.089	0.109	+0.019
Alternative samples			
First quintile dropped	0.083	0.112	+0.029
Top 5% only	0.043	0.051	+0.008

U.S. Progressivity Estimates over Time

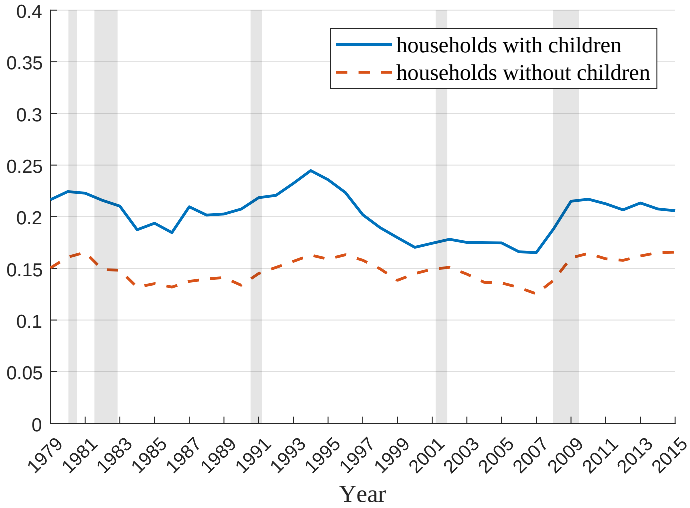


Figure: Estimates obtained by OLS on CBO quantiles

Avg. net tax rates, households w/children

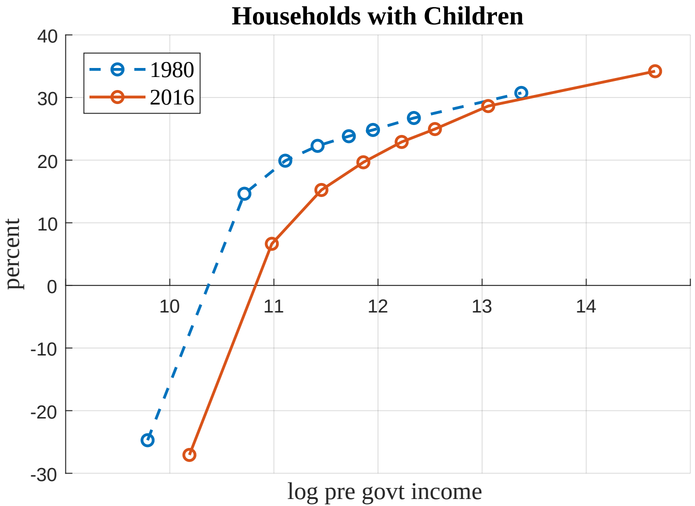


Figure: Average net tax rates (taxes minus transfers)/income. Data: CBO

Avg. net tax rates, households w/o children

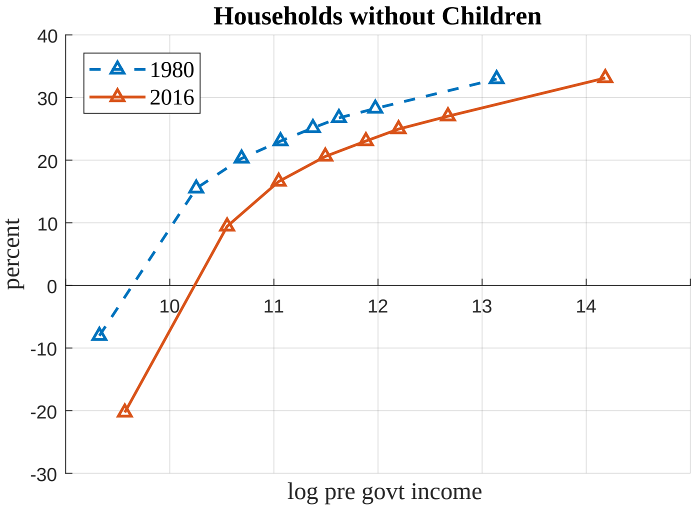


Figure: Average net tax rates (taxes minus transfers)/income. Data: CBO

Progressivity (τ) over time

- Conclusion: progressivity of the U.S. tax- and transfer system has remained approximately constant since 1980
- Some differences across demographic groups:
  - slight increase in progressivity for households w/o children
  - slight decline in progressivity for households with children
- Progressivity at bottom quintile is larger than for top 80%
  - increased progressivity for top 80% of population
  - evidence for “floor” in disposable income
  - transfers matter!
- Top 5%: higher average tax rates & lower progressivity than bottom 95%. Slight increase over time

Comparison with other studies

- Findings are line with Slemrod and Bakija (2017)
- Piketty and Saez (2007) claim US tax system became less progressive 1960-2004. Differently from us, they ...
  - Focus on top 0.1% of income distribution
  - Incorproate estate taxes and make different assumptions from CBO on corporate taxes

How should tax system respond to rising inequality?

Simple macro model with heterogeneity

Technology

Output Y is a constant elasticity of substitution aggregate of effective hours supplied by the continuum of skill types $s \in [0, \infty)$,

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

- $\theta > 1$ is the elasticity of substitution across skill types
- $m(s)$ is the density of individuals with skill level s.
- Note: all skill levels enter symmetrically in the production technology, and thus any equilibrium differences in skill prices will reflect relative scarcity in the context of a model in which different skill types are imperfect substitutes. In

Preferences

Preferences over private consumption and skill investment effort

$$\begin{aligned} U_i &= v_i(s_i) + u(c_i) \\ &= -\frac{(\kappa_i)^{-1/\psi}}{1+1/\psi}(s_i)^{1+1/\psi} + \ln(c_i) \end{aligned}$$

- $\psi > 0$ is the elasticity of substitution for skill effort
- $\kappa_i \geq 0$ is a parameter is individual's utility cost of acquiring skills.
  - Note: The larger is κ_i, the smaller is the cost, so κ_i is "innate learning ability"
  - Assume $\kappa_i \sim Exp(\eta)$, an exponential distribution with parameter η.

Budget constraint

- Each worker provides one unit of labor of skill type s_i
- $p(s)$ is the price schedule for skill
- Individual's budget constraint

$$c_i = p(s_i),$$

Optimal skill choice

- Optimal skill choice:

$$
\begin{aligned}
\frac{\partial v_i(s)}{\partial s} &= \frac{\partial}{\partial s} u_i(c_i) \\
\left(\frac{s}{\kappa_i}\right)^{\frac{1}{\psi}} &= \frac{\partial}{\partial s}\left\{\ln p(s)\right\} \\
&\Rightarrow \\
s_i &= \left[\frac{\partial}{\partial s}\left\{\ln p(s)\right\}\right]^{\psi} \cdot \kappa_i
\end{aligned}
$$

- Problem: must find the schedule $p(s)$ before computing optimal skill effort

Finding p(s) (cont.)

- Guess that $\ln p(s)$ is linear in s, i.e.,

$$\ln p(s) = \pi_0 + \pi_1 s$$

- The optimal skill choice then becomes

$$s_i = (\pi_1)^{\psi} \cdot \kappa_i$$

- skills inherit the exponential distrbution from κ, with parameter

$$\begin{aligned} \zeta &\equiv \eta [\pi_1]^{-\psi} \\ m(s) &= \zeta \exp(-\zeta s) \end{aligned}$$

Finding p(s) (cont.)

