Notes on Heathcote, Storesletten, and Violante (2017)

Alex von Hafften

December 13, 2022

My notes follow the organization of the paper. Initial focus particularly on sections 3, 4, and 5. Appendix has notes from Job's lecture on related papers, Heathcote, Storesletten, and Violante (2014) and Boerma and Karabarbounis (2021).

1 Introduction

- What is the optimal degree of progressivity of tax/transfer system?
- Reasons for more progressivity
  - Counteracts inequality in initial conditions
  - Substitutes for imperfect private insurance against idiosyncratic earnings risk
- Reasons for less progressivity
  - Reduces incentives to work
  - Reduces incentives to invest in skills
- Develop general equilibrium model with these four trade-offs
- Results for U.S. calibration:
  - Baseline model indicates current system is too progressive
  - Adding that poverty constrains skill investment indicates the current system is close to optimal

2 Tax Function

- Tax revenue for pre-tax income y is

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

- After-tax income $\tilde{y}_i$ for pre-tax income y_i

$$\begin{aligned}\tilde{y}_i &= y_i - T(y_i) \\ &= y_i - y_i + \lambda y_i^{1-\tau} \\ &= \lambda y_i^{1-\tau} \end{aligned} \tag{2}$$

Thus, $1-\tau$ is elasticity of after-tax income to pre-tax income.

- If $\tau > 0$ then marginal rates exceed average rates (progressive), if $\tau < 0$, marginal rates are smaller than average rates (regressive), and, if $\tau = 0$, marginal rates and average are the same $\implies$ flat tax.

$$\begin{aligned}\frac{1-T'(y_i)}{1-\frac{T(y_i)}{y_i}} &= \frac{1-(1-\lambda(1-\tau)y^{-\tau})}{1-(1-\lambda y^{-\tau})} \\ &= 1-\tau \end{aligned} \tag{3}$$

- With balanced budget $gY = \int T(y_i)di$, so income-weighted marginal rate is

$$\begin{aligned}\int T'(y_i)\left(\frac{y_i}{Y}\right)di &= \int (1-(1-\tau)\lambda y_i^{-\tau})\left(\frac{y_i}{Y}\right)di \\ &= \frac{1}{Y}\int y_i di - (1-\tau)\frac{1}{Y}\int \lambda y_i^{1-\tau}di \\ &= 1-(1-\tau)\frac{1}{Y}\int y_i - T(y_i)di \\ &= 1-(1-\tau)\frac{1}{Y}(Y-gY) \\ &= 1-(1-\tau)(1-g) \end{aligned} \tag{4}$$

2.1 Empirical Fit

2.2 Robustness

2.3 Discussion

3 Economic Environment

- Economy is steady state, so time subscripts omitted.

3.1 Demographics

- Agents indexed by $i = [0,1]$.
- Yaari perpetual youth structure:
  - At every age a, an agent survives to next period with probability $\delta < 1$.
  - Each period newborn agents of size $1-\delta$ enter economy.

3.2 Life Cycle

- At beginning of live, agent i chooses initial investment in skills s_i.
- Agent i then enters labor market and faces random fluctuations in labor productivity z_i.
- Every period, agent i chooses hours of work $h_i \geq 0$ and consumption of private good c_i.

3.3 Technology

- $\theta > 1$ denotes elasticity of substitution across skill types
- $N(s)$ denotes average effective hours worked by individuals with skill type s
- $m(s)$ denotes the density of individuals with skill type s
- Output Y is constant elasticity of substitution aggregate of effective hours supplied by continuum of skill types $s \in [0, \infty)$:

$$Y = \left(\int_0^\infty [N(s) \cdot m(s)]^{\frac{\theta-1}{\theta}} ds \right)^{\frac{\theta}{\theta-1}} \tag{5}$$

- Marginal product from an additional unit of effective hours of skill type s is

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

- Resource constraint is

$$Y = \int_0^1 c_i di + G \tag{6}$$

3.4 Preferences

- β is the "pure" discount factor
- Agent i has preferences over private consumption, hours worked, public goods, and skill investment effort:

$$U_i = -v_i(s_i) + (1 - \beta\delta) E_0 \sum_{a=0}^{\infty} (\beta\delta)^a u_i(c_{ia}, h_{ia}, G) \tag{7}$$

  where expectations are taken over the future idiosyncratic productivity shocks.

- $\psi \geq 0$ denotes the elasticity of skill investment wrt to the return to skill

- $\kappa_i \geq 0$ is an individual-specific parameter that determine the utility cost of acquiring skills (larger $\kappa_i \implies$ the skills are cheaper, so κ_i can be thought of as learning ability); κ_i is assumed to be distributed exponential with parameter η: $\kappa_i \sim Exp(\eta)$
- The disutility of the initial skill investment $s_i \geq 0$ is

$$v_i(s_i) = \frac{\psi}{1+\psi}\kappa_i^{-\frac{1}{\psi}} s_i^{\frac{1+\psi}{\psi}} \tag{8}$$

