Insurance and Taxation Over the Life Cycle

Emmanuel Farhi   Iván Werning

January 23, 2020

Introduction

- Uncertainty of lifetime earnings
- Gradually resolved over time
- How to set taxes to insure?
- Static models (Mirrlees 1971, Diamond 1998, Saez 2002):
  - symmetric treatment of redistribution and insurance
  - how to interpret a period?
- Dynamic context?

Introduction

- Most progress: particular cases or focus on saving distortions
- This paper: labor distortions and saving distortions in general setting
  - theoretical: novel formula for labor taxes
  - numerical simulations

Introduction

- Questions regarding optimum:
  - taxes depend on age?
  - taxes depend on past history?
  - tax system progressive or regressive

Introduction

- Optimum requires sophisticated taxes
- Simpler taxes?
  - use optimum to construct simpler taxes
  - finding: relatively simple taxes get most gains

References

- Static Taxation: Mirrlees (1971), Diamond (1998), Saez (2002), Werning (2007).
- Dynamic Taxation: Diamond-Mirrlees (1978); Albanesi-Sleet (2004), Shimer-Werning (2008), Ales-Maziero (2009), Golosov-Troshkin-Tsyvinsky (2010).
- Method: Fernandes-Phelan (2000), Werning (2002), Abraham-Pavoni (2008), Kapicka (2009), Williams (2009), Pavan-Segal-Toikka (2009).
- Age Dependent taxation: Erosa-Gervais (2002), Kremer (2002), Weinzierl (2008).

Preferences and Technology

- Utility

$$U(\{c,y\}) = \mathbb{E}_0 \sum_{t=1}^{T} \beta^{t-1} u^t(c_t, y_t; \theta_t)$$

- Cost

$$\mathbb{E}_0 \sum_{t=1}^{T} R^{-(t-1)}(c_t - y_t)$$

- Life cycle:
  - work phase

$$t \leq T_E \qquad u^t(c,y;\theta) = \tilde{u}(c, y/\theta)$$

  - retirement

$$T_E < t \leq T \qquad u^t(c,y;\theta) = \begin{cases} \tilde{u}(c,0) & y = 0 \\ -\infty & y > 0 \end{cases}$$

Uncertainty and Information

- θ_t private info
- $\{\theta\}$ markov:
  - support: $[\underline{\theta}_t(\theta_{t-1}), \overline{\theta}_t(\theta_{t-1})]$
  - differentiable density: $f^t(\theta_t|\theta_{t-1})$
- Start with:
  - fixed support $[\underline{\theta}, \overline{\theta}]$
  - relax later...

Planning Problem

$$K_0(v) \equiv \min_{\{c,y\}} \mathbb{E}_0 \sum_{t=1}^{T} R^{-(t-1)}(c_t - y_t)$$

$$\text{s.t.} \quad U(\{c,y\}) \geq v$$

$$U(\{c,y\}) \geq U(\{c^\sigma, y^\sigma\}) \quad \forall \sigma \in \Sigma$$

- Not tractable except special cases (for example, i.i.d.)
- Approach here:
  - solve relaxed program with local ICs
  - verify global ICs

Local ICs

- Continuation utility

$$w(\theta^t) = u(c(\theta^t), y(\theta^t); \theta_t) + \beta v(\theta^t)$$

$$v(\theta^t) \equiv \int w(\theta^{t+1}) f^{t+1}(\theta_{t+1}|\theta_t) d\theta_{t+1}.$$

- Necessary conditions for IC:

$$\frac{\partial}{\partial \theta_t} w(\theta^t) = u_\theta(c(\theta^t), y(\theta^t); \theta_t) + \beta \Delta(\theta^t)$$

$$\Delta(\theta^t) \equiv \int w(\theta^{t+1}) f^{t+1}_{\theta_t}(\theta_{t+1}|\theta_t) d\theta_{t+1}.$$

- Dynamic generalization of Envelope condition of Mirrlees (1971) and Milgrom and Segal (2002)
- Kapicka (2009), Williams (2009), Pavan, Segal and Toikka (2009)

A Recursive First-Order Approach

$$K(v,\Delta,\theta_-,t) = \min \int \left[c(\theta) - y(\theta) + \frac{1}{R} K(v(\theta),\Delta(\theta),\theta,t+1)\right] f^t(\theta|\theta_-)\,d\theta$$

s.t.

