Optimal Tax Progressivity: An Analytical Framework

Jonathan Heathcote
Federal Reserve Bank of Minneapolis

Kjetil Storesletten
Oslo University and Federal Reserve Bank of Minneapolis

Gianluca Violante
New York University
Motivation

The hardest thing in the world to understand is income taxes.

(Albert Einstein)
How progressive should labor income taxation be?
How progressive should labor income taxation be?

- Argument in favor of progressivity: missing markets
  - Social insurance of privately-uninsurable lifecycle shocks
  - Redistribution with respect to unequal initial conditions
How progressive should labor income taxation be?

- Argument in favor of progressivity: missing markets
  - Social insurance of privately-uninsurable lifecycle shocks
  - Redistribution with respect to unequal initial conditions

- Argument I against progressivity: distortions
  - Labor supply
  - Human capital investment

Heathcote-Storesletten-Violante, “Optimal Tax Progressivity”
How progressive should labor income taxation be?

• Argument in favor of progressivity: missing markets
  ▶ Social insurance of privately-uninsurable lifecycle shocks
  ▶ Redistribution with respect to unequal initial conditions

• Argument I against progressivity: distortions
  ▶ Labor supply
  ▶ Human capital investment

• Argument II against progressivity: externality
  ▶ Financing of public good provision

Heathcote-STORESLETTS-VIOLANTE, “Optimal Tax Progressivity”
Overview of the approach

• Model ingredients:

  1. partial insurance against labor-income risk [ex-post heter.]
  2. differential diligence & (learning) ability [ex-ante heter.]
Overview of the approach

Model ingredients:

1. partial insurance against labor-income risk [ex-post heter.]
2. differential diligence & (learning) ability [ex-ante heter.]
3. flexible labor supply
4. endogenous skill investment + multiple-skill technology
Overview of the approach

- **Model ingredients:**

1. partial insurance against labor-income risk  \[\text{[ex-post heter.]}\]
2. differential diligence & (learning) ability  \[\text{[ex-ante heter.]}\]
3. flexible labor supply
4. endogenous skill investment + multiple-skill technology
5. government expenditures valued by households
Overview of the approach

- **Model ingredients:**
  1. partial insurance against labor-income risk [ex-post heter.]
  2. differential diligence & (learning) ability [ex-ante heter.]
  3. flexible labor supply
  4. endogenous skill investment + multiple-skill technology
  5. government expenditures valued by households

- **Ramsey approach:** mkt structure & tax instruments taken as given
Overview of the approach

- Model ingredients:

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

- Ramsey approach: mkt structure & tax instruments taken as given

  $\rightarrow$ closed-form Social Welfare Function
TAX/TRANSFER FUNCTION
The tax/transfer function

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

- The parameter \( \tau \) measures the degree of progressivity:
  - \( \tau = 1 \): full redistribution \( T(y) = y - \lambda \)
  - \( 0 < \tau < 1 \): progressivity \( T'(y) > \frac{T(y)}{y} \)
  - \( \tau = 0 \): no redistribution \( T'(y) = \frac{T(y)}{y} = 1 - \lambda \)
  - \( \tau < 0 \): regressivity \( T'(y) < \frac{T(y)}{y} \)

- Break-even income level: \( y^0 = \lambda^{\frac{1}{\tau}} \)
The tax/transfer function

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

- The parameter \( \tau \) 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
Measurement of $\tau^{US}$

- PSID 2000-06, age of head of hh 25-60, $N = 12,943$

- Pre gov. income: income minus deductions (medical expenses, state taxes, mortgage interest and charitable contributions)

- Post-gov income: ... minus taxes (TAXSIM) plus transfers
Measurement of $\tau^{US}$

- PSID 2000-06, age of head of hh 25-60, $N = 12,943$

- Pre gov. income: income minus deductions (medical expenses, state taxes, mortgage interest and charitable contributions)