- Let $\exp[(1+\sigma)\varphi_i]$ be the disutility of work effort with $\varphi_i \sim N(\frac{v_\varphi}{2}, v_\varphi)$ and κ_i and φ_i are uncorrelated.
- Let $\sigma > 0$ determine aversion to hours fluctuations
- Let $\chi \geq 0$ measure the taste for public goods relative to private consumption goods
- The period utility function

$$u_i(c_{ia}, h_{ia}, G) = \log c_{ia} - \frac{\exp[(1+\sigma)\varphi_i]}{1+\sigma}(h_{ia})^{1+\sigma} + \chi \log G \tag{9}$$

- Define the tax-modified Frisch elasticity $\hat{\sigma}$ as:

$$\frac{1}{\hat{\sigma}} \equiv \frac{1-\tau}{\sigma+\tau} \tag{10}$$

3.5 Labor Productivity and Earnings

- Log labor efficiency z_{ia} is sum of a random walk and orthogonal white noise

$$\begin{aligned} \log z_{ia} &= \alpha_{ia} + \varepsilon_{ia} \\ \text{where } \alpha_{ia} &= \alpha_{i,a-1} + \omega_{ia} \\ \omega_{ia} &\sim^{iid} N(-\frac{v_w}{2}, v_w) \\ \alpha_{i0} &= 0 \\ \varepsilon_{ia} &\sim^{iid} N(-\frac{v_\varepsilon}{2}, v_\varepsilon) \end{aligned} \tag{11}$$

- Individual earnings y_{ia} is the product of the equilibrium price for labor from skill type s_i, individual labor productivity, and the number of hours worked.

$$y_{ia} = p(s_i) \times \exp(\alpha_{ia} + \varepsilon_{ia}) \times h_{ia} \tag{12}$$

3.6 Financial Assets

- Partial insurance structure with ε shocks fully insurable but α shocks are not
- Full set of state-contingent claims indexed by the ε shocks.
- Let $B(\mathbf{E})$ and $Q(\mathbf{E})$ denote the quantity and the price of insurance claims purchased that pay one unit of consumption iff $\varepsilon \in \mathbf{E} \subset \mathbb{R}$.
- Insurance claims in zero net supply and newborn agents start with zero initial holdings
- Notice special cases of autarky with $v_\varepsilon = 0$ and full insurance with $v_\omega = 0$

3.7 Markets

- Competitive market for final consumption good, all types of labor services, and financial claims.
- Final consumption good is numeraire

3.8 Government

- Government chooses (g, τ) where g is government consumption as fraction of aggregate output and τ determines the degree of progressivity of the tax system.
- Given (g, τ), the average level of taxation λ balances its budget:

$$G \equiv gY = g \int_0^1 y_i di = \int_0^1 (y_i - \lambda y_i^{1-\tau}) di \tag{13}$$

3.9 Agent's Problem

- At $a = 0$, given (κ_i, φ_i), agent chooses a skill level given. FOC of (7) wrt s_i

$$\frac{\partial v_i(s_i)}{\partial s_i} = (1 - \beta\delta) E_0 \sum_{a=0}^{\infty} (\beta\delta)^a \frac{\partial u_i(c_{ia}, h_{ia}, G)}{\partial s_i}$$

In words, the marginal disutility of skill investment must equal the present discounted value of the higher associated expected lifetime wages. The derivative of (8) is

$$\frac{\partial v_i(s_i)}{\partial s_i} = \frac{\psi}{1+\psi} \frac{1+\psi}{\psi} \kappa_i^{-\frac{1}{\psi}} s_i^{\frac{1}{\psi}} = \left(\frac{s_i}{\kappa_i}\right)^{\frac{1}{\psi}}$$

Combining

$$\left(\frac{s_i}{\kappa_i}\right)^{\frac{1}{\psi}} = (1 - \beta\delta) E_0 \sum_{a=0}^{\infty} (\beta\delta)^a \frac{\partial u_i(c_{ia}, h_{ia}, G)}{\partial s_i} \tag{14}$$

- At $a > 0$, timing is:
    1. ω_{ia} realized
    2. Insurance market for ε shocks opens and individual buys claims $B(\cdot)$
    3. ε_{ia} is realized
    4. Chooses hours h_{ia}, receives y_{ia}, pays taxes, and chooses private consumption c_{ia}
- When insurance purchases are made (middle of period), budget constraint is

$$\int_E Q(\varepsilon) B(\varepsilon) d\varepsilon = 0 \tag{15}$$

- When paying taxes and choosing consumption (end of period), budget constraint is

$$\begin{aligned} c_{ia} &= y_{ia} - T(y_{ia}) + B(\varepsilon_{ia}) \\ &= y_{ia} - y_{ia} + \lambda y_{ia}^{1-\tau} + B(\varepsilon_{ia}) \\ &= \lambda[p(s_i)\exp(\alpha_{ia} + \varepsilon_{ia})h_{ia}]^{1-\tau} + B(\varepsilon_{ia}) \end{aligned} \tag{16}$$

- Thus, given s_i, an agent solves

$$\begin{aligned} \max_{\{c_{ia},h_{ia}\}_{a=1,\dots,\infty}} & E_0 \sum_{a=0}^{\infty} (\beta\delta)^a u_i(c_{ia}, h_{ia}, G) \\ \text{s.t. } 0 &= \int_E Q(\varepsilon)B(\varepsilon)d\varepsilon \\ c_{ia} &= \lambda[p(s_i)\exp(\alpha_{ia} + \varepsilon_{ia})h_{ia}]^{1-\tau} + B(\varepsilon_{ia}) \end{aligned}$$