- Market clearing in labor markets for each skill type $\Rightarrow p(s)$ is the marginal product from an additional unit of effective hours of skill type s :

$$\begin{aligned} p(s) &= \frac{\partial}{\partial m(s)} \left\{ \left(\int_0^\infty [m(s)]^{\frac{\theta-1}{\theta}} ds \right)^{\frac{\theta}{\theta-1}} \right\} \\ &= \left(\frac{Y}{m(s)} \right)^{\frac{1}{\theta}} \end{aligned}$$

- Recall that, according to the guess, $m(s) = \zeta \exp(-\zeta s)$ so

$$\begin{aligned} \ln p(s) &= \frac{1}{\theta} \ln Y - \frac{1}{\theta} \ln m(s) \\ &= \frac{1}{\theta} \ln Y - \frac{1}{\theta} \ln(\zeta) + \frac{1}{\theta} \zeta s \end{aligned}$$

Finding p(s) (cont.)

- Equate coefficients:

$$\pi_0 + \pi_1 s = \frac{1}{\theta} \ln Y - \frac{1}{\theta} \ln \left(\eta \left[\pi_1 \right]^{-\psi} \right) + \frac{1}{\theta} \eta \left[\pi_1 \right]^{-\psi} s$$

- Implies

$$\begin{aligned} \pi_1 &= \frac{1}{\theta} \eta \left[\pi_1 \right]^{-\psi} \\ &\Rightarrow \\ \pi_1 &= \left(\frac{\eta}{\theta} \right)^{\frac{1}{1+\psi}} \end{aligned}$$

- Calculate the skill premium:

$$\pi_1 \cdot s(\kappa) = \pi_1 \cdot \left[\pi_1 \right]^{\psi} \cdot \kappa = \frac{\eta}{\theta} \cdot \kappa$$

Household's problem revisited

- Guess again that $p(s) = \pi_0 + \pi_1 s$
- Budget constraint is

$$\begin{aligned} c_i &= \lambda \left(p\left(s_i\right)\right)^{1-\tau} \\ &\Rightarrow \\ \ln c_i &= \ln \lambda + (1-\tau) \ln \left(p\left(s_i\right)\right) \\ &= \ln \lambda + (1-\tau)\, \pi_0 + (1-\tau)\, \pi_1 s_i \end{aligned}$$

- First-order condition for skill choice is

$$s_i = \left((1-\tau)\, \pi_1\right)^{\psi} \cdot \kappa_i$$

- The skill distribution becomes

$$\zeta \equiv \eta \left[(1-\tau)\, \pi_1\right]^{-\psi}$$

Solve for return to skill and skill inv.

- Equating coefficients yields the return to skill π_1:

$$\pi_1(\tau) = \left(\frac{\eta}{\theta}\right)^{\frac{1}{1+\psi}} (1-\tau)^{-\frac{\psi}{1+\psi}}$$

- The skill investment allocation is given by

$$s(\kappa;\tau) = \left[\frac{\eta}{\theta}(1-\tau)\right]^{\frac{\psi}{1+\psi}} \cdot \kappa$$

- Consider comparative statics on the return to skill π_1 and skill investment s, letting $\psi \to 0$ and $\psi \to \infty$

Cross-sectional inequality

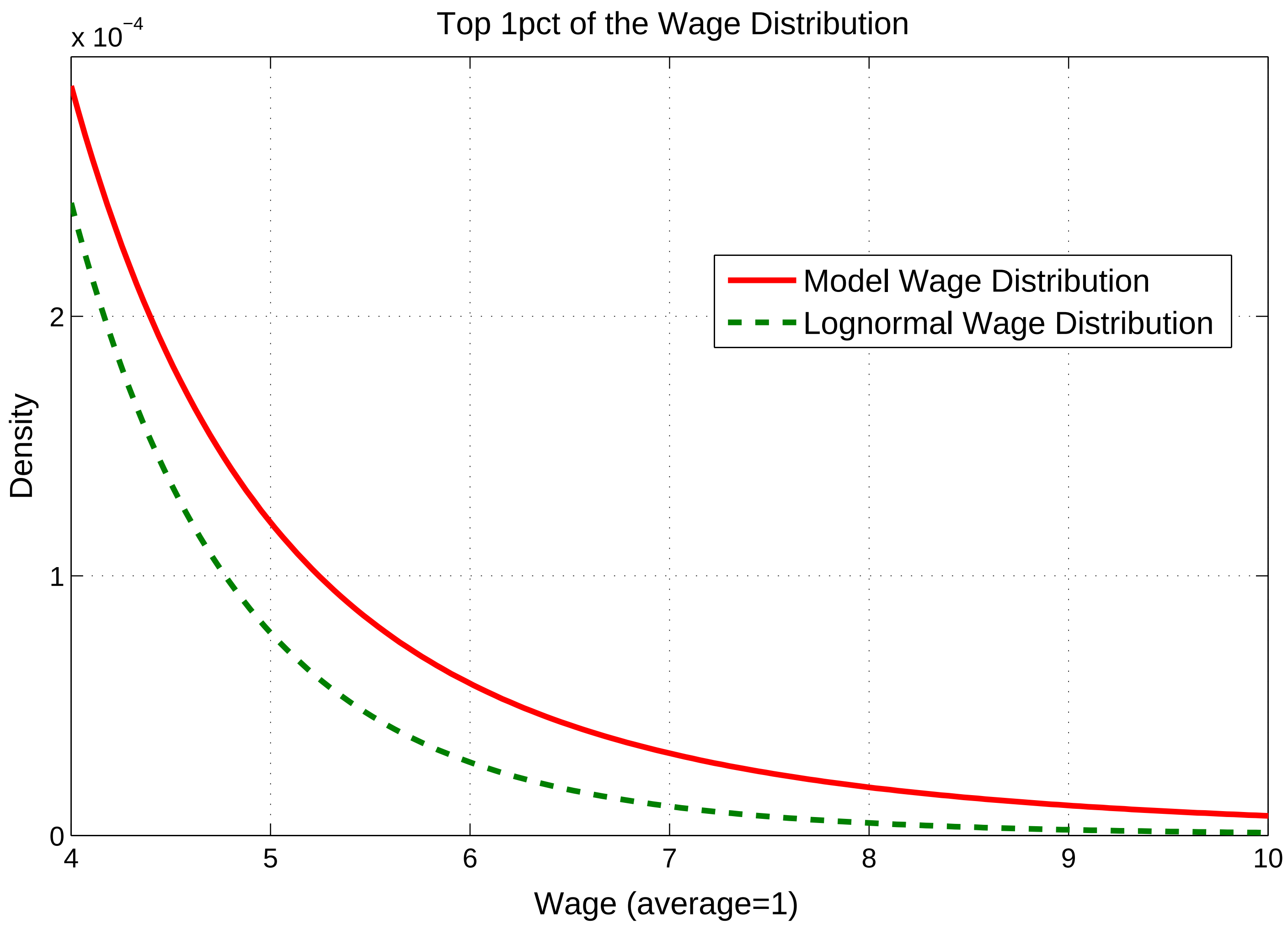
- Recalculate the skill premium:

$$\pi_1 \cdot s(\kappa) = \pi_1 \cdot \left[(1-\tau)\,\pi_1\right]^{\psi} \cdot \kappa = \frac{\eta}{\theta} \cdot \kappa$$

- The variance of log wages is then

$$var\left(\ln p\left(s_i\right)\right) = var\left(\pi_1 \cdot s(\kappa)\right) = \left(\frac{\eta}{\theta}\right)^2 \cdot var\left(\kappa\right) = \frac{1}{\theta^2}$$

Upper tail of wage distribution



Calculate lambda in equilibrium

- Recall that $\tilde{y}_i = \lambda\left(y_i\right)^{1-\tau}$. If $\tilde{Y} = \int \tilde{y}_i di = \int y_i^{1-\tau} di$, we have
$$\lambda = \frac{Y}{\tilde{Y}}. \tag{1}$$
- Aggregate output equals aggregate labor earnings

$$\begin{aligned} Y\left(\tau\right) &= \int p(s;\tau)\, dF_s \\ &= \int p\left(s;\tau\right) dF_s \\ &= \frac{\theta}{\theta-1}\cdot \exp\left(\pi_0\left(\tau\right)\right) \end{aligned} \tag{2}$$