- Post-gov income: ... minus taxes (TAXSIM) plus transfers

Heathcote-Storesletten-Violante, "Optimal Tax Progressivity"
MODEL
Demographics and preferences

- **Perpetual youth** demographics with constant survival probability $\delta$

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

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

Heathcote-Storesletten-Violante, "Optimal Tax Progressivity"
Demographics and preferences

- **Perpetual youth** demographics with constant survival probability $\delta$
- **Preferences** over consumption ($c$), hours ($h$), publicly-provided goods ($G$), and skill-investment ($s$) effort:

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

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

\[
\kappa_i \sim \text{Exp}(1)
\]

\[
u_i (c_{it}, h_{it}, G) = \log c_{it} - \exp(\varphi_i) \frac{h_{it}^{1+\sigma}}{1 + \sigma} + \chi \log G
\]

\[
\varphi_i \sim \mathcal{N}\left(\frac{v_{\varphi}}{2}, v_{\varphi}\right), \quad \varphi_i \perp \kappa_i
\]
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:

\[
Y = \left[ \int_{0}^{\infty} N(s) \frac{\theta - 1}{\theta} \, ds \right] \frac{\theta}{\theta - 1}, \quad \theta \in (1, \infty)
\]

- **Determination of skill price:** \( p(s) = MPN(s) \)

- **Aggregate resource constraint:**

\[
Y = \int_{0}^{1} c_i \, di + G
\]

Heathcote-Storesletten-Violante, "Optimal Tax Progressivity"
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]
- \( \varepsilon_{it} \) i.i.d. over time with \( \varepsilon_{it} \sim \mathcal{N} \left( -\frac{v_\varepsilon}{2}, v_\varepsilon \right) \) [transitory]
- \( \omega_{it} \perp \varepsilon_{it} \) cross-sectionally and longitudinally

Heathcote-Storesletten-Violante, "Optimal Tax Progressivity"
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]
  +
- \( \varepsilon_{it} \) i.i.d. over time with \( \varepsilon_{it} \sim \mathcal{N} \left( -\frac{v_\varepsilon}{2}, v_\varepsilon \right) \) [transitory]
- \( \omega_{it} \perp \varepsilon_{it} \) cross-sectionally and longitudinally
- Pre-government earnings:
  \[
y_{it} = p(s_i) \times \exp(\alpha_{it} + \varepsilon_{it}) \times h_{it}
  \]
  determined by skill, fortune, and diligence
Government

- Government budget constraint (no government debt):
  \[ G = \int_0^1 \left[ y_i - \lambda y_i^{1-\tau} \right] di \]

- Government chooses \((G, \tau)\), and \(\lambda\) balances the budget residually

- Without loss of generality, we let the government choose:
  \[ g \equiv \frac{G}{\overline{Y}} \]
Market structure

- Final good (numeraire) market and labor markets are competitive
- Perfect annuity markets against survival risk
Market structure

- Final good (numeraire) market and labor markets are competitive
- Perfect annuity markets against survival risk
- Full set of insurance claims against $\varepsilon$ shocks
- No market to insure $\omega$ shock [microfoundation with bond]
Market structure

• Final good (numeraire) market and labor markets are competitive

• Perfect annuity markets against survival risk

• Full set of insurance claims against $\varepsilon$ shocks

• No market to insure $\omega$ shock  [microfoundation with bond]

- $v_\varepsilon > 0, v_\omega > 0 \rightarrow \text{partial insurance economy}$

- $v_\omega = 0 \rightarrow \text{full insurance economy}$

- $v_\omega = v_\varepsilon = v_\varphi = 0 \ & \ \theta = \infty \rightarrow \text{RA economy}$
Special case: representative agent economy

\[
\max_{C,H} \quad U = \log C - \frac{H^{1+\sigma}}{1 + \sigma} + \chi \log gY \\
\text{s.t.} \\
C = \lambda Y^{1-\tau} \\
Y = H \\
C + G = Y
\]
Special case: representative agent economy