$$v = \int w(\theta) f^t(\theta|\theta_-)\,d\theta \qquad \text{where} \qquad w(\theta) = u(c(\theta),y(\theta);\theta) + \beta v(\theta)$$

$$\dot{w}(\theta) = u_\theta(c(\theta),y(\theta);\theta) + \beta \Delta(\theta)$$

$$\Delta = \int w(\theta) f^t_{\theta_-}(\theta|\theta_-)\,d\theta$$

- Kapicka (2009), Williams (2009), Pavan, Segal and Toikka (2009)

Fernandes-Phelan

- warm up: finite shocks $\theta^1 < \theta^2 < \cdots < \theta^N$
- Fernandes-Phelan:

$$K(v_1, v_2, \dots, v_N, \theta^l, t) = \min \sum_n [c(\theta^n) - y(\theta^n) + \frac{1}{R} K(v_1(\theta^n), v_2(\theta^n), \dots, v_N(\theta^n), \theta^n, t+1)] f^t(\theta^n | \theta^l)$$

s.t. $\forall n, m$:

$$u(c(\theta^n), y(\theta^n); \theta) + \beta v_n(\theta^n) \geq u(c(\theta^m), y(\theta^m); \theta^n) + \beta v_n(\theta^m)$$

$$v_k = \sum_\theta w(\theta) f(\theta | \theta^k) \qquad k = 1, 2, \dots, N$$

$$w(\theta^n) = u(c(\theta^n), y(\theta^n); \theta^n) + \beta v_n(\theta^n)$$

Fernandes-Phelan Local ICs

- warm up: finite shocks $\theta^1 < \theta^2 < \cdots < \theta^N$
- Fernandes-Phelan: (relaxed)

$$K(v_l, v_{l+1}, \theta^l, t) = \min \sum_n [c(\theta^n) - y(\theta^n) + \frac{1}{R} K(v_n(\theta^n), v_{n+1}(\theta^n), \theta^n, t+1)] f^t(\theta^n | \theta^l)$$

s.t. $\forall n$:

$$u(c(\theta^n), y(\theta^n), \theta^n) + \beta v_n(\theta^{n-1}) \geq u(c(\theta^{n-1}), y(\theta^{n-1}), \theta^n) + \beta v_n(\theta^{n-1})$$

$$v_k = \sum_\theta w(\theta) f(\theta | \theta^k) \qquad k = l, l+1$$

$$w(\theta^n) = u(c(\theta^n), y(\theta^n); \theta^n) + \beta v_n(\theta^n)$$

Wedges

- Intertemporal wedge

$$1 = \beta R(1 - \tau_{K,t-1})\mathbb{E}_{t-1}\left[\frac{\hat{u}_c^t(c_t, y_t; \theta_t)}{\hat{u}_c^{t-1}(c_{t-1}, y_{t-1}; \theta_{t-1})}\right]$$

- Labor wedge

$$1 = (1 - \tau_{L,t})\frac{\hat{u}_c^t(c_t, y_t; \theta_t)}{-\hat{u}_y^t(c_t, y_t; \theta_t)}$$

Intertemporal Wedge: Inverse Euler

Assumption

Separable utility: $u^t(c, y, \theta) = \hat{u}^t(c) - \hat{h}^t(y, \theta)$.

Proposition

Inverse Euler holds:

$$\frac{1}{\hat{u}^{t-1\prime}(c_{t-1})} = \frac{1}{\beta R}\mathbb{E}_{t-1}\left[\frac{1}{\hat{u}^{t\prime}(c_t)}\right]$$

- Intertemporal wedge

$$\tau_{K,t-1} \geq 0$$

Labor Wedge: A Simple Formula

Assumption

Isoelastic disutility of work $\hat{h}^t(y,\theta) = (\kappa/\alpha)(y/\theta)^{\alpha}$.