Calculate pi_0

- From the production function and distribution for s,

$$
\begin{aligned}
Y &= \left\{ \int_0^\infty \left[\zeta \exp\left(-\zeta s\right)\right]^{\frac{\theta-1}{\theta}} ds \right\}^{\frac{\theta}{\theta-1}} \\
&= \left(\frac{\theta}{\theta-1}\right)^{\frac{\theta}{\theta-1}} \left(\eta\right)^{\frac{1}{1+\psi}\frac{1}{1-\theta}} \left(\frac{\theta}{1-\tau}\right)^{\frac{\psi}{\psi+1}\frac{1}{1-\theta}}. \qquad (3)
\end{aligned}
$$

- Comparing this to $Y = \frac{\theta}{\theta-1}\exp\left(\pi_0\right)$ and get

$$
\pi_0\left(\tau\right) = \frac{1}{\theta-1}\frac{\psi}{1+\psi}\log\left(1-\tau\right) + \bar{\pi}
$$

where the constant $\bar{\pi}$ is

$$
\bar{\pi} \equiv -\frac{1}{\theta-1}\frac{1}{1+\psi}\left(\log\eta + \psi\log\theta\right) + \frac{1}{(\theta-1)}\log\left(\frac{\theta}{\theta-1}\right)
$$

Computing lambda (cont.)

- Compute $\tilde{Y}$; the average per capita disposable income for households

$$\begin{aligned}\tilde{Y} &= \int [y(s)]^{1-\tau} m(s)\, ds \\ &= \int [\exp(p(s))]^{1-\tau} m(s)\, ds \\ &= \exp[(1-\tau)\pi_0(\tau)] \int \exp[(1-\tau)\,\pi_1(\tau)s]\, m(s)\, ds \\ &= \frac{\theta}{\theta-1+\tau} \cdot \exp[(1-\tau)\pi_0(\tau)]\end{aligned}$$

- So

$$\begin{aligned}\ln\lambda &= \ln \frac{\frac{\theta}{\theta-1}\cdot\exp(\pi_0)}{\frac{\theta}{\theta-1+\tau}\cdot\exp[(1-\tau)\pi_0(\tau)]} \\ &= \ln\left(\frac{\theta-1+\tau}{\theta-1}\right) + \tau\pi_0,\end{aligned}$$

Computing lambda (cont.)

- ... implying

$$\begin{aligned}\ln\lambda &= \ln\left(\frac{\theta-1+\tau}{\theta-1}\right)+\tau\pi_0 \\ &= \frac{\psi}{1+\psi}\frac{\tau}{\theta-1}\log\left(\frac{1-\tau}{\theta}\right)-\frac{1}{1+\psi}\frac{\tau}{\theta-1}\log\left(\eta\right) \\ &\quad +\frac{\theta+\tau-1}{\theta-1}\log\left(\frac{\theta}{\theta-1}\right)+\log\left(\frac{\theta-1+\tau}{\theta}\right).\end{aligned}$$

Compute aggregate welfare

- Focus on equally-weighted social welfare function

$$W = \int \left[-v_i\left(s_i\right) + u\left(c_i\right) \right] dF_i$$

- Substituting the equilibrium allocations into utility yields

$$\begin{aligned} u(c) &= \log \lambda + (1-\tau)\,\pi_1 s + (1-\tau)\,\pi_0 \\ &= \log \lambda + (1-\tau)\kappa \frac{\eta}{\theta} \\ &\quad + \frac{1-\tau}{(\theta-1)(\psi+1)} \left(\psi \log(1-\tau) + \log \left(\frac{1}{\eta} \frac{\theta}{(\theta-1)^{(1+\psi)}} \right. \right. \end{aligned}$$

- The disutility cost from investing in education is

$$v\left(s\left(\kappa\right)\right) = \frac{\kappa^{-1/\psi}}{1+1/\psi} \cdot s(\kappa;\tau)^{1+1/\psi} = (1-\tau)\,\kappa \frac{\eta}{(1+1/\psi)\,\theta}.$$

Compute aggregate welfare (cont.)

- Add things up

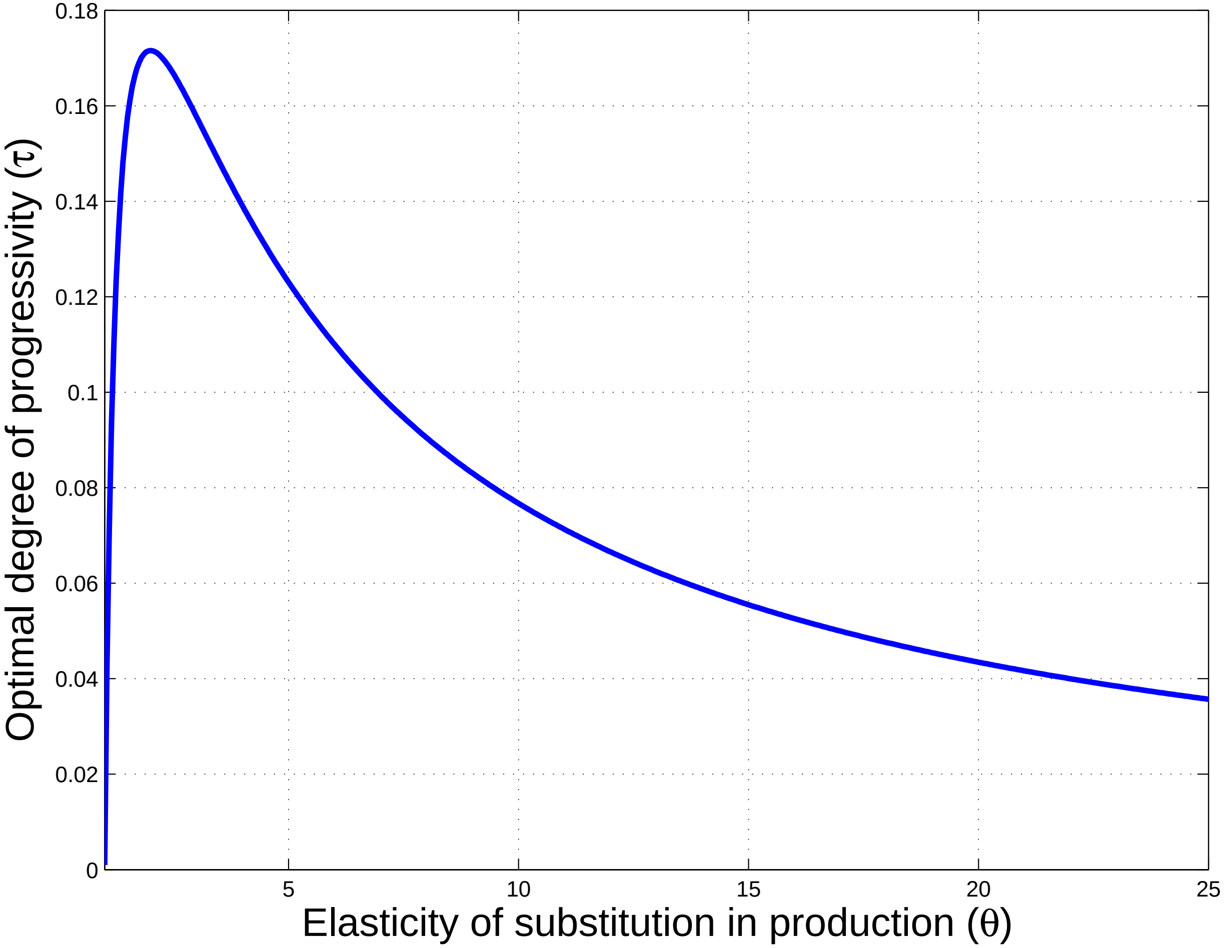
$$
\begin{aligned}
W &= \int \left[-v_i\left(s_i\right)+u\left(c_i\right)\right] dF_i \\
&= \int \left[-\frac{(\kappa)^{-1/\psi}}{1+1/\psi}\left(s\left(\kappa\right)\right)^{1+1/\psi}+\ln\lambda+\left(1-\tau\right)\ln\left(p\left(s\left(\kappa\right)\right)\right)\right] \\
&= \ln\lambda+\int\left[-\frac{(\kappa)^{-1/\psi}}{1+1/\psi}\left(\left[\frac{\eta}{\theta}\left(1-\tau\right)\right]^{\frac{\psi}{1+\psi}}\cdot\kappa\right)^{1+1/\psi}\right] dF\left(\kappa\right) \\
&\quad +\int\left[\left(1-\tau\right)\pi_0+\left(1-\tau\right)\pi_1 s\left(\kappa\right)\right] dF\left(\kappa\right)
\end{aligned}
$$

Compute aggregate welfare (cont.)