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

s.t.

\[
C = \lambda Y^{1-\tau}
\]

\[
Y = H
\]

\[
C + G = Y
\]

- Substitute equilibrium allocations into \( U \) to obtain:

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

- Ramsey planner chooses \((g, \tau)\) to maximize \(\mathcal{W}^{RA}\)
Optimal policy in the RA economy

\[ g^* = \frac{\chi}{1 + \chi} \]

- Samuelson condition: \( MRS_{C,G} = MRT_{C,G} = 1 \)
- This result will extend to the general model

Heathcote-Storesletten-Violante, "Optimal Tax Progressivity"
Optimal policy in the RA economy

\[ g^* = \frac{\chi}{1 + \chi} \]

- Samuelson condition: \( MRS_{C,G} = MRT_{C,G} = 1 \)

- This result will extend to the general model

\[ \tau^* = -\chi \]

- Regressivity corrects the externality linked to valued G

- Allocations are first best, i.e., same as with lump-sum taxation
Equilibrium skill choice and skill price
Equilibrium skill choice and skill price

- Skill price has **Mincerian shape**: \( \log p(s; \tau) = \pi_0(\tau) + \pi_1(\tau)s(\kappa; \tau) \)

\[
s(\kappa; \tau) = \left( \frac{1 - \tau}{\theta} \right)^{\frac{\psi}{1+\psi}} \cdot \kappa \quad \text{skill choice}
\]
\[
\pi_1(\tau) = \left( \frac{1}{\theta} \right)^{\frac{1}{1+\psi}} (1 - \tau)^{-\frac{\psi}{1+\psi}} \quad \text{marginal return to skill}
\]
Equilibrium skill choice and skill price

• Skill price has **Mincerian shape**: \( \log p(s; \tau) = \pi_0(\tau) + \pi_1(\tau)s(\kappa; \tau) \)

\[
s(\kappa; \tau) = \left( \frac{1 - \tau}{\theta} \right)^{\frac{\psi}{1 + \psi}} \cdot \kappa \quad \text{skill choice}
\]

\[
\pi_1(\tau) = \left( \frac{1}{\theta} \right)^{\frac{1}{1 + \psi}} (1 - \tau)^{-\frac{\psi}{1 + \psi}} \quad \text{marginal return to skill}
\]

• **Direct effect**: \( \tau \) reduces skill accumulation

• **Equilibrium (Stiglitz) effect**: \( \tau \) raises skill premium through scarcity

\[
\text{Neutrality} \rightarrow \text{var}(\log p(s; \tau)) = \frac{1}{\theta^2}
\]
Equilibrium skill choice and skill price

- Skill price has **Mincerian shape**: \( \log p(s; \tau) = \pi_0(\tau) + \pi_1(\tau) s(\kappa; \tau) \)

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

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

- Direct effect: \( \tau \) reduces skill accumulation
- Equilibrium (Stiglitz) effect: \( \tau \) raises skill premium through scarcity

**Neutrality**

\[
\text{Neutrality} \rightarrow \text{var}(\log p(s; \tau)) = \frac{1}{\theta^2}
\]

- Distribution of skill prices \( p \) is **Pareto** with parameter \( \theta \)
Equilibrium consumption and hours allocation

\[
\log c(\alpha, \varphi, s; g, \tau) = \log C^{RA}(g, \tau) + (1 - \tau) \log p(s; \tau) \\
+ (1 - \tau) \alpha - (1 - \tau) \varphi + M(v_\varepsilon; \tau)
\]