3.10 A Special Case: The Representative Agent's Problem

- Consider the representative agent case of this model.
- The rep agent problem entails no cross-sectional dispersion in disutility of work effort ($v_\varphi = 0$), no idiosyncratic labor efficiency permanent shocks ($v_\omega = 0$), no idiosyncratic labor efficiency transitory shocks ($v_\varepsilon = 0$), and perfectly elastic production across skill types ($\theta = \infty$).
- The relevant state variable for the general model is the skill investment and the current level of the productivity random walk, but the rep agent both are degenerate, thus the rep agent problem a static consumption/leisure choice where (λ, g, τ) are given:

$$\max_{C,H} \log C - \frac{H^{1+\sigma}}{1+\sigma} + \chi \log G \tag{17}$$

$$\text{s.t. } C = \lambda H^{1-\tau}$$

$$\implies \max_{H} \log(\lambda H^{1-\tau}) - \frac{H^{1+\sigma}}{1+\sigma} + \chi \log G$$

- FOC wrt H

$$\begin{aligned} \frac{(1-\tau)\lambda H^{-\tau}}{\lambda H^{1-\tau}} &= H^\sigma \\ \implies (1-\tau) &= H^{\sigma+1} \\ \implies \log H^{RA}(\tau) &= \frac{1}{1+\sigma}\log(1-\tau) \end{aligned} \tag{18}$$

  Taking logs of both sides of the budget constraint and substituting:

$$\begin{aligned} \log C^{RA}(g,\tau,\lambda) &= \log\lambda + (1-\tau)\log H^{RA}(\tau) \\ &= \log\lambda + \frac{1-\tau}{1+\sigma}\log(1-\tau) \end{aligned} \tag{19}$$

- With degenerate skill distribution, aggregate output equals aggregate hours $Y = H$. The government budget constraint implies

$$
\begin{aligned}
G = gY = gH = Y - \lambda Y^{1-\tau} = H - \lambda H^{1-\tau} \\
\implies g = 1 - \lambda H^{-\tau} \\
\implies \lambda(g, \tau) = H^{\tau}(1-g)
\end{aligned}
$$

Substituting into (19),

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

- More progressivity lowers rep agent labor supply:

$$
\frac{\partial \log H^{RA}(\tau)}{\partial \tau} = \frac{-1}{(1+\sigma)(1-\tau)} < 0
$$

- More progressivity lowers rep agent consumption:

$$
\frac{\partial \log C^{RA}(\tau)}{\partial \tau} = \frac{-1}{(1+\sigma)(1-\tau)} < 0
$$

- At high enough levels of progressivity, the rep agent stops working

$$
\begin{aligned}
\lim_{\tau \to 1} H^{RA}(\tau) &= \lim_{\tau \to 1} \max\{0, \frac{1}{1+\sigma}\log(1-\tau)\} \\
&= \max\{0, -\infty\} \\
&= 0
\end{aligned}
$$

4 Equilibrium

- Recursive formulation to define stationary competitive equilibrium
- Individual state variables:
  - (κ, φ) for skill accumulation decision at $a = 0$
  - (φ, α, s) for beginning-of-period insurance claims purchasing decisions
  - $(\varphi, \alpha, \varepsilon, \bar{B})$ for end-of-period consumption and labor supply decisions where $\bar{B} = B(\varepsilon; \varphi, \alpha, s)$

- Given (g, τ), a stationary recursive competitive equilibrium is a tax level λ; asset prices $Q(\cdot)$; skill prices $p(s)$; decision rules $s(\kappa, \varepsilon), c(\varphi, \alpha, \varepsilon, s), h(\varphi, \alpha, \varepsilon, s)$, and $B(\cdot; \varphi, \alpha, s)$; and aggregate quantities $N(s)$ such that:

  1. HHs solve their problem and $s(\kappa, \varepsilon), c(\varphi, \alpha, \varepsilon, s), h(\varphi, \alpha, \varepsilon, s)$, and $B(\cdot; \varphi, \alpha, s)$ are the associated decision rules.
  2. Labor markets for each skill type clear and $p(s)$ is the value of the marginal product from an additional unit of effective hours of skill type s:

$$p(s) = \left(\frac{Y}{N(s) \cdot m(s)} \right)^{\frac{1}{\theta}}$$

  3. Asset markets clear, and the prices $Q(\cdot)$ of insurance claims are actuarially fair.
  4. Government budget constraint holds (i.e., λ satisfies (13)).