- Add things up

$$
\begin{aligned}
W &= \ln\lambda + (1-\tau)\,\pi_0 + \int \Big[-\frac{\psi}{1+\psi}\frac{\eta}{\theta}\,(1-\tau)\,(\kappa) + (1-\tau)\,\frac{\eta}{\theta} \\
&= \ln\left(\frac{\theta-1+\tau}{\theta-1}\right) + \tau\pi_0 + (1-\tau)\,\pi_0 - \frac{\psi}{1+\psi}\frac{1-\tau}{\theta} + \frac{1-}{\theta} \\
&= \ln\left(\frac{\theta-1+\tau}{\theta-1}\right) + \pi_0 - \frac{\psi}{1+\psi}\frac{1-\tau}{\theta} + \frac{1-\tau}{\theta}
\end{aligned}
$$

Skill investment component



Optimal marginal tax rate at the top

- Diamond-Saez formula (Mirlees approach):

$$\bar{t} = \frac{1}{1 + \zeta^u + \zeta^c(\theta - 1)} = \frac{1 + \sigma}{\theta + \sigma}$$

  - Decreasing in θ (increasing in thickness of Pareto tail)

- Our model: there is a range where τ is increasing in θ
  - When complementarity is high, cost of distorting skill investment is large
  - Key difference: exogenous vs endogenous skill distribution

Simple model framework

- Model ingredients:

    1. partial insurance against labor-income risk [ex-post heterogeneity]
    2. differential diligence & (learning) ability [ex-ante heterogeneity]
    3. flexible labor supply
    4. endogenous skill investment + multiple-skill technology [investment distortion and endogenous inequality]
    5. government consumption valued by households

- Ramsey approach: fix mkt structure & tax instruments

Demographics and preferences

- Perpetual youth model: constant survival rate δ
- Preferences over consumption (c), hours (h), publicly-provided goods (G), and skill-investment (s) effort:

$$U_i = -v_i(s_i) + \mathbb{E}_0 \sum_{t=0}^{\infty} (\beta\delta)^t u_i(c_{it}, h_{it}, G)$$

$$\begin{aligned}
v_i(s_i) &= \frac{1}{(\kappa_i)^{1/\psi}} \cdot \frac{s_i^{1+1/\psi}}{1+1/\psi} \\
\kappa_i &\sim Exp\,(1) \\
u_i\,(c_{it}, h_{it}, G) &= \log c_{it} - \exp(\varphi_i)\frac{h_{it}^{1+\sigma}}{1+\sigma} + \chi \log G \\
\varphi_i &\sim \mathcal{N}\left(\frac{v_\varphi}{2}, v_\varphi\right), \quad \varphi_i \perp \kappa_i
\end{aligned}$$

Technology

- Aggregate effective hours by skill type:

$$N(s) = \int_0^1 \mathbb{I}_{\{s_i = s\}} \, z_i h_i \, di$$

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

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

  where ρ is "absolute advantage" of skill s
- Skill price is marginal product of skill:

$$\log p(s) = \frac{1}{\theta} \ln \left(\frac{Y}{N} \right) + \rho s - \frac{1}{\theta} \ln \left[m(s) \right].$$

Individual efficiency units of labor

$$\log z_{it} = \alpha_{it} + \varepsilon_{it}$$

- $\alpha_{it} = \alpha_{i,t-1} + \omega_{it}$ with $\omega_{it} \sim \mathcal{N}\left(-\frac{v_\omega}{2}, v_\omega\right)$ [permanent]
- ε_{it} i.i.d. over time with $\varepsilon_{it} \sim \mathcal{N}\left(-\frac{v_\varepsilon}{2}, v_\varepsilon\right)$ [transitory]
- Pre-government earnings:

$$y_{it} = \underbrace{p(s_i)}_{\text{skill price}} \times \underbrace{\exp(\alpha_{it} + \varepsilon_{it})}_{\text{efficiency}} \times \underbrace{h_{it}}_{\text{hours}}$$

determined by human capital investment, luck, and effort

Market structure

- Competitive markets for labor and final good (numeraire)
- annuity markets against survival risk
- A risk-free bond in zero net supply
- Full insurance against transitory ε shocks
- No markets to insure permanent ω shock
  - $v_\varepsilon > 0, v_\omega > 0 \rightarrow$ partial insurance economy
  - $v_\omega = 0 \rightarrow$ full insurance economy
  - $v_\omega = v_\varepsilon = v_\varphi = 0 \quad \& \quad \theta = \infty \rightarrow$ RA economy

Aggregate resource constraint

- No capital and no government debt
  $\Rightarrow$ zero aggregate wealth
- Aggregate resource constraint:

$$Y = \int_0^1 c_i \, di + G$$

Planner

- Government budget constraint (no government debt):

$$G = \int_0^1 T\left(y_i \mid \lambda, \tau\right) di$$

- Government chooses (G, τ),
  ... and λ balances the budget residually
- Reversible investments $\Rightarrow$ immediate transition to st.st.
- Utilitarian planner (equal weights), discounts future generations by β
- Equilibrium: efficient public good provision, given Y
  Note: must correct externality caused by public goods

EQUILIBRIUM ALLOCATIONS

A special case: the representative agent

$$\begin{aligned}\max_{C,H} \quad U \quad &= \quad \log C - \frac{H^{1+\sigma}}{1+\sigma} + \chi \log gY \\ &s.t. \\ C \quad &= \quad \lambda Y^{1-\tau} \text{ and } Y = H\end{aligned}$$

A special case: the representative agent

$$\max_{C,H} \quad U = \log C - \frac{H^{1+\sigma}}{1+\sigma} + \chi \log gY$$

$$s.t.$$

$$C = \lambda Y^{1-\tau} \text{ and } Y = H$$

- Market clearing: $C + G = Y$

A special case: the representative agent

$$\begin{aligned}\max_{C,H} \quad U &= \log C - \frac{H^{1+\sigma}}{1+\sigma} + \chi \log gY \\ s.t.& \\ C &= \lambda Y^{1-\tau} \text{ and } Y = H\end{aligned}$$

- Market clearing: $C + G = Y$
- Equilibrium allocations:

$$\begin{aligned}\log H^{RA}(\tau) &= \frac{1}{1+\sigma}\log(1-\tau) \\ \log C^{RA}(g,\tau) &= \log(1-g) + \log H^{RA}(\tau)\end{aligned}$$

Optimal policy in the RA economy

- Welfare function:

$$\mathcal{W}^{RA}(g,\tau) = \log(1-g) + \chi \log g + (1+\chi)\frac{\log(1-\tau)}{1+\sigma} - \frac{1-\tau}{1+\sigma}$$

- Welfare maximizing (g,τ) pair:

$$g^* = \frac{\chi}{1+\chi}$$

$$\tau^* = -\chi$$

- Solution for g^* will extend to heterogeneous-agent setup

- Allocations are first best (same as with lump-sum taxes)

