OPTIMAL TAX PROGRESSIVITY

An analytical framework

Heathcote, J Storesletten, G Violante, J

Alexandre Gaillard
Toulouse School of Economics

May 3, 2017

Introduction

Key idea: use realistic parametric tax function to get analytical optimal tax formula

A rich environment with:

1. **Public good provision with spillover effects.**
   Agents fail to internalize their contribution to welfare. Need to tax them to correct the gap.
2. **Endogenous skill investment.**
   Higher tax at the top will reduce incentive to get high skill $\rightarrow$ lower productivity.
3. **Partial insurance against labor income risk.**
   More insurable earnings variations $\rightarrow$ lower needs for redistribution $\rightarrow$ lower tax progressivity.

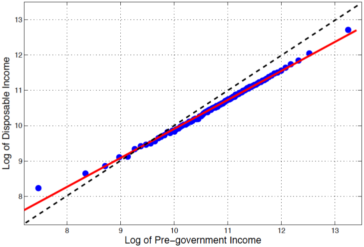
At the center: parametric tax function

Benabou (2000, 2002) and many papers from the same authors ("HSV" function).

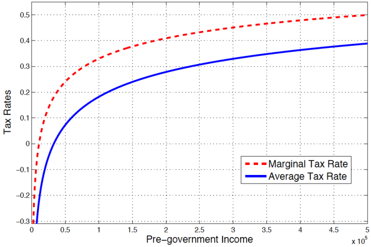
$$T(y) = y - \lambda y^{1-\tau} \qquad y \equiv \text{income} \tag{1}$$

- τ control the progressivity of the tax function (ex: < 0 regressive).
- λ is a scale parameter, used to balance gov' budget.
- $y^0 = \lambda^{\frac{1}{\tau}} > 0$ defines the threshold at which taxes are positive.

HSV function and empirical fit



(a) Statistical fit on U.S. data



(b) Implied average and marginal tax rates

- Using PSID: estimate $\tau_{US} = 0.181$ using OLS (so progressive).
- Implied income-weighted average marginal tax equal to 0.34.
- Underestimate marginal tax rates at low income levels.

Motivation of the paper

Given the tax system currently in place:

1. By how much τ should be moved to get optimality ?
2. What is the associated welfare gain ?

Key features: everything is parametric. Two limits with $T(y)$

1. $T(y_i)$ is either globally convex ($\tau > 0$) or concave ($\tau < 0$).
2. Does not allow for lump-sum cash transfers ($T(0) = 0$).

Motivation of the paper

Given the tax system currently in place:

1. By how much τ should be moved to get optimality ?
2. What is the associated welfare gain ?

Key features: everything is parametric. Two limits with $T(y)$

1. $T(y_i)$ is either globally convex ($\tau > 0$) or concave ($\tau < 0$).
2. Does not allow for lump-sum cash transfers ($T(0) = 0$).

Why not use more complex tax function ?

(i). tractability

(ii). welfare gain (Heathcote and Tsujiyama (2016): move to Mirelees taxation implies $+0.20\%$ gains relative to HSV).

Economic Environment

4 dimensions of heterogeneity for an individual $i \in [0, 1]$ with age a

1. Labor productivity: permanent (α_{ai}) + transitory (ϵ_{ai}) components.
2. Disutility of work effort (ϕ_i).
3. Innate learning ability (κ_i).

Choice variables: s_i: skill , c_i: consumption, h_i: hours worked.

Economic Environment

4 dimensions of heterogeneity for an individual $i \in [0,1]$ with age a

1. Labor productivity: permanent (α_{ai}) + transitory (ϵ_{ai}) components.
2. Disutility of work effort (ϕ_i).
3. Innate learning ability (κ_i).

Choice variables: s_i: skill , c_i: consumption, h_i: hours worked.

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

- θ: elasticity of substitution across skills.
- N(s): average effective hours worked by individuals with skill s.
- m(s): density of individuals with skill s.