Heathcote-Storesletten-Violante, "Optimal Tax Progressivity"
Equilibrium consumption and hours allocation

\[
\log c(\alpha, \varphi, s; g, \tau) = \log C^{RA}(g, \tau) + (1 - \tau) \log p(s; \tau)
\]

\[
\quad + (1 - \tau) \alpha - (1 - \tau) \varphi + \underline{\mathcal{M}(v_{\varepsilon}; \tau)}
\]

\[
\text{skill price} \quad \text{unins. shock} \quad \text{pref. het.} \quad \text{welf. gain from ins. variation}
\]

\[
\log h(\varepsilon, \varphi; \tau) = \log H^{RA}(\tau) - \varphi + \frac{1}{\hat{\sigma}} \varepsilon - \frac{1}{\hat{\sigma}(1 - \tau)} \mathcal{M}(v_{\varepsilon}; \tau)
\]

\[
\text{pref. het.} \quad \text{ins. shock} \quad \text{welf. gain from ins. variation}
\]

\[
\frac{1}{\hat{\sigma}} := \frac{1 - \tau}{\sigma + \tau} \text{ is the tax-modified Frisch elasticity}
\]
SOCIAL WELFARE FUNCTION
Social Welfare Function

Economy is in steady-state with pair \((g_{-1}, \tau_{-1})\)

Planner chooses, once and for all, a new pair \((g^*, \tau^*)\)

We make two assumptions:

1. Planner puts equal weight on all currently alive agents, discounts \(U\) of future cohorts at rate \(\beta\)

2. Skill investments are reversible
Social Welfare Function

Economy is in steady-state with pair \((g_{-1}, \tau_{-1})\)

Planner chooses, once and for all, a new pair \((g^*, \tau^*)\)

We make two assumptions:

1. Planner puts equal weight on all currently alive agents, discounts \(U\) of future cohorts at rate \(\beta\)

2. Skill investments are reversible

   ▶ SWF becomes average period-utility in the cross-section

   ▶ \(\tau^*\) does not depend on the pre-existing skill distribution

   ▶ The transition to the new steady-state is instantaneous
The exact expression for SWF is given by:

\[ \mathcal{W}(g, \tau) = \log(1 + g) + \chi \log g + (1 + \chi) \frac{\log(1 - \tau)}{(1 + \hat{\sigma})(1 - \tau)} - \frac{1}{(1 + \hat{\sigma})} \]

\[ + (1 + \chi) \left( \frac{\psi}{1 + \psi} \right) \frac{1}{\theta - 1} \log(1 - \tau) \]

\[ - \left( \frac{\psi}{1 + \psi} \right) \frac{1}{\theta} (1 - \tau) - \left[ - \log \left( 1 - \left( \frac{1 - \tau}{\theta} \right) \right) - \left( \frac{1 - \tau}{\theta} \right) \right] \]

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

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

\[ + (1 + \chi) \frac{1}{\hat{\sigma}} v_\varepsilon - (1 + \chi)\sigma \frac{1}{\hat{\sigma}^2} \frac{v_\varepsilon}{2} \]
Representative Agent component

\[ \mathcal{W}(g, \tau) = \log(1 + g) + \chi \log g + (1 + \chi) \frac{\log(1 - \tau)}{(1 + \hat{\sigma})(1 - \tau)} - \frac{1}{(1 + \hat{\sigma})} \]

Representative Agent Welfare = \( \mathcal{W}^{RA}(g, \tau) \)

\[ + (1 + \chi) \left( \frac{\psi}{1 + \psi} \right) \frac{1}{\theta - 1} \log (1 - \tau) \]

\[ - (\frac{\psi}{1 + \psi}) \frac{1}{\theta} (1 - \tau) - \left[ \frac{\log \left( 1 - \left( \frac{1 - \tau}{\theta} \right) \right) - \left( \frac{1 - \tau}{\theta} \right) }{1 - \tau} \right] \]