Recursive stationary equilibrium

- Given (G, τ), a stationary RCE is a value λ^*, asset prices $\{Q(\cdot), q\}$, skill prices $p(s)$, decision rules $s(\varphi, \kappa, 0)$, $c(\alpha, \varepsilon, \varphi, s, b)$, $h(\alpha, \varepsilon, \varphi, s, b)$, and aggregate quantities $N(s)$ such that:
  - households optimize
  - markets clear
  - the government budget constraint is balanced

Budget constraints

1. Beginning of period: innovation ω to α shock is realized

2. Middle of period: buy insurance against ε:

$$b = \int_E Q(\varepsilon)B(\varepsilon)d\varepsilon,$$

where $Q(\cdot)$ is the price of insurance and $B(\cdot)$ is the quantity

3. End of period: ε realized, consumption and hours chosen:

$$c + \delta q b' = \lambda \left[p(s)\exp(\alpha+\varepsilon)h\right]^{1-\tau} + B(\varepsilon)$$

Recursive stationary equilibrium

- Given (G, τ), a stationary RCE is a value λ^*, asset prices $\{Q(\cdot), q\}$, skill prices $p(s)$, decision rules $s(\varphi, \kappa, 0)$, $c(\alpha, \varepsilon, \varphi, s, b)$, $h(\alpha, \varepsilon, \varphi, s, b)$, and aggregate quantities $N(s)$ such that:
  - households optimize
  - markets clear
  - the government budget constraint is balanced
- In equilibrium, bonds are not traded
  - $b = 0 \rightarrow$ allocations depend only on exogenous states
  - α shocks remain uninsured, ε shocks fully insured

Skill prices and skill investment

- Guess: skill price has *Mincerian form*,
  $\log p(s) = \pi_0 + \pi_1 s(\kappa; \tau)$
- Optimal skill investment linear in κ, falling in τ:

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

- Assume $\psi = 1$. Analytical solution for equilibrium return to skill

$$\pi_1 = \frac{\rho}{2} + \sqrt{\left(\frac{\rho}{2}\right)^2 + \frac{1}{\theta\,(1-\tau)}}$$

- Implies a Stigliz g.e. effect: larger $\tau \Rightarrow$ larger π_1

Consumption and hours

$$\log c_{it} = \log \lambda + (1-\tau)\left[\frac{\log(1-\tau)}{1+\sigma} - \varphi_i + \log p(s_i) + \alpha_{it}\right] + \mathcal{C}$$

- Progressivity determines the pass-through of shocks/inequality

$$\log h_{it} = \frac{\log(1-\tau)}{1+\sigma} - \varphi_i + \left(\frac{1-\tau}{\sigma+\tau}\right)\varepsilon_{it} - \mathcal{H}$$

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

Note: insurable productivity shocks (ε_{it}) enters h_{it} but not c_{it}

Optimal taxation depends on source of inequality

- Analytical characterization of social welfare function
- Larger uninsurable risk v_α implies $\tau^* \uparrow$
- Larger insurable risk v_ε implies $\tau^* \downarrow$
- Lareger skill price dispersion: ambiguous effect on τ^*:
  - If $\psi = 0$ (zero elasticity of skill choice to $p(s)$) then skill price dispersion is like uninsurable risk
  - If $\psi > 0$ then tradeoff between insurance and distortion of skill choice

Exact expression for SWF

$$
\begin{aligned}
\mathcal{W}(g,\tau) \quad = \quad & \log(1+g)+\chi\log g+(1+\chi)\frac{\log(1-\tau)}{(1+\hat{\sigma})(1-\tau)}-\frac{1}{(1+\hat{\sigma})} \\
& +(1+\chi)\left[\frac{-1}{\theta-1}\log\left(\sqrt{\frac{\eta\theta}{\mu\,(1-\tau)}}\right)+\frac{\theta}{\theta-1}\log\left(\frac{\theta}{\theta-1}\right)\right] \\
& -\frac{1}{2\theta}(1-\tau)-\left[-\log\left(1-\left(\frac{1-\tau}{\theta}\right)\right)-\left(\frac{1-\tau}{\theta}\right)\right] \\
& -(1-\tau)^2\,\frac{v_\varphi}{2} \\
& -\left[(1-\tau)\,\frac{\delta}{1-\delta}\,\frac{v_\omega}{2}-\log\left(\frac{1-\delta\exp\left(\frac{-\tau(1-\tau)}{2}v_\omega\right)}{1-\delta}\right)\right] \\
& -(1+\chi)\sigma\frac{1}{\hat{\sigma}^2}\,\frac{v_\varepsilon}{2}+(1+\chi)\frac{1}{\hat{\sigma}}v_\varepsilon
\end{aligned}
$$

Representative Agent component

$$\mathcal{W}(g,\tau) = \underbrace{\log(1+g) + \chi\log g + (1+\chi)\frac{\log(1-\tau)}{(1+\hat{\sigma})(1-\tau)} - \frac{1}{(1+\hat{\sigma})}}_{\text{Representative Agent Welfare} = \mathcal{W}^{RA}(g,\tau)}$$

$$+(1+\chi)\left[\frac{-1}{\theta-1}\log\left(\sqrt{\frac{\eta\theta}{\mu(1-\tau)}}\right) + \frac{\theta}{\theta-1}\log\left(\frac{\theta}{\theta-1}\right)\right]$$

$$-\frac{1}{2\theta}(1-\tau) - \left[-\log\left(1-\left(\frac{1-\tau}{\theta}\right)\right) - \left(\frac{1-\tau}{\theta}\right)\right]$$

$$-(1-\tau)^2\frac{v_\varphi}{2}$$

$$-\left[(1-\tau)\frac{\delta}{1-\delta}\frac{v_\omega}{2} - \log\left(\frac{1-\delta\exp\left(\frac{-\tau(1-\tau)}{2}v_\omega\right)}{1-\delta}\right)\right]$$

$$-(1+\chi)\sigma\frac{1}{\hat{\sigma}^2}\frac{v_\varepsilon}{2} + (1+\chi)\frac{1}{\hat{\sigma}}v_\varepsilon$$

Exact expression for SWF

$$
\begin{aligned}
\mathcal{W}(\tau) \quad = \quad & \chi \log \chi - (1+\chi)\log(1+\chi) + (1+\chi)\frac{\log(1-\tau)}{(1+\hat{\sigma})(1-\tau)} - \frac{1}{(1+\hat{\sigma})} \\
& +(1+\chi)\left[\frac{-1}{\theta-1}\log\left(\sqrt{\frac{\eta\theta}{\mu\,(1-\tau)}}\right) + \frac{\theta}{\theta-1}\log\left(\frac{\theta}{\theta-1}\right)\right] \\
& -\frac{1}{2\theta}(1-\tau) - \left[-\log\left(1-\left(\frac{1-\tau}{\theta}\right)\right) - \left(\frac{1-\tau}{\theta}\right)\right] \\
& -(1-\tau)^2\,\frac{v_\varphi}{2} \\
& -\left[(1-\tau)\,\frac{\delta}{1-\delta}\,\frac{v_\omega}{2} - \log\left(\frac{1-\delta\exp\left(\frac{-\tau(1-\tau)}{2}v_\omega\right)}{1-\delta}\right)\right] \\
& -(1+\chi)\sigma\frac{1}{\hat{\sigma}^2}\frac{v_\varepsilon}{2} + (1+\chi)\frac{1}{\hat{\sigma}}v_\varepsilon
\end{aligned}
$$