Assumption

AR(1) productivity

$$\log(\theta_t) = \rho \log(\theta_{t-1}) + \bar{\theta}_t + \varepsilon_t$$

Proposition

$$\mathbb{E}_{t-1}\left[\frac{\tau_{L,t}}{1-\tau_{L,t}}\frac{1}{\beta R}\frac{\hat{u}^{t-1\prime}(c_{t-1})}{\hat{u}^{t\prime}(c_t)}\right] = \rho\frac{\tau_{L,t-1}}{1-\tau_{L,t-1}} + \alpha Cov_{t-1}\left(\log(\theta_t), \frac{1}{\beta R}\frac{\hat{u}^{t-1\prime}(c_{t-1})}{\hat{u}^{t\prime}(c_t)}\right)$$

Labor Wedge: A Simple Formula

- Labor wedge formula:

$$\mathbb{E}_{t-1}\left[\frac{\tau_{L,t}}{1-\tau_{L,t}}\frac{1}{\beta R}\frac{\hat{u}^{t-1\prime}(c_{t-1})}{\hat{u}^{t\prime}(c_t)}\right]$$
$$=\rho\frac{\tau_{L,t-1}}{1-\tau_{L,t-1}}+\alpha \text{Cov}_{t-1}\left(\log(\theta_t),\frac{1}{\beta R}\frac{\hat{u}^{t-1\prime}(c_{t-1})}{\hat{u}^{t\prime}(c_t)}\right)$$

Labor Wedge: A Simple Formula

- Labor wedge formula:

$$\mathbb{E}_{t-1}\left[\frac{\tau_{L,t}}{1-\tau_{L,t}}\frac{1}{\beta R}\frac{\hat{u}^{t-1\prime}(c_{t-1})}{\hat{u}^{t\prime}(c_t)}\right]$$

$$=\rho\frac{\tau_{L,t-1}}{1-\tau_{L,t-1}}+\alpha \text{Cov}_{t-1}\left(\log(\theta_t),\frac{1}{\beta R}\frac{\hat{u}^{t-1\prime}(c_{t-1})}{\hat{u}^{t\prime}(c_t)}\right)$$

- LHS: risk-adjusted conditional expectation of $\tau_{L,t}/(1-\tau_{L,t})$

Labor Wedge: A Simple Formula

- Labor wedge formula:

$$\mathbb{E}_{t-1}\left[\frac{\tau_{L,t}}{1-\tau_{L,t}}\frac{1}{\beta R}\frac{\hat{u}^{t-1\prime}(c_{t-1})}{\hat{u}^{t\prime}(c_t)}\right] = \rho\frac{\tau_{L,t-1}}{1-\tau_{L,t-1}} + \alpha \mathrm{Cov}_{t-1}\left(\log(\theta_t), \frac{1}{\beta R}\frac{\hat{u}^{t-1\prime}(c_{t-1})}{\hat{u}^{t\prime}(c_t)}\right)$$

- LHS: risk-adjusted conditional expectation of $\tau_{L,t}/(1-\tau_{L,t})$
- RHS(1): $\{\tau_L/(1-\tau_L)\}$ inherits mean reversion of $\{\theta\}$

Labor Wedge: A Simple Formula

- Labor wedge formula:

$$\mathbb{E}_{t-1}\left[\frac{\tau_{L,t}}{1-\tau_{L,t}}\frac{1}{\beta R}\frac{\hat{u}^{t-1\prime}(c_{t-1})}{\hat{u}^{t\prime}(c_t)}\right] = \rho\frac{\tau_{L,t-1}}{1-\tau_{L,t-1}} + \alpha \text{Cov}_{t-1}\left(\log(\theta_t), \frac{1}{\beta R}\frac{\hat{u}^{t-1\prime}(c_{t-1})}{\hat{u}^{t\prime}(c_t)}\right)$$

- LHS: risk-adjusted conditional expectation of $\tau_{L,t}/(1-\tau_{L,t})$
- RHS(1): $\{\tau_L/(1-\tau_L)\}$ inherits mean reversion of $\{\theta\}$
- RHS(2): positive drift of $\{\tau_L/(1-\tau_L)\}$
  - benefit of added insurance: $\text{Cov}_{t-1}\left(\log(\theta_t), \frac{1}{\beta R}\frac{\hat{u}^{t-1\prime}(c_{t-1})}{\hat{u}^{t\prime}(c_t)}\right)$
  - incentive cost increases with Frisch elasticity $1/(\alpha-1)$

Labor Wedge: A Simple Formula

- If θ_t predictable...