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

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

\[ + (1 + \chi) \frac{1}{\hat{\sigma}} v_\varepsilon - (1 + \chi) \sigma \frac{1}{\hat{\sigma}^2} \frac{v_\varepsilon}{2} \]
Exact expression for \( \text{SWF}(\tau) \)

\[
\mathcal{W}(\tau) = \chi \log \chi - (1 + \chi) \log(1 + \chi) + (1 + \chi) \left\{ \frac{\log(1 - \tau)}{(1 + \hat{\sigma})(1 - \tau)} - \frac{1}{(1 + \hat{\sigma})} \right\} \\
+ (1 + \chi) \left( \frac{\psi}{1 + \psi} \right) \frac{1}{\theta - 1} \log(1 - \tau) \\
- \left( \frac{\psi}{1 + \psi} \right) \frac{1}{\theta} (1 - \tau) - \left[ -\log \left( 1 - \left( \frac{1 - \tau}{\theta} \right) \right) - \left( \frac{1 - \tau}{\theta} \right) \right] \\
- (1 - \tau)^2 \frac{\nu_\varphi}{2} \\
- \left[ (1 - \tau) \frac{\delta}{1 - \delta} \frac{\nu_\omega}{2} - \log \left( \frac{1 - \delta \exp \left( \frac{-\tau(1 - \tau)\nu_\omega}{2} \right)}{1 - \delta} \right) \right] \\
+ (1 + \chi) \frac{1}{\hat{\sigma}} \nu_\varepsilon - (1 + \chi) \sigma \frac{1}{\hat{\sigma}^2} \frac{\nu_\varepsilon}{2}
\]
Skill investment component

\[ W(\tau) = \chi \log \chi - (1 + \chi) \log(1 + \chi) + (1 + \chi) \frac{\log(1 - \tau)}{(1 + \hat{\sigma})(1 - \tau)} - \frac{1}{(1 + \hat{\sigma})} + (1 + \chi) \left( \frac{\psi}{1 + \psi} \right) \frac{1}{\theta - 1} \log (1 - \tau) \]

- \left( \frac{\psi}{1 + \psi} \right) \frac{1}{\theta} (1 - \tau) - \left[ -\log \left( 1 - \left( \frac{1 - \tau}{\theta} \right) \right) - \left( \frac{1 - \tau}{\theta} \right) \right] \]

productivity gain = \log E [(p(s))] = \log (Y/N)

avg. education cost

consumption dispersion across skills

- \left( 1 - \tau \right)^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) \frac{1}{\hat{\sigma}} v \varepsilon - (1 + \chi) \sigma \frac{1}{\hat{\sigma}^2} \frac{v \varepsilon}{2}
Skill investment component

Heathcote-Storesletten-Violante, "Optimal Tax Progressivity"
Skill investment component

- Diamond-Saez formula for top marginal rate: \( \bar{t} = \frac{1+\sigma}{\theta+\sigma} \)
  - Lower \( \theta \): thicker Pareto tail in \( y \) dist. \( \rightarrow \) more redistribution

- Our model: endogenous skill accumulation
  - Lower \( \theta \): strong skill complementarity \( \rightarrow \) more skill investment
Uninsurable component

\[ \mathcal{W}(\tau) = \chi \log \chi - (1 + \chi) \log(1 + \chi) + (1 + \chi) \frac{\log(1 - \tau)}{(1 + \hat{\sigma})(1 - \tau)} - \frac{1}{(1 + \hat{\sigma})} \]

\[ + (1 + \chi) \left( \frac{\psi}{1 + \psi} \right) \frac{1}{\theta - 1} \log (1 - \tau) \]

\[ - \left( \frac{\psi}{1 + \psi} \right) \frac{1}{\theta} (1 - \tau) - \left[ - \log \left( 1 - \left( \frac{1 - \tau}{\theta} \right) \right) \right] - \left( \frac{1 - \tau}{\theta} \right) \]

\[ - \left( (1 - \tau)^2 \frac{\nu_\varphi}{2} \right) \]

cons. disp. due to prefs.