Skill investment component

$$
\begin{aligned}
\mathcal{W}(\tau) \quad = \quad & \chi \log \chi - (1+\chi)\log(1+\chi) + (1+\chi)\frac{\log(1-\tau)}{(1+\hat{\sigma})(1-\tau)} - \frac{1}{(1+\hat{\sigma})} \\
& +(1+\chi)\underbrace{\left[\frac{-1}{\theta-1}\log\left(\sqrt{\frac{\eta\theta}{\mu(1-\tau)}}\right) + \frac{\theta}{\theta-1}\log\left(\frac{\theta}{\theta-1}\right)\right]}_{\text{productivity gain} = \log E[(p(s))] = \log(Y/N)} \\
& \underbrace{-\frac{1}{2\theta}(1-\tau)}_{\text{avg. education cost}} - \underbrace{\left[-\log\left(1-\left(\frac{1-\tau}{\theta}\right)\right) - \left(\frac{1-\tau}{\theta}\right)\right]}_{\text{consumption dispersion across skills}} \\
& -(1-\tau)^2\frac{v_\varphi}{2} \\
& - \left[(1-\tau)\frac{\delta}{1-\delta}\frac{v_\omega}{2} - \log\left(\frac{1-\delta\exp\left(\frac{-\tau(1-\tau)}{2}v_\omega\right)}{1-\delta}\right)\right] \\
& -(1+\chi)\sigma\frac{1}{\hat{\sigma}^2}\frac{v_\varepsilon}{2} + (1+\chi)\frac{1}{\hat{\sigma}}v_\varepsilon
\end{aligned}
$$

Uninsurable component

$$
\begin{aligned}
\mathcal{W}(\tau) \;=\;& \chi \log \chi - (1+\chi)\log(1+\chi) + (1+\chi)\frac{\log(1-\tau)}{(1+\hat{\sigma})(1-\tau)} - \frac{1}{(1+\hat{\sigma})} \\
&+(1+\chi)\left[\frac{-1}{\theta-1}\log\left(\sqrt{\frac{\eta\theta}{\mu\,(1-\tau)}}\right) + \frac{\theta}{\theta-1}\log\left(\frac{\theta}{\theta-1}\right)\right] \\
&-\frac{1}{2\theta}(1-\tau) - \left[-\log\left(1-\left(\frac{1-\tau}{\theta}\right)\right) - \left(\frac{1-\tau}{\theta}\right)\right] \\
&- \underbrace{(1-\tau)^2\,\frac{v_\varphi}{2}}_{\text{cons. disp. due to prefs}} \\
&- \underbrace{\left[(1-\tau)\,\frac{\delta}{1-\delta}\,\frac{v_\omega}{2} - \log\left(\frac{1-\delta\exp\left(\frac{-\tau(1-\tau)}{2}v_\omega\right)}{1-\delta}\right)\right]}_{\text{consumption dispersion due to uninsurable shocks} \,\approx\, (1-\tau)^2\frac{v_\alpha}{2}} \\
&-(1+\chi)\sigma\frac{1}{\hat{\sigma}^2}\frac{v_\varepsilon}{2} + (1+\chi)\frac{1}{\hat{\sigma}}v_\varepsilon
\end{aligned}
$$

Insurable component

$$
\begin{aligned}
\mathcal{W}(\tau) \quad = \quad & \chi \log \chi - (1+\chi)\log(1+\chi) + (1+\chi)\frac{\log(1-\tau)}{(1+\hat{\sigma})(1-\tau)} - \frac{1}{(1+\hat{\sigma})} \\
& +(1+\chi)\left[\frac{-1}{\theta-1}\log\left(\sqrt{\frac{\eta\theta}{\mu(1-\tau)}}\right) + \frac{\theta}{\theta-1}\log\left(\frac{\theta}{\theta-1}\right)\right] \\
& -\frac{1}{2\theta}(1-\tau) - \left[-\log\left(1-\left(\frac{1-\tau}{\theta}\right)\right) - \left(\frac{1-\tau}{\theta}\right)\right] \\
& -(1-\tau)^2\,\frac{v_\varphi}{2} \\
& -\left[(1-\tau)\frac{\delta}{1-\delta}\frac{v_\omega}{2} - \log\left(\frac{1-\delta\exp\left(\frac{-\tau(1-\tau)}{2}v_\omega\right)}{1-\delta}\right)\right] \\
& -(1+\chi)\sigma \underbrace{\frac{1}{\hat{\sigma}^2}\frac{v_\varepsilon}{2}}_{\text{hours dispersion}} + (1+\chi)\underbrace{\frac{1}{\hat{\sigma}}v_\varepsilon}_{\text{prod. gain from ins. shock}=\log(N/H)}
\end{aligned}
$$

Quantitative evaluation

Parameterization

- Parameter vector $\{\chi, \sigma, \theta, v_{\varphi}, v_{\omega}, v_{\varepsilon}\}$

Parameterization

- Parameter vector $\{\chi, \sigma, \theta, v_\varphi, v_\omega, v_\varepsilon\}$
- To match $G/Y = 0.19$: $\rightarrow \chi = 0.23$

Parameterization

- Parameter vector $\{\chi, \sigma, \theta, v_\varphi, v_\omega, v_\varepsilon\}$
- To match $G/Y = 0.19$: $\rightarrow \chi = 0.23$
- Frisch elasticity (micro-evidence): $\rightarrow \sigma = 2$

Parameterization

- Parameter vector $\{\chi, \sigma, \theta, v_\varphi, v_\omega, v_\varepsilon\}$
- To match $G/Y = 0.19$: $\rightarrow \chi = 0.23$
- Frisch elasticity (micro-evidence): $\rightarrow \sigma = 2$

$$\begin{aligned}
cov(\log h, \log w) &= \frac{1}{\hat{\sigma}} v_\varepsilon \\
var(\log h) &= v_\varphi + \frac{1}{\hat{\sigma}^2} v_\varepsilon \\
var^0(\log c) &= (1-\tau)^2 \left(v_\varphi + \frac{1}{\theta^2} \right) \\
var(\log w) &= \frac{1}{\theta^2} + \frac{\delta}{1-\delta} v_\omega
\end{aligned}$$

Parameterization

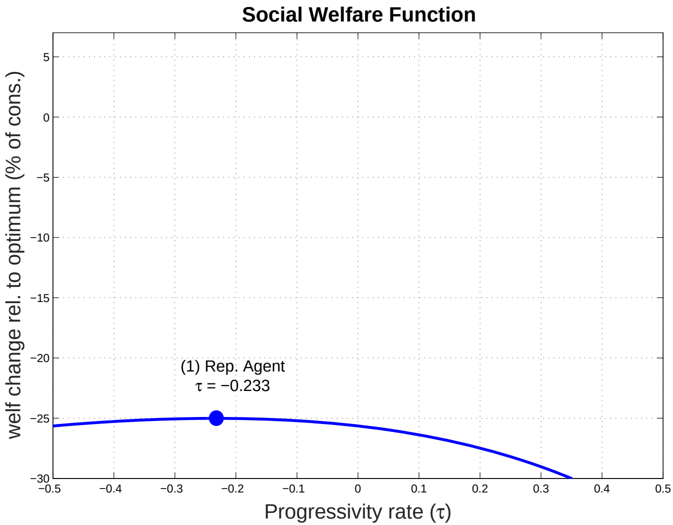
- Parameter vector $\{\chi, \sigma, \theta, v_\varphi, v_\omega, v_\varepsilon\}$
- To match $G/Y = 0.19$: $\rightarrow \chi = 0.23$
- Frisch elasticity (micro-evidence): $\rightarrow \sigma = 2$

$$
\begin{aligned}
cov(\log h, \log w) &= \frac{1}{\hat{\sigma}} v_\varepsilon & \rightarrow v_\varepsilon = 0.17 \\
var(\log h) &= v_\varphi + \frac{1}{\hat{\sigma}^2} v_\varepsilon & \rightarrow v_\varphi = 0.035 \\
var^0(\log c) &= (1-\tau)^2 \left(v_\varphi + \frac{1}{\theta^2} \right) & \rightarrow \theta = 2.16 \\
var(\log w) &= \frac{1}{\theta^2} + \frac{\delta}{1-\delta} v_\omega & \rightarrow v_\omega = 0.003
\end{aligned}
$$