Proposition 1. *The equilibrium hours-worked allocation is*

$$\log h(\varphi, \varepsilon; \tau) = \log H^{RA}(\tau) - \varphi + \frac{1}{\hat{\sigma}} \varepsilon - \frac{1}{\hat{\sigma}(1-\tau)} \mathcal{M}(v_\varepsilon; \tau) \tag{20}$$

where $\mathcal{M}(v_\varepsilon; \tau) = \frac{(1-\tau)(1-\tau(1+\hat{\sigma}))}{\hat{\sigma}} \frac{v_\varepsilon}{2}$. *The consumption allocation is*

$$\log c(\varphi, \alpha, s; g, \tau) = \log[C^{RA}(\tau)\vartheta(\tau)] + (1-\tau)[\log p(s; \tau) + \alpha - \varphi] + \mathcal{M}(v_\varepsilon; \tau) \tag{21}$$

where $\vartheta(\tau)$ *is common across agents.*

Proof of Proposition 1:

- The roadmap of the proof is to (1) separate the economy into "island" for each (φ, s, α), (2) solve the static "island planner problem", and (3) verify that this allocation satisfies the conditions of the CE. The intuition behind this result is that, since the market is complete wrt to ε, so the CE allocation is efficient, so it can be computed as the solution to an island-specific planner problem.

- Since each island transfers zero net financial wealth between periods and preferences are time separable, each island-specific planner problem is static.

- Island-specific resource constraint is island-aggregate consumption equals island-aggregate after-tax income:

$$\begin{aligned}
\int_E c(\varepsilon) dF_\varepsilon &= \lambda \int_E [p(s) \exp(\alpha + \varepsilon) h(\varepsilon)]^{1-\tau} dF_\varepsilon \\
&= \lambda \int_E \exp[(1-\tau)(\log p(s) + \alpha + \varepsilon)] h(\varepsilon)^{1-\tau} dF_\varepsilon \\
&= \lambda \exp[(1-\tau)(\log p(s) + \alpha)] \int_E \exp[(1-\tau)\varepsilon] h(\varepsilon)^{1-\tau} dF_\varepsilon
\end{aligned}$$

- Taking (φ, s, α), (G, λ, τ), and $p(s)$ as given, the island planner problem is

$$\begin{aligned}
\max_{\{c(\varepsilon), h(\varepsilon)\}} & \int_E \left\{ \log c(\varepsilon) - \frac{\exp[(1+\sigma)\varphi]}{1+\sigma} h(\varepsilon)^{1+\sigma} + \chi \log G \right\} dF_\varepsilon \\
\text{s.t.} & \int_E c(\varepsilon) dF_\varepsilon = \lambda \exp[(1-\tau)(\log p(s) + \alpha)] \int_E \exp[(1-\tau)\varepsilon] h(\varepsilon)^{1-\tau} dF_\varepsilon
\end{aligned}$$

- Let γ be the multiplier on the resource constraint. The lagrangian is

$$\begin{aligned}
\mathcal{L} = & \int_E \left\{ \log c(\varepsilon) - \frac{\exp[(1+\sigma)\varphi]}{1+\sigma} h(\varepsilon)^{1+\sigma} + \chi \log G \right\} dF_\varepsilon \\
& + \gamma \left[\lambda \exp[(1-\tau)(\log p(s) + \alpha)] \int_E \exp[(1-\tau)\varepsilon] h(\varepsilon)^{1-\tau} dF_\varepsilon - \int_E c(\varepsilon) dF_\varepsilon \right]
\end{aligned}$$

- FOC wrt $c(\varepsilon)$:

$$\frac{1}{c(\varepsilon)} = \gamma \implies c(\varepsilon) = c = \frac{1}{\gamma} \quad \forall \varepsilon$$

Thus, consumption is the same across agents on an island regardless of their realization of $\varepsilon \implies$ perfect consumption risk sharing

- FOC wrt $h(\varepsilon)$:

$$\begin{aligned}
\exp[(1+\sigma)\varphi] h(\varepsilon)^{\sigma} &= \gamma\lambda(1-\tau) \exp[(1-\tau)(\log p(s) + \alpha)] \exp[(1-\tau)\varepsilon] h(\varepsilon)^{-\tau} \\
\implies h(\varepsilon)^{\sigma+\tau} &= \gamma\lambda(1-\tau) \exp[(1-\tau)(\log p(s) + \alpha) - (1+\sigma)\varphi] \exp[(1-\tau)\varepsilon] \\
\implies h(\varepsilon) &= \gamma^{\frac{1}{\sigma+\tau}} [\lambda(1-\tau)]^{\frac{1}{\sigma+\tau}} \exp\left[\frac{1-\tau}{\sigma+\tau} (\log p(s) + \alpha) - \frac{1+\sigma}{\sigma+\tau} \varphi \right] \exp\left[\frac{1-\tau}{\sigma+\tau} \varepsilon \right]
\end{aligned}$$

- Substituting both conditions into the resource constraint The resource constraint can be expressed as