Output is shared between $G(= gY)$ and C: $Y = \int_0^\infty c_i di + G$

Preferences

$$U_i = \frac{(\kappa_i)^{\frac{-1}{\psi}}}{1+\frac{1}{\psi}}(s_i)^{1+\frac{1}{\psi}} + (1-\beta\delta)\mathbb{E}_0\sum_{a=0}^{\infty}(\beta\delta)^a u_i(c_{ia}, h_{ia}, G) \quad (2)$$

$$u_i(c_{ia}, h_{ia}, G) = log c_{ia} - \frac{exp[(1+\sigma)\phi_i]}{1+\sigma}(h_{ia})^{1+\sigma} + \chi log G \quad (3)$$

$$\phi_i \sim \mathcal{N}(\nu_\phi/2, \nu_\phi), \quad \kappa_i \sim exp(\eta) \quad (4)$$

- χ is the taste for public good relative to private consumption.
- s_i is chosen before the individual enters the model, given κ_i

Labor productivity and earnings

From labor economics (Meghir and Pistaferri, 2011)

$$log(z_{ia}) = \alpha_{ia} + \epsilon_{ia} \tag{5}$$

$$\alpha_{ia} \sim \mathcal{N}(-\nu_\omega/2, \nu_\omega) \qquad \epsilon_{ia} \sim \mathcal{N}(-\nu_\epsilon/2, \nu_\epsilon) \tag{6}$$

Labor earnings is therefore: $y_{ia} = \underbrace{p(s_i)}_{\text{skill price}} \underbrace{exp(\alpha_{ia} + \epsilon_{ia})}_{\text{shocks}} \underbrace{h_{ia}}_{\text{hours worked}}$

- $p(s) = (Y/[N(s).m(s)])^{1/\theta}$.
- difference across individual with same s_i and h_i is given by shocks (luck).

Financial assets

1. There is a full set of state-contingent claims indexed by the ϵ shock $\rightarrow$ this shock is fully insurable.
   - Hours worked perfectly react to ϵ shocks (without taxes)
2. α shocks can not be insured
   - No saving here to smooth.

(i) Autarky if $\nu_\epsilon = 0$,
(ii) full insurance if $\nu_\omega = 0$
(iii) partial insurance if both are positive.

Intuition: if large part of the earnings is uninsurable, then government redistribution motive will be higher, τ rises.

Agent's problem: solution to investment in skills

At age $a = 0$, i chooses s : $\left(\frac{s}{\kappa_i}\right)^{\frac{1}{\phi}} = (1 - \beta\delta)\mathbb{E}_0 \sum_{a=0}^{\infty}(\beta\delta)^a \frac{\partial u_i(.)}{\partial s}$

In subsequent period:

1. α_{ia} is realized
2. Individuals buys insurance claims $B(.)$.
3. ϵ_{ia} is realized and individuals choose h_{ia}, c_{ia}.

$$\int_E Q(\epsilon)B(\epsilon)d\epsilon = 0 \tag{7}$$

$$c_{ia} = \lambda[p(s_i)exp(\alpha_{ia} + \epsilon_{ia})h_{ia}]^{1-\tau} + B(\epsilon_{ia}) \tag{8}$$

Representative agent problem

Set $\nu_\phi = \nu_\omega = \nu_\epsilon = 0$ and $\theta = \infty$ implies no dispersion in taste for leisure, labor productivity. Skill are perfect substitute.

$$\max_{C,H} \left\{ logC - \frac{H^{1+\sigma}}{1+\sigma} + \chi logG \right\} \quad s.t \quad C = \lambda H^{1-\tau} \quad G = gY = gH$$

$$logH^{RA}(\tau) = \frac{1}{1+\sigma} log(1-\tau) \tag{9}$$

$$logC^{RA}(g,\tau) = log(1-g) + \frac{1}{1+\sigma} log(1-\tau) \tag{10}$$

- Tax reduces hours worked and thus consumption = behavioral effect. As $\tau \to 0$ then $H^{RA}(\tau) \to 0$.

Heterogenous agent model - consumption / hours worked

$$logh(\phi,\epsilon;\tau) = \underbrace{logH^{RA}(\tau)^{-}}_{\text{behavioral effect}} - \underbrace{\phi}_{\text{disutility}} + \underbrace{\frac{1}{\hat{\sigma}}\epsilon}_{\text{shock}} + \underbrace{\frac{1}{\hat{\sigma}}C(\nu_\epsilon;\tau)^{-}}_{\text{dispersion effect}}$$

$$logc(\phi,\alpha,s;g,\tau) = \underbrace{log[C^{RA}(g,\tau)^{-}\Theta(\tau)]}_{\text{behavioral effect}}$$

$$+ (1-\tau)\underbrace{[logp(s;\tau)+\alpha-\phi]}_{\text{labor income - disutility}} + \underbrace{(1-\tau)C(\nu_\epsilon;\tau)^{-}}_{\text{dispersion effect}}$$

Dispersion effect: more insurable wage variation → need to work more when high productivity and less with low productivity.