Start with zero heterogeneity

- Only distortion: labor supply
- Trade off distortion with desire to fund public good
- Optimal policy: $\tau^* = -\chi$
- First best is regressive: can be implemented with lump sum tax and zero marginal tax
  $\Rightarrow$ marginal tax < avg. tax

Optimal progressivity: zero heterogeneity



Skill price dispersion

- Skill price dispersion in model:

$$var\left(\ln(p(s))\right) = \frac{1}{\theta^2}$$

- Data: in 1980, between-group inequality was 0.051,

$$\frac{1}{\theta_{1980}^2} = 0.051 \Rightarrow \theta_{1980} = 4.42$$

- Identify skill-investment price elasticity $\psi \approx 1$ ... by assuming that all increase in avg. educational attainment after 1980 is due to increase in (Mincer) return to education

Optimal progressivity: skill price dispersion

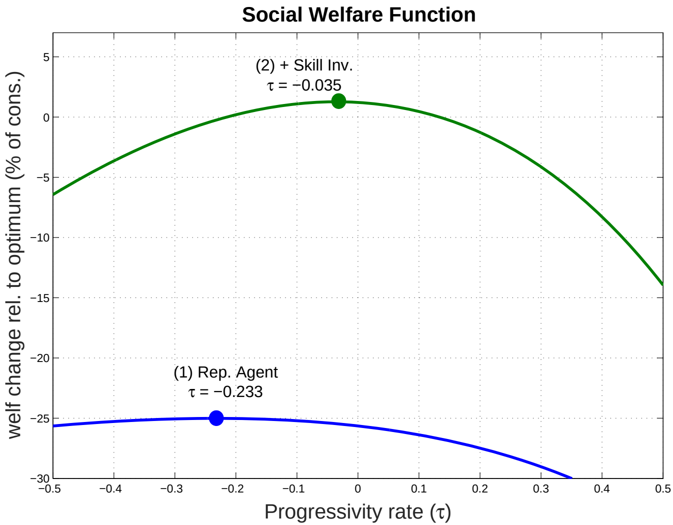
- Recall: $s(\kappa) = [(1-\tau)\,\pi_1]^{\psi} \cdot \kappa$
- Estimate ψ from changes in π_1 and average s

$$\ln\left(\frac{E[s_{2016}]}{E[s_{1980}]}\right) = \psi \ln\left(\frac{(1-\tau_{2016})\pi_{2016}}{(1-\tau_{1980})\pi_{1980}}\right)$$

- For 26-30 years old individuals:
  - average years of education: 12.6 in 1980 and 14.0 in 2016
  - Mincer return to education: $\pi_{1980} = 8.2\%$ and $\pi_{2016} = 11.7\%$

$$\psi = \ln\left(\frac{14.0-10}{12.6-10}\right) / \ln\left(\frac{(1-0.186)\cdot 0.117}{(1-0.186)\cdot 0.082}\right) = 1.21$$

Optimal progressivity: skill price dispersion



Preference heterogeneity

- Recall: consumption & hours worked

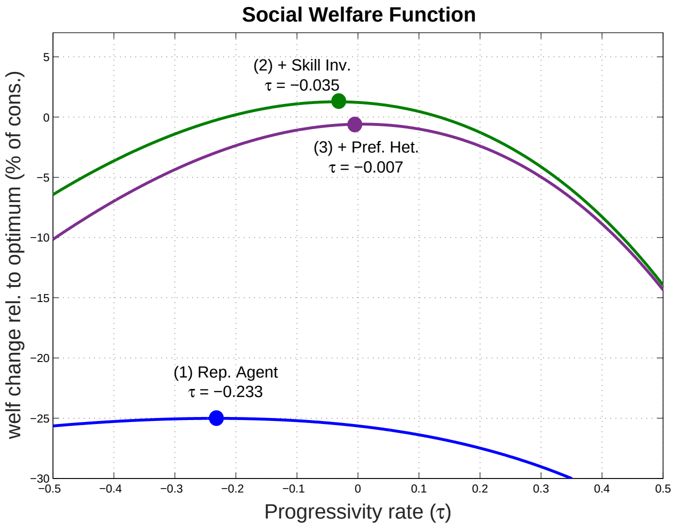
$$\begin{aligned} \ln h_{it} &= -\varphi_i + \frac{1-\tau}{\sigma+\tau} \cdot \varepsilon_{it} + const. \\ \ln c_{it} &= (1-\tau)\left[\alpha_{it} - \varphi_i + \ln p(s_i)\right] + const. \end{aligned}$$

Question: How can $var(\varphi_i)$ be identified? Answer:

$$cov(\ln c_{it}, \ln h_{it}) = (1-\tau_{1980}) \cdot var(\varphi_i)$$

- CEX: $cov(\ln c_{it}, \ln h_{it}) \approx 0.034$ in both 1980 and 2016
  $\Rightarrow \; var(\varphi) = 0.043$
  - Accounts for 3/4 of $var(\ln h_{it})$

Optimal progressivity: preference heterogeneity



Uninsurable risk

- Uninsurable risk $var\,(\alpha)$ increases optimal progressivity

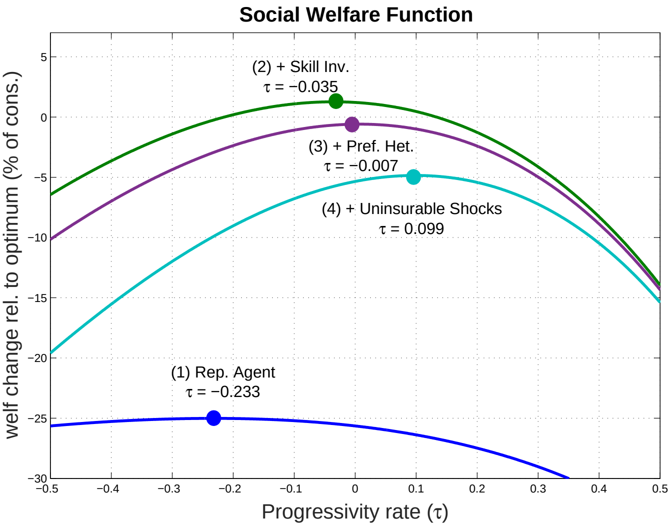
Question: How can $var\,(\alpha)$ be identified?

Answer: From dispersion of consumption

- Blundell-Preston QJE 1998; Heathcote-Storesletten-Violante AER 2014

$$
\begin{aligned}
var\,(\ln c) &= (1-\tau)^2\left[var\,(\alpha) + var\,(\varphi) + var\,(\ln p\,(s))\right] \\
&\Rightarrow \\
var\,(\alpha) &= \frac{var\,(\ln c)}{(1-\tau)^2} - var\,(\varphi) - var\,(\ln p\,(s)) \\
&= \frac{0.215}{(1-0.186)^2} - 0.043 - 0.051 = 0.23
\end{aligned}
$$

Optimal progressivity: uninsurable risk



Insurable risk

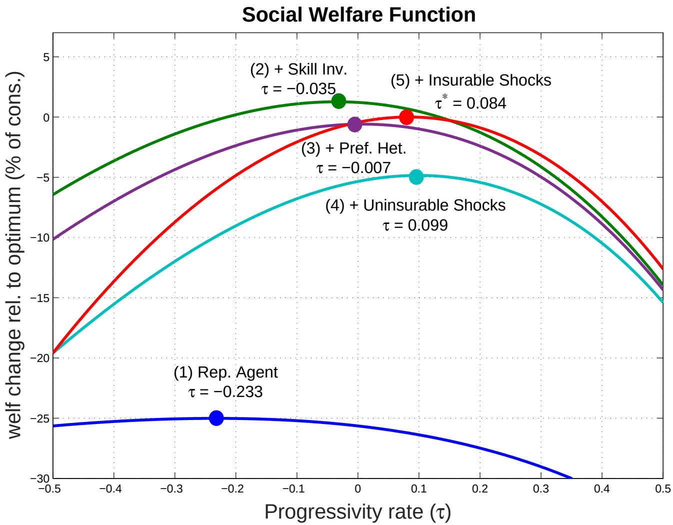
Question: How can insurable risk be identified?