$$
\begin{aligned}
c &= \lambda \exp[(1-\tau)(\log p(s)+\alpha)] \int_E \exp[(1-\tau)\varepsilon] h(\varepsilon)^{1-\tau} dF_\varepsilon \\
\implies \gamma^{-1} &= \lambda \exp[(1-\tau)(\log p(s)+\alpha)] \\
&\quad \cdot \int_E \exp[(1-\tau)\varepsilon]\gamma^{\frac{1-\tau}{\sigma+\tau}}[\lambda(1-\tau)]^{\frac{1-\tau}{\sigma+\tau}} \\
&\quad \exp\left[\frac{(1-\tau)^2}{\sigma+\tau}(\log p(s)+\alpha) - \frac{(1+\sigma)(1-\tau)}{\sigma+\tau}\varphi\right] \exp\left[\frac{(1-\tau)^2}{\sigma+\tau}\varepsilon\right] dF_\varepsilon \\
\implies \gamma^{\frac{-\sigma-1}{\sigma+\tau}} &= \lambda^{\frac{1+\sigma}{\sigma+\tau}}(1-\tau)^{\frac{1-\tau}{\sigma+\tau}} \exp[(1-\tau)(\log p(s)+\alpha)] \exp\left[\frac{(1-\tau)^2}{\sigma+\tau}(\log p(s)+\alpha) - \frac{(1+\sigma)(1-\tau)}{\sigma+\tau}\varphi\right] \\
&\quad \cdot \int_E \exp\left[\left(\frac{(1-\tau)^2+(1-\tau)(\sigma+\tau)}{\sigma+\tau}\right)\varepsilon\right] dF_\varepsilon \\
\gamma^{\frac{-\sigma-1}{\sigma+\tau}} &= \lambda^{\frac{1+\sigma}{\sigma+\tau}}(1-\tau)^{\frac{1-\tau}{\sigma+\tau}} \exp\left[\frac{(1-\tau)(1+\sigma)}{\sigma+\tau}(\log p(s)+\alpha) - \frac{(1+\sigma)(1-\tau)}{\sigma+\tau}\varphi\right] \\
&\quad \cdot \int_E \exp\left[\left(\frac{(1-\tau)(1+\sigma)}{\sigma+\tau}\right)\varepsilon\right] dF_\varepsilon
\end{aligned}
$$

- Focusing in on the last term, we can solve using MGF of normal distribution:[1]

$$
\begin{aligned}
\int_E \exp\left[\frac{(1-\tau)(1+\sigma)}{\sigma+\tau}\varepsilon\right] dF_\varepsilon &= \exp\left[\frac{(1-\tau)(1+\sigma)}{\sigma+\tau}\frac{-v_\varepsilon}{2} + \frac{1}{2}\frac{(1-\tau)^2(1+\sigma)^2}{(\sigma+\tau)^2}v_\varepsilon\right] \\
&= \exp\left[\frac{v_\varepsilon}{2}\left(\frac{(1-\tau)^2(1+\sigma)^2}{(\sigma+\tau)^2} - \frac{(1-\tau)(1+\sigma)}{\sigma+\tau}\right)\right] \\
&= \exp\left[\frac{v_\varepsilon}{2}\left(\frac{(1-\tau)^2(1+\sigma)^2-(1-\tau)(1+\sigma)(\sigma+\tau)}{(\sigma+\tau)^2}\right)\right] \\
&= \exp\left[\frac{v_\varepsilon}{2}\left(\frac{(1-\tau)(1+\sigma)[(1-\tau)(1+\sigma)-(\sigma+\tau)]}{(\sigma+\tau)^2}\right)\right] \\
&= \exp\left[\frac{v_\varepsilon}{2}\left(\frac{(1-\tau)(1+\sigma)(1-2\tau-\tau\sigma)}{(\sigma+\tau)^2}\right)\right]
\end{aligned}
$$

[1] If $X \sim N(\mu, \sigma^2)$, then $E[\exp(aX)] = \exp(a\mu + \frac{a^2\sigma^2}{2})$.

Plugging back into resource constraint:

$$\gamma^{\frac{-\sigma-1}{\sigma+\tau}} = \lambda^{\frac{1+\sigma}{\sigma+\tau}}(1-\tau)^{\frac{1-\tau}{\sigma+\tau}} \exp\left[\frac{(1-\tau)(1+\sigma)}{\sigma+\tau}(\log p(s)+\alpha-\varphi)\right]$$
$$\cdot \exp\left[\frac{v_\varepsilon}{2}\left(\frac{(1-\tau)(1+\sigma)(1-2\tau-\tau\sigma)}{(\sigma+\tau)^2}\right)\right]$$
$$\implies \gamma^{-1} = \lambda(1-\tau)^{\frac{1-\tau}{1+\sigma}} \exp\left[(1-\tau)(\log p(s)+\alpha-\varphi)\right] \exp\left[\frac{v_\varepsilon}{2}\left(\frac{(1-\tau)(1-2\tau-\tau\sigma)}{(\sigma+\tau)}\right)\right]$$
$$\log c = \log\lambda + \frac{1-\tau}{1+\sigma}\log(1-\tau) + (1-\tau)[\log p(s)+\alpha-\varphi] + \frac{v_\varepsilon}{2}\left(\frac{(1-\tau)(1-2\tau-\tau\sigma)}{(\sigma+\tau)}\right)$$

- Defining the last term as $\mathcal{M}(v_\varepsilon;\tau)$:

$$\begin{aligned}\mathcal{M}(v_\varepsilon;\tau) &= \frac{v_\varepsilon}{2}\left(\frac{(1-\tau)(1-2\tau-\tau\sigma)}{(\sigma+\tau)}\right)\\ &= \frac{v_\varepsilon}{2}\left(\frac{1-2\tau-\tau\sigma}{\hat{\sigma}}\right)\\ &= \frac{v_\varepsilon}{2}\left(\frac{1-\tau-\tau(1+\sigma)}{\hat{\sigma}}\right)\\ &= \frac{(1-\tau)(1-\tau(1+\hat{\sigma}))}{\hat{\sigma}}\frac{v_\varepsilon}{2}\end{aligned}$$

- Recall that $\log C^{RA} = \log(1-g) + \frac{1}{1+\sigma}\log(1-\tau)$, so we can rewrite consumption as

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

where $\log\vartheta \equiv \log\lambda - \log(1-g) - \frac{\tau}{1+\sigma}\log(1-\tau)$. This equation matches (21).