→ This rises productivity, average leisure, and welfare (since productivity is in line with earnings)

→ Progressivity: weakens this channel.

Heterogenous agent model: skill investment

$$log(p(s;\tau)) = \pi_0(\tau)^{-} + \pi_1(\tau)^{+} s(\kappa;\tau)^{-} \tag{11}$$

$$s(\kappa;\tau) = \left[\frac{\eta}{\theta}(1-\tau)\right]^{\frac{\phi}{1+\phi}} \kappa \tag{12}$$

- π_0 is the skill price of the lowest skill type.
- Higher τ reduces the incentive to acquire skills.
- Stiglitz effect: compresses the skill distribution toward zero, but imperfect substitutability across skills increases $p(s;\tau)$.
- $var(log(p(s;\tau)) = \frac{1}{\theta^2}$ independent on τ. The two effects cancel.

Optimal degree of progressivity

Equilibrium with $\tau = 0$ is (in general) inefficient.

- No private market for insuring α shocks. $(\tau > 0)$
- Free-riding problem. If all agents work more, the quantity of public goods increases. $(\tau < 0)$

Welfare - when skills are fully reversible

Goal: compute the welfare gain from (g_{-1}, τ_{-1}) to (g, τ).

Skill reversibility implies that transition is immediate. No need to keep track individuals.

Utilitarian social welfare is used as baseline.

- puts equal weight on all agents within a cohort (but weight less old people).
- implies that gov' would want to redistribute income across generations (matter if skills are not reversible).

$$W(g, \tau; \tau_{-1}) \equiv (1 - \gamma)\Gamma \sum_{j=-\infty}^{\infty} \gamma^j U_{j,0}(g, \tau; \tau_{-1}) \tag{13}$$

Closed-form social welfare

Method to solve:

1. Solve for the value $\lambda(g, \tau)$ that balances the gov' budget.
2. Next, plug consumption, hours, and skill allocations in (14)

Using the RA model simplification, they get:

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

This implies that:

- Samuelson condition: $g^* = \frac{\chi}{1+\chi}$ $\left(MRS_{priv./pub.} = MRT_{priv./pub.}\right)$
- $\tau^* = -\chi < 0$ due to the free-riding effect.
- independent of τ_{-1}

Closed form social welfare - Heterogeneous agents

$$W(g,\tau;\tau_{-1}) = W^{RA}(g,\tau) + f(\tau) + h(\tau_{-1})^{-} + q(\tau)$$

Corollary 4.2 $W(g,\tau;\tau_{-1})$ is globally concave in g and if $\sigma \geq 2$, then globally concave in τ.

$\rightarrow$ satisfied in reasonable calibration $\rightarrow$ FOC are sufficient !

- τ^* is independent of g^*, $g^* = g^*_{RA}$ only determined by Samuelson condition.
- τ^* is independent of τ_{-1} , due to full reversibility.
- $f(\tau)$ contains the effect of τ on uninsurable differences across individuals.
- $q(\tau)$ contains the effect of τ on insurable risk.

Closed form social welfare - Heterogeneous agents

$$f(\tau) = \underbrace{\underbrace{(1-\chi)\Theta_1(\tau)^{-}}_{\text{productivity effect}} + \underbrace{\Theta_2(\tau)^{+}}_{\text{Cost}} + \underbrace{\Theta_3(\tau)^{+}}_{\text{Price effect}}}_{\text{Effet on skill investment}} + \underbrace{\underbrace{\Theta_4(\tau)^{+}}_{\text{disutility}} + \underbrace{\Theta_5(\tau)^{+}}_{\text{uninsurable shocks}}}_{\text{Effect on consumption dispersion}}$$

Effect of increasing τ:

- Lowers productivity because wage dispersion is lower $\rightarrow$ reduces "superstars" size.
- High skill types become scarce $\rightarrow$ increases p
- Decreases consumption inequalities through differences in disutility of effort and uninsurable shocks

Question: which effect dominates for skills?