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

consumption dispersion due to uninsurable shocks \( \approx (1 - \tau)^2 \frac{\nu_\varphi}{2} \)

\[ + (1 + \chi) \frac{1}{\hat{\sigma}} \nu_\varepsilon - (1 + \chi) \sigma \frac{1}{\hat{\sigma}^2} \frac{\nu_\varepsilon}{2} \]
Insurable component

\[ W(\tau) = \chi \log \chi - (1 + \chi) \log(1 + \chi) + (1 + \chi) \frac{\log(1 - \tau)}{(1 + \hat{\sigma})(1 - \tau)} - \frac{1}{(1 + \hat{\sigma})} \]

\[ + (1 + \chi) \left( \frac{\psi}{1 + \psi} \right) \frac{1}{\theta - 1} \log (1 - \tau) \]

\[ - \left( \frac{\psi}{1 + \psi} \right) \frac{1}{\theta} (1 - \tau) - \left[ - \log \left( 1 - \left( \frac{1 - \tau}{\theta} \right) \right) - \left( \frac{1 - \tau}{\theta} \right) \right] \]

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

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

\[ + (1 + \chi) \frac{1}{\hat{\sigma}} v_\varepsilon \]

prod. gain from ins. shock = \( \log(N/H) \)

\[ - (1 + \chi) \sigma \frac{1}{\hat{\sigma}^2} \]

hours dispersion
QUANTITATIVE IMPLICATIONS
Parameterization

- Parameter vector \( \{ \chi, \sigma, \psi, \theta, v_\varphi, v_\omega, v_\varepsilon \} \)
Parameterization

- Parameter vector \( \{\chi, \sigma, \psi, \theta, v_\varphi, v_\omega, v_\varepsilon\} \)

- Assume observed \( G/Y = 0.19 = g^* \) \( \rightarrow \chi = 0.233 \)

- Frisch elasticity (micro-evidence \( \sim 0.5 \)) \( \rightarrow \sigma = 2 \)
Parameterization

• Parameter vector \( \{\chi, \sigma, \psi, \theta, v_\phi, v_\omega, v_\varepsilon\} \)

• Assume observed \( G/Y = 0.19 = g^* \) \( \rightarrow \chi = 0.233 \)

• Frisch elasticity (micro-evidence \( \sim 0.5 \)) \( \rightarrow \sigma = 2 \)

• Price-elasticity of skill investment \( \rightarrow \psi = 0.65 \)
Parameterization

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

- Assume observed $G/Y = 0.19 = g^*$ \quad $\rightarrow \chi = 0.233$

- Frisch elasticity (micro-evidence $\sim 0.5$) \quad $\rightarrow \sigma = 2$

- Price-elasticity of skill investment \quad $\rightarrow \psi = 0.65$

\[
\begin{align*}
\text{cov}(\log h, \log w) &= \frac{1}{\sigma} v_{\varepsilon} \\
\text{var}(\log h) &= v_{\varphi} + \frac{1}{\sigma^2} v_{\varepsilon} \\
\text{var}^0(\log c) &= (1 - \tau)^2 \left( v_{\varphi} + \frac{1}{\theta^2} \right) \\
\text{var}(\log w) &= \frac{1}{\theta^2} + \frac{\delta}{1 - \delta} v_{\omega} + v_{\varepsilon}
\end{align*}
\]
Parameterization

- Parameter vector \( \{\chi, \sigma, \psi, \theta, v_\varphi, v_\omega, v_\varepsilon\} \)
- Assume observed \( G/Y = 0.19 = g^* \) \( \rightarrow \chi = 0.233 \)
- Frisch elasticity (micro-evidence \( \sim 0.5 \)) \( \rightarrow \sigma = 2 \)
- Price-elasticity of skill investment \( \rightarrow \psi = 0.65 \)

\[
\text{cov}(\log h, \log w) = \frac{1}{\hat{\sigma}} v_\varepsilon \quad \rightarrow v_\varepsilon = 0.17
\]

\[
\text{var}(\log h) = v_\varphi + \frac{1}{\hat{\sigma}^2} v_\varepsilon \quad \rightarrow v_\varphi = 0.035
\]

\[
\text{var}^0(\log c) = (1 - \tau)^2 \left( v_\varphi + \frac{1}{\theta^2} \right) \quad \rightarrow \theta = 3.12
\]

\[
\text{var}(\log w) = \frac{1}{\theta^2} + \frac{\delta}{1 - \delta} v_\omega + v_\varepsilon \quad \rightarrow v_\omega = 0.003
\]
Optimal progressivity