- Turning back to the intratemporal FOC and taking logs and substituting in consumption:

$$\begin{aligned}\log h(\varepsilon) &= \frac{1}{\sigma+\tau}\log\gamma + \frac{1}{\sigma+\tau}\log\lambda + \frac{1}{\sigma+\tau}\log(1-\tau) + \frac{1-\tau}{\sigma+\tau}(\log p(s)+\alpha) - \frac{1+\sigma}{\sigma+\tau}\varphi + \frac{1-\tau}{\sigma+\tau}\varepsilon\\ &= \frac{-1}{\sigma+\tau}\left[\log\lambda + \frac{1-\tau}{1+\sigma}\log(1-\tau) + (1-\tau)[\log p(s)+\alpha-\varphi] + \mathcal{M}(v_\varepsilon;\tau)\right]\\ &+ \frac{1}{\sigma+\tau}\log\lambda + \frac{1}{\sigma+\tau}\log(1-\tau) + \frac{1-\tau}{\sigma+\tau}(\log p(s)+\alpha) - \frac{1+\sigma}{\sigma+\tau}\varphi + \frac{1-\tau}{\sigma+\tau}\varepsilon\\ &= \left[\frac{-1}{\sigma+\tau} + \frac{1}{\sigma+\tau}\right]\log\lambda + \left[\frac{-1(1-\tau)}{(\sigma+\tau)(1+\sigma)} + \frac{1}{\sigma+\tau}\right]\log(1-\tau)\\ &+ \left[\frac{-(1-\tau)}{(\sigma+\tau)} + \frac{(1-\tau)}{(\sigma+\tau)}\right][\log p(s)+\alpha] + \left[\frac{-1(1-\tau)}{\sigma+\tau} - \frac{1+\sigma}{\sigma+\tau}\right]\varphi + \frac{1-\tau}{\sigma+\tau}\varepsilon - \frac{1}{\sigma+\tau}\mathcal{M}(v_\varepsilon;\tau)\\ &= \frac{1}{1+\sigma}\log(1-\tau) - \varphi + \frac{1-\tau}{\sigma+\tau}\varepsilon - \frac{1}{\sigma+\tau}\mathcal{M}(v_\varepsilon;\tau)\end{aligned}$$

- Recall that $\log H^{RA}(\tau) = \frac{1}{1+\sigma}\log(1-\tau)$, so this equation can be rewritten to match (20):

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

□

Proposition 2. *Hello*

Corollary 1. *Hello*

Corollary 2. *Hello*

4.1 Efficiency

Proposition 3. *Hello*

5 Welfare Effects of Tax Reform

5.1 Social Welfare Function

Proposition 4. *Hello*

Corollary 3. *Hello*

Corollary 4. *Hello*

Corollary 5. *Hello*

Corollary 6. *Hello*

Corollary 7. *Hello*

Corollary 8. *Hello*

5.2 Decomposition of the Social Welfare Function

5.2.1 Welfare of the Representative Agent

Proposition 5. *Hello*

5.2.2 Welfare from Skill Investment

5.2.3 Welfare from Preference Heterogeneity and Uninsurable Wage Risk

5.2.4 Welfare from Insurable Wage Risk

5.3 When Should Taxes Be Progressive?

Proposition 6. *Hello*

5.4 Optimal Marginal Tax Rate at the Top

6 Quantitative Analysis

6.1 Parameterization

6.2 Results

6.3 Progressivity When Past Skill Investment is Fixed

6.4 Modeling Public Consumption

6.5 Inequality Aversion

Proposition 7. *Hello*

6.6 Political-Economic Determination of Progressivity

Proposition 8. *Hello*

Proposition 9. *Hello*

7 Skill Investment Constraints

8 Empirical Evidence

9 Conclusions

Appendix: Lecture Notes on HSV (2014) and BK (2021)

Job presented a simplified version of HSV (2014) and BK (2021), which are closely related papers. I have included these lecture notes because I found them helpful in understanding the more complicated HSV (2017).

Question: We have data on consumption, hours, and wages $\{c_i, h_i, w_i\}_i$. Can we perfectly rationalize these data?