Effect on skill investment: all is about θ

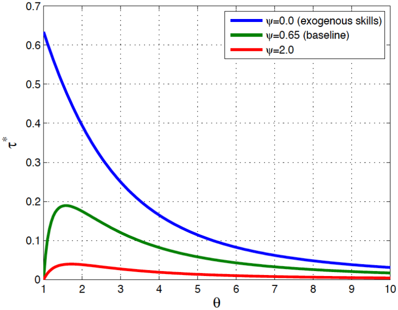


FIGURE II: Optimal τ as a Function of θ in the Special Case $v_\varphi = v_\omega = v_\varepsilon = \chi = 0$ and $\sigma = 2$

- If $\psi = 0$, then τ is decreasing in θ. Higher substitutability across skills means less inequalities in wages.
- if $\psi > 0$, then increasing τ for low θ lowers W since productivity decreases.

Insurable risk effect

The effect of $q(\tau)$ in $W(.)$ is due to insurable wage variation.

$$q(\tau) = log\left(\frac{N(\tau)}{H(\tau)}\right) - var_\epsilon(logh)$$
$$= \frac{1}{\hat{\sigma}}\nu_\epsilon - \frac{1}{\hat{\sigma}^2}\nu_\epsilon$$

Increasing insurable wage variation:

- increases productivity because induces individuals to work more when endowed with good shock, and work less with bad shock.
- As hours is costly (effort), this reduces welfare.

The sum of the two effects is max when $\tau = 0$ (hours worked respond efficiently to insurable shocks), so push it toward 0.

When taxes should be progressive ?

By differentiating W, τ is strictly positive if and only if

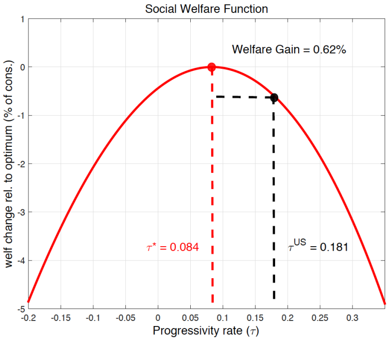
$$\underbrace{\frac{\psi}{(1+\psi)\theta}}_{\text{lower investment cost}} + \underbrace{\frac{1}{(\theta-1)\theta}}_{\text{less skill price inequality}} + \underbrace{(\nu_\phi + \nu_\alpha)}_{\text{less } \alpha \text{ and } \phi \text{ inequality}} > \underbrace{\frac{\psi}{(1+\psi)(\theta-1)}}_{\text{lower productivity due to less inv.}} + \underbrace{\chi\left(\frac{1}{1+\sigma} + \frac{1}{(1+\psi)(\theta-1)}\right)}_{\text{lower } G \text{ due to less hours and less inv.}}$$

LHS: marginal benefit of increasing τ at $\tau = 0$, RHS: marginal cost.

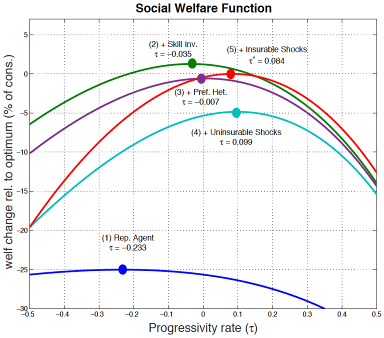
- If $\chi = 0$ then $\tau = 0$. Increasing χ decrease τ.
  That is, the behavioral response limits the temptation to compress inequality by increasing progressivity.
- Effect of θ is ambiguous, as discussed before.
- More dispersion in α and ϕ makes τ more progressive.

Quantitative analysis

Optimal progressivity: $\tau^* = 0.084$ against $\tau_{US} = 0.181$



(a) Social welfare function



(b) Decomposition of the social welfare function

Robustness check

1. Skill irreversibility: government can expropriate past generations. Lower τ to 0.061.
2. Public consumption: the way the public consumption is modeled has large impacts on τ^*. Support a single model with tax and public good provision.
3. Inequality aversion: a model with gov' inequality aversion (ex: Rawlsian type) increases τ.
4. Voting model

Conclusion

To conclude, does the parameters $(\chi, \theta, \nu_\alpha, \nu_\phi, \nu_\epsilon)$ of interest matter in reality ?

Empirical estimates (across countries) suggest that:

- τ is decreasing with g (or χ).
- τ displays the predicted hump-shaped curve w.r.t θ
- τ increases with inequality.