Social Welfare Function

Welfare Gain = 0.4%

$\tau^* = 0.084$

$\tau^{US} = 0.161$

Heathcote-Storelletten-Violante, "Optimal Tax Progressivity"
Optimal progressivity: decomposition

Social Welfare Function

(1) Rep. Agent
\[ \tau = -0.233 \]

Heathcote-Storeisletten-Violante, "Optimal Tax Progressivity"
Optimal progressivity: decomposition

Social Welfare Function

(1) Rep. Agent
\( \tau = -0.233 \)

(2) + Skill Inv.
\( \tau = -0.035 \)

Heathcote-Storesletten-Violante, "Optimal Tax Progressivity"
Optimal progressivity: decomposition

Social Welfare Function

(1) Rep. Agent
\( \tau = -0.233 \)

(2) + Skill Inv.
\( \tau = -0.035 \)

(3) + Pref. Het.
\( \tau = -0.007 \)

Heathcote-Storesletten-Violante, "Optimal Tax Progressivity"
Optimal progressivity: decomposition

Heathcote-Storelsetten-Violante, "Optimal Tax Progressivity"
Optimal progressivity: decomposition

Social Welfare Function

- (1) Rep. Agent: $\tau = -0.233$
- (2) + Skill Inv.: $\tau = -0.035$
- (3) + Pref. Het.: $\tau = -0.007$
- (4) + Uninsurable Shocks: $\tau = 0.099$
- (5) + Insurable Shocks: $\tau^* = 0.084$

Heathcote-Storesletten-Violante, "Optimal Tax Progressivity"
Actual and optimal progressivity

Income-weighted average marginal: down from 32% to 26%

Heathcote-Storesletten-Violante, "Optimal Tax Progressivity"
If you believe that...

- $G$ does not yield any utility ($\chi = 0$):
  
  $\tau^* = 0.20 \implies y$-weighted average MTR: 36 pct
If you believe that...

- $G$ does not yield any utility $(\chi = 0)$:
  \[ \tau^* = 0.20 \rightarrow \text{y-weighted average MTR: 36 pct} \]

- All uninsurable wage ineq. due to exogenous shocks $(\theta = \infty)$:
  \[ \tau^* = 0.21 \rightarrow \text{y-weighted average MTR: 37 pct} \]
If you believe that...

- $G$ does not yield any utility ($\chi = 0$):
  \[ \tau^* = 0.20 \rightarrow \text{y-weighted average MTR: 36 pct} \]

- All uninsurable wage ineq. due to exogenous shocks ($\theta = \infty$)
  \[ \tau^* = 0.21 \rightarrow \text{y-weighted average MTR: 37 pct} \]

- All uninsurable wage ineq. is due to endogenous choices ($v_{\omega} = 0$)
  \[ \tau^* = 0.06 \rightarrow \text{y-weighted average MTR: 24 pct} \]
EXTENSIONS
Role of weight on future vs. current cohorts
Role of weight on future vs. current cohorts

Lower weight $\rightarrow$ more concern for current inequality and redistribution

Heathcote-Storesletten-Violante, "Optimal Tax Progressivity"
Irreversible skill investment
Irreversible skill investment

- Progressivity does not distort \textit{sunk} skill inv. of existing cohorts

Heathcote-Storesletten-Violante, "Optimal Tax Progressivity"
• Progressivity does not distort *sunk* skill inv. of existing cohorts

• As weight → 1, (ir)-reversibility does not matter
Age-dependent progressivity

- Give the planner ability to index the pair $(\lambda, \tau)$ on individual age $a$