Environment

1. Demography
   - HHs
   - Time: $t = 0, 1, ...$
   - Perpetual youth: $P(death) = 1 - \delta$ and $P(survival) = \delta$
   - Every time t, mass $1 - \delta$ is born
   - Age $a = 0, 1, ...$
   - Cohort $j \in (-\infty, \infty)$
2. Preferences

$$E_j \sum_{t=j}^{\infty} (\beta\delta)^{t-j} u^j(c_t^j, h_t^j) = E_j \sum_{t=j}^{\infty} (\beta\delta)^{t-j} \left[\frac{(c_t^j)^{1-\gamma} - 1}{1-\gamma} - \frac{[\exp(B^j) h_t^j]^{1+\frac{1}{\eta}}}{1+\frac{1}{\eta}} \right]$$

   - η is the Frisch elasticity of labor supply
   - Here, we generally focus on $\gamma \to 1 \implies$ log preferences over consumption

$$u^j(c_t^j, h_t^j) = \log c_t^j - \frac{[\exp(B^j) h_t^j]^{1+\frac{1}{\eta}}}{1+\frac{1}{\eta}}$$

3. Technology
   - Production
   - $y = zh$ where y is earnings, z is wage rate per hour, and h is hours
   - $z = \exp(\alpha + \varepsilon)$ where α^j is permanent, $\varepsilon_t^j = \nu_t^\varepsilon$ varies in every time period and across agent (idiosyncratic and iid across time and agents)
   - Distributions are $\nu_t^\varepsilon \sim \Phi_t^\varepsilon$, $\alpha^j \sim \Phi_j^\alpha$, $B^j \sim \Phi_j^B$
4. Information structure
   - Perfect foresight: $\{\Phi_t^\varepsilon, \Phi_j^\alpha, \Phi_j^B\}$ are known

5. Equilibrium

- A HH denoted by $i = \{j, B^j, \alpha^j, \{\nu_t^j\}\}$
- Define island ℓ as the collection of households with identical $\{j, B^j, \alpha^j\}$
- Define states $s_t^j := \varepsilon_t^j$ and histories $s^{j,t} = (s_j^j, ..., s_t^j)$
- Agents can trade one-period bonds with people on their islands. Arrow securities with corresponding prices $q_b^\ell(\varepsilon_{t+1}^j)$
- Agents can also trade one-period bonds across islands. Arrow securities with corresponding prices $q_x(\zeta_{t+1}^j)$ where $\zeta_t^j := \varepsilon_t^j$
- HH problem:

$$\max_{\{c_t^j, h_t^j, b^\ell(s_{t+1}^j), x(\zeta_{t+1}^j)\}_{t=j}^\infty} E_j \sum_{t=j}^\infty (\beta\delta)^{t-j} u^j(c_t^j, h_t^j)$$

$$\text{s.t. } c_t^j + \int_{s_{t+1}^j} q_b^\ell(s_{t+1}^j) b^\ell(s_{t+1}^j) ds_{t+1}^j + \int_{\zeta_{t+1}^j} q_x(s_{t+1}^j) x(\zeta_{t+1}^j) d\zeta_{t+1}^j \leq y_t^j + b^\ell(s_t^j) + x(\zeta_t^j)$$

- An equilibrium is an allocation is x_j and prices $\{q_b^\ell(s_{t+1}^j\}_{\ell,t}$ and $\{q_x(\zeta_{t+1}^j\}_t$ such that
  - HH solves HH problem given prices
  - Asset markets clear

$$\int b^\ell(s_{t+1}^j) d\Phi_t(i) = 0 \quad \forall s_{t+1}^j, \ell$$

$$\int x(\zeta_{t+1}^j) d\Phi_t(i) = 0 \quad \forall \zeta_{t+1}^j$$

- Steps: (1) formulate auxiliary problem (island planner problem) and (2) verify that this auxiliary problem satisfy the conditions of an competitive equilibrium

Island planner problem (auxiliary problem)

- For every island ℓ, the island planner solves static problem

$$\max_{\{c,h\}_{\varepsilon_t^j}} \int u^j(c_t^j, h_t^j) d\Phi_t(\varepsilon_t)$$

$$\text{s.t. } \int c(\varepsilon_t^j) d\Phi = \int z(\varepsilon_t^j) h(\varepsilon_t^j) d\Phi$$

- Attach λ on the RC
- FOC wrt $c(\varepsilon_t^j)$:

$$c(\varepsilon_t^j)^{-1} = \lambda$$

  Thus, there is perfect insurance on an island

- FOC wrt $h(\varepsilon_t^j)$:

$$(\exp(B_j))^{\frac{1}{\eta}+1} h_j^{1/\eta} = z(\varepsilon_t^j)\lambda$$

  Thus, hours change with ε_t^j. High ε_t^j work more.

- The RC and consumption FOC imply

$$\begin{aligned}
\frac{1}{\lambda} &= \int z^{1+\eta}\lambda^\eta \exp(-(1+\eta)B) d\Phi \\
&= \lambda^\eta \exp(-(1+\eta)B)\exp((1+\eta)\alpha) \int \exp((1+\eta)\varepsilon) d\Phi \\
\implies 1 &= \lambda \exp(-B)\exp(\alpha)\left[\int \exp((1+\eta)\varepsilon) d\Phi\right]^{\frac{1}{1+\eta}} \\
\implies c &= \exp(\alpha - B)\underbrace{\left[\int \exp((1+\eta)\varepsilon) d\Phi\right]^{\frac{1}{1+\eta}}}_{\equiv \mathcal{C}}
\end{aligned}$$

$$\begin{aligned}
\exp((1+\eta)B)h(\varepsilon) &= [\exp(\alpha+\varepsilon-\alpha+B)]^\eta / \mathcal{C}^\eta \\
h(\varepsilon) &= \frac{\varepsilon\eta}{\mathcal{C}^\eta \exp(B)}
\end{aligned}$$