Answer: Determined residually:
within-group inequality minus uninsurable risk

$$\begin{aligned} var(\varepsilon) &= var(\text{within group inequality}) - var(\alpha) \\ &= 0.26 - 0.23 = 0.03 \end{aligned}$$

- 1/8 of within-group inequality in 1980 was insurable

Optimal progressivity: insurable risk



Thought experiment

- How should progressivity respond to rising inequality?
- Thought experiment: assume that tax progressivity of 0.181 in U.S. was optimal (close to empirical $\tau_{1980}^{US} = 0.186$)
- Study how each driver of inequality from 1980 to 2016 changes τ^* – one factor at a time
- Recall: wage inequality incrased by 0.15;
  - between-group inequality: up by 0.05
  - within-group inequality: up by 0.10

Increased skill premium

- How should taxes repsond to increased skill premium? (skill-biased technical change)
- Baseline: SBTC stems from change in θ
- 1980-2016 increase in between-group inequality: 0.05

$$\begin{aligned} \frac{1}{(\theta_{2010})^2} - \frac{1}{(\theta_{1980})^2} &= 0.05 \\ &\Rightarrow \\ \theta_{2016} &= 3.1 \end{aligned}$$

- Implication: optimal progressivity $\tau^*_{2016} = 0.161$

- Conclusion: τ^* should have fallen by 0.02

Increased uninsurable risk

- Within-group wage dispersion increased by 0.10

Question: how much accounted for by $\Delta var\,(\alpha)$?

- Identify increase in uninsurable risk $\Delta var\,(\alpha)$ from increase in consumption dispersion
- CEX: $\Delta var(\log(c)) = 0.06$ (use Heathcote et al. 2010 data)
- Result: $\Delta var(\alpha) = 0.032$

- Result: $\Delta var(\alpha) = 0.032$

- Implication: optimal progressivity $\tau^*_{2016} = 0.201$

■ Conclusion: τ^* should have increased by 0.02

Increased insurable risk:

- 1980-2016: Residual wage dispersion increased by 0.10
- Increase in insurable risk set residually:

$$\Delta var\left(\varepsilon\right) = 0.10 - 0.032 = 0.068$$

$\Rightarrow$ 2/3 of increase in residual inequality was insurable

- Implication: optimal progressivity $\tau^*_{2016} = 0.167$

■ Conclusion: τ^* should have fallen by 0.014

Overall change

- Consider all changes simultaneously
  (Δskill price dispersion, Δuninsurable risk, Δinsurable risk)

- Implication: optimal progressivity $\tau^*_{2016} = 0.169$

- Conclusion: when all channels are active, τ^* should have fallen slightly by 0.012 between 1980 and 2016

- ... approximately in line with the empirical changes in U.S. progressivity

Robustness analysis

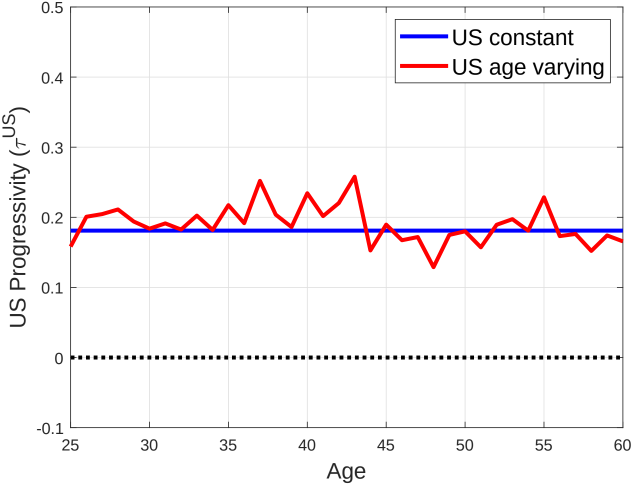
- alternative model of Skill-Biased Technical Change (increase in absolute skill advantage ρ): optimal response is to lower τ^* to 0.161
- exogenous skills (no human capital distortion): $\Delta\tau^* = +0.033$
- low skill investment elasticity $\psi = 0.5$: $\Delta\tau^* = +0.003$
- larger increase in consumption inequality (Attanasio et al. 2007): $\Delta var(\log c) = +0.09$ instead of 0.06 (HPV) implies $\Delta\tau^* = +0.007$

Conclusion

- How should taxes respond to rising inequality?
- Answer depends on the source of rise in inequality
    - if increase is due to uninsurable residual risk: increase τ magnitude is disciplined by rise in $var(\log c)$
    - if increase is due to privately insurable residual risk OR Skill-Biased technical Change: lower τ
    - size of effect hinges on the size of the human capital distortion
- Quantitative exercise (and equally-weighted planenr) suggests that progressivity should have remained approximately constant, in line with the actual evolution of tax progressivity in U.S.

ADDITIONAL INSTRUMENTS

Progressivity over life cycle



Progressivity over life cycle

Heathcote, Storesletten, and Violante (2019): age-dependent taxation

- Suppose log-linear tax system vary with age: $\{\lambda_a, \tau_a\}_{a=1}^{A}$

Finding 1: large potential gains relative to no age variation

Finding 2: optimal progressicity τ u-shaped over life cycle
... driven by two opposing forces:

  - uninsurable dispersion increases with age $\Rightarrow \tau_a \uparrow$ with a
  - age-profiles for productivity (hours) are hump-shaped (flat) $\Rightarrow \tau_a$ u-shaped in age a

- Results hold up in calibrated model with financial wealth
... but uninsurable channel dominates if borrowing is generous

History-dependent taxation

- Dynamic optimal taxation generally depends on one's full income history (Golosov et al. 2003)
- Kapicka (2017): problem simplifies if tax system is log-linear
- Optimal history dependent taxes: policy:
  - High marginal tax on current income
  - Negative marginal tax on past incomes
  - ... but zero history dependence if shocks are permanent

Human capital investments

- Optimal taxation changes when skill investments are endogenous
  - Stancheva (2017); Badel, Huggett, Luo (2019); HSV
- Education subsidies: complement to progressive taxes

  $\Rightarrow$ need more education and training subsidies when tax progressivity is larger
  - Krueger and Ludwig 2016; Findeisen and Sachs 2016

Individual vs. joint taxation

- Some countries have joint taxation (USA, Germany) ... while others have separate taxation (Scandinavia)
- Joint taxation has large negative distortion on labor-force participation for secondary earner
- Moving to separate taxation can have large positive effects on female employment
  - Guner, Kaygusuz, and Ventura (2011)
- Employment differences Continental Europe - USA - Scandinavia explained by differences in joint vs. separate taxation
  - Bick and Fuchs-Schuendeln 2017; Chakraborty et al. 2015
- Message: high avg. taxes require efficient tax systems ...holds true in OECD data!

Next Steps

What caused rising inequality in the first place?

- Appropriate policy prescription requires correct diagnose
- Explanation must be consistent with
  - Large increase in inequality in Anglo-Saxon countries after 1970s
  - Minor increase in inequality in Continental Europe, Scandinavia, Japan
- The usual suspects:
  - Trade exposure, China shock
  - Technical change
  - The culprit must be more subtle
    - Some interaction between human capital policies and trade/technology trends?
    - Luck?
    - Norms and institutions?