- Link with dynamic Mirrlees approach: age-dependent tax scheme realizes most of gains from fully history-dependent tax reform
Age-dependent progressivity

• Give the planner ability to index the pair \((\lambda, \tau)\) on individual age \(a\)

• Link with dynamic Mirrlees approach: age-dependent tax scheme realizes most of gains from fully history-dependent tax reform

• Three results:
  
  ▶ Optimal public good provision \(g^*\) is unchanged

  ▶ The sequence \(\{\lambda_a^*, \tau_a^*\}\) is independent of age iff \(v_\omega = 0\)

  ▶ With \(v_\omega > 0\), the sequence \(\{\lambda_a^*, \tau_a^*\}\) is strictly increasing in \(a\)
Age-dependent progressivity

Welfare gains from making $\tau^*$ age dependent near 5%!

Heathcote-Storesletten-Violante, "Optimal Tax Progressivity"
Three lessons on optimal progressivity
Three lessons on optimal progressivity

1. The **endogeneity of the skill distribution** limits optimal progressivity

   • **Key:** skill-complementarity in production \((\theta)\), price-elasticity of skill investment \((\psi)\), alterability of past skill choices
Three lessons on optimal progressivity

1. The endogeneity of the skill distribution limits optimal progressivity

   • Key: skill-complementarity in production ($\theta$), price-elasticity of skill investment ($\psi$), alterability of past skill choices

2. The externality in the provision of public goods limits progressivity

   • Low progressivity induces higher labor supply, output, and $G$
Three lessons on optimal progressivity

1. The endogeneity of the skill distribution limits optimal progressivity

   • Key: skill-complementarity in production ($\theta$), price-elasticity of skill investment ($\psi$), alterability of past skill choices

2. The externality in the provision of public goods limits progressivity

   • Low progressivity induces higher labor supply, output, and $G$

3. Age-dependent progressivity delivers large welfare gains

   • Low progressivity at young ages induces skill investment
   • High progressivity at old ages redistributes against shocks
Alternative drastic solution to increase welfare...

THANKS!
Inequality aversion

- **Utilitarian planner**: equal concern for redistributing across individuals and for insuring consumption fluctuations over time

- New inequality aversion parameter $\nu \in (0, \infty)$ to vary the strength of the concern for redistribution
Inequality aversion

- **Utilitarian planner**: equal concern for redistributing across individuals and for insuring consumption fluctuations over time.

- **New inequality aversion parameter** $\nu \in (0, \infty)$ to vary the strength of the concern for redistribution.

<table>
<thead>
<tr>
<th>$\nu$</th>
<th>Planner</th>
<th>$\tau^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\to 0$</td>
<td>Rawlsian</td>
<td>$1.0$</td>
</tr>
<tr>
<td>$1$</td>
<td>Utilitarian</td>
<td>$0.084$</td>
</tr>
<tr>
<td>$\to \infty$</td>
<td>Inequality-neutral</td>
<td>$-0.159$</td>
</tr>
</tbody>
</table>
Inequality aversion

- **Utilitarian planner**: equal concern for redistributing across individuals and for insuring consumption fluctuations over time

- New inequality aversion parameter $\nu \in (0, \infty)$ to vary the strength of the concern for redistribution

<table>
<thead>
<tr>
<th>$\nu$</th>
<th>Planner</th>
<th>$\tau^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\to 0$</td>
<td>Rawlsian</td>
<td>1.0</td>
</tr>
<tr>
<td>1</td>
<td>Utilitarian</td>
<td>0.084</td>
</tr>
<tr>
<td>$\to \infty$</td>
<td>Inequality-neutral</td>
<td>−0.159</td>
</tr>
</tbody>
</table>

- Planner only concerned with consumption insurance ($\nu \to \infty$) chooses an income-weighted average marginal tax rate of 6%

Heathcote-Storesletten-Violante, "Optimal Tax Progressivity"