- Comparative statistics:$\uparrow \varepsilon \implies \uparrow h(\varepsilon)$ and $\uparrow B \implies \downarrow h(\varepsilon)$
- α doesn't show up in h (kind of weird). Why? Income and substitution effects perfectly cancel out.

$$\begin{aligned}
q_\ell(\varepsilon_{t+1}^j) &= \frac{u_c(c_{t+1}(\varepsilon_{t+1}^j))}{u_c(c_t(\varepsilon_t^j))}\beta\delta \qquad \underbrace{\pi_t(\varepsilon_{t+1})}_{\text{Doesn't vary by islands}} \\
&= \frac{c_t^j(\varepsilon_t^j)}{c_{t+1}^j(\varepsilon_{t+1}^j)}\beta\delta\pi_t(\varepsilon_{t+j}) \\
&= \exp(\alpha - B - \alpha + B)\frac{\mathcal{C}_t}{\mathcal{C}_{t+1}}\beta\delta\pi_t(\varepsilon_{t+1}) \\
&= \frac{\mathcal{C}_t}{\mathcal{C}_{t+1}}\beta\delta\pi_t(\varepsilon_{t+1})
\end{aligned}$$

  Price is the same across islands.

- Risk-free asset price (let $\mathcal{E}_{t+1}$ be the support of ε_{t+1}^j):

$$\begin{aligned}
q_\ell(\mathcal{E}_{t+1}) &= \int \frac{u_c^j(c_{t+1}^j(\varepsilon_{t+1}^j))}{u_c^j(c_t^j(\varepsilon_t^j))}\beta\delta d\Phi_j^\varepsilon \\
&= \beta\delta\frac{\mathcal{C}_t}{\mathcal{C}_{t+1}}
\end{aligned}$$

Equilibrium Characterization

- From the equilibrium conditions, we can map data on wages, hours, and consumption $(\hat{w}_i, \hat{h}_i, \hat{c}_i)$ to structural parameters $(\alpha_i, \varepsilon_i, B_i)$:

$$\underbrace{\hat{w}_i}_{\text{data}} = \underbrace{\exp(\alpha_i + \varepsilon_i)}_{\text{fn of structural parameter}}$$

$$\underbrace{\hat{h}_i}_{\text{data}} = \underbrace{\frac{\exp(\eta\varepsilon_i)}{\mathcal{C}^\eta \exp(B_i)}}_{\text{fn of structural parameter}}$$

$$\underbrace{\hat{c}_i}_{\text{data}} = \underbrace{\exp(\alpha_i - B_i)\mathcal{C}}_{\text{fn of structural parameter}}$$

- We can rewrite structural parameters $(\alpha_i, \varepsilon_i, B_i)$ as functions of data $(\hat{w}_i, \hat{h}_i, \hat{c}_i)$:

$$\frac{\hat{c}_i \hat{w}_i^\eta}{\hat{h}_i} = \frac{\exp(\alpha_i - B_i)\exp(\eta(\alpha_i + \varepsilon_i))}{\exp(-B_i)\exp(\eta\varepsilon_i)}$$

$$= \exp((1+\eta)\alpha_i)$$

$$\implies \underbrace{\frac{1}{1+\eta}\log\left[\frac{\hat{c}_i \hat{w}_i^\eta}{\hat{h}_i}\right]}_{\text{fn of data}} = \underbrace{\alpha_i}_{\text{structural parameter}} \tag{22}$$

$$\log(\hat{w}_i) - \alpha_i = \varepsilon_i$$

$$\implies \underbrace{\log(\hat{w}_i) - \frac{1}{1+\eta}\log\left[\frac{\hat{c}_i \hat{w}_i^\eta}{\hat{h}_i}\right]}_{\text{fn of data}} = \underbrace{\varepsilon_i}_{\text{structural parameter}} \tag{23}$$

$$\log(\hat{c}_i) = \alpha_i - B_i + \log(\mathcal{C})$$

$$\implies \underbrace{\frac{1}{1+\eta}\log\left[\frac{\hat{c}_i \hat{w}_i^\eta}{\hat{h}_i}\right] + \log(\mathcal{C}) - \log(\hat{c}_i)}_{\text{fn of data}} = \underbrace{B_i}_{\text{structural parameter}} \tag{24}$$

- We can use common estimates of the Frisch elasticity of labor supply η from the literature
- (24) Recall that

$$\mathcal{C} = \left[\int \exp((1+\eta)\varepsilon)d\Phi\right]^{\frac{1}{1+\eta}}$$

Which can be estimated as

$$\mathcal{C} \approx \left[\frac{1}{n}\sum \exp((1+\eta)\varepsilon_i)\right]^{\frac{1}{1+\eta}}$$

$$= \left[\frac{1}{n}\sum \exp\left((1+\eta)\left(\log(\hat{w}_i) - \frac{1}{1+\eta}\log\left[\frac{\hat{c}_i \hat{w}_i^\eta}{\hat{h}_i}\right]\right)\right)\right]^{\frac{1}{1+\eta}}$$