$$\frac{\tau_{L,t}}{1-\tau_{L,t}} = \rho \frac{\tau_{L,t-1}}{1-\tau_{L,t-1}}$$

- Tax smoothing: $\rho = 1 \Rightarrow \tau_{L,t} = \tau_{L,t-1}$
- Mean reversion: $\rho < 1 \Rightarrow \{\tau_L\}$ reverts to zero at rate ρ

General Stochastic Process

- Define

$$\phi_t^{\log}(\theta_{t-1}) \equiv \int \log(\theta_t)\, f^t(\theta_t|\theta_{t-1})\, d\theta_t$$

Proposition

$$\mathbb{E}_{t-1}\left[\frac{\tau_{L,t}}{1-\tau_{L,t}} \frac{1}{\beta R} \frac{\hat{u}^{t-1\prime}(c_{t-1})}{\hat{u}^{t\prime}(c_t)}\right] =$$

$$\theta_{t-1} \frac{d\phi_t^{\log}}{d\theta_{t-1}} \frac{\tau_{L,t-1}}{1-\tau_{L,t-1}} + \alpha Cov_{t-1}\left(\log(\theta_t), \frac{1}{\beta R} \frac{\hat{u}^{t-1\prime}(c_{t-1})}{\hat{u}^{t\prime}(c_t)}\right)$$

Moving Support

- Moving support:

$$[\underline{\theta}_t(\theta_{t-1}), \overline{\theta}_t(\theta_{t-1})]$$

- Only difference

$$\Delta(\theta^t) = \int w(\theta^{t+1}) f_{\theta_t}^{t+1}(\theta_{t+1}|\theta_t) d\theta_{t+1}$$
$$+ \frac{d\overline{\theta}_{t+1}}{d\theta_t} w\left(\overline{\theta}_{t+1}\right) f^{t+1}\left(\overline{\theta}_{t+1}|\theta_t\right)$$
$$- \frac{d\underline{\theta}_{t+1}}{d\theta_t} w\left(\underline{\theta}_{t+1}\right) f^{t+1}\left(\underline{\theta}_{t+1}|\theta_t\right)$$

Labor Wedge at Top and Bottom

Proposition

$$
\begin{aligned}
\frac{\bar{\tau}_{L,t}}{1-\bar{\tau}_{L,t}} &= \frac{\tau_{L,t-1}}{1-\tau_{L,t-1}} \beta R \frac{\hat{u}^{t\prime}}{\hat{u}^{t-1\prime}} \frac{d \log \overline{\theta}_t}{d \log \theta_{t-1}} \\
\frac{\underline{\tau}_L}{1-\underline{\tau}_{L,t}} &= \frac{\tau_{L,t-1}}{1-\tau_{L,t-1}} \beta R \frac{\hat{u}^{t\prime}}{\hat{u}^{t-1\prime}} \frac{d \log \underline{\theta}_t}{d \log \theta_{t-1}}
\end{aligned}
$$

- Generalizes Mirrlees (1971):
  - fixed support...

$$\tau_L\left(\theta^{t-1}, \overline{\theta}_t\right) = \tau_L\left(\theta^{t-1}, \underline{\theta}_t\right) = 0$$

  - $\theta_t = \varepsilon_t \theta_{t-1}$ and $\varepsilon_t \in [\underline{\varepsilon}, \overline{\varepsilon}]$...

$$\tau_L\left(\theta^{t-1}, \overline{\theta}_t\right) \leq \tau_L\left(\theta^{t-1}\right) \leq \tau_L\left(\theta^{t-1}, \underline{\theta}_t\right)$$

Continuous Time: Approach

- Productivity $\{\theta\}$ follows a Brownian diffusion:
$$d \log \theta_t = \hat{\mu}_t^{\log}(\theta_t) d\theta_t + \hat{\sigma}_t dW_t$$
- Stochastic control formulation:
  - Laws of motion for state variables v_t and Δ_t...
  - ...HJB equation for cost function $K(v_t, \Delta_t, \theta_t, t)$

Continuous Time: Dynamics

Proposition

1. *Dynamics*

$$\frac{d\lambda_t}{\lambda_t} = \sigma_{\lambda,t}\hat{\sigma}_t dW_t$$

$$d\gamma_t = [-\theta_t \sigma_{\lambda,t}\hat{\sigma}_t^2 + (\hat{\mu}_t + \theta_t \frac{d\hat{\mu}_t^{\log}}{d\theta})\gamma_t]dt + \gamma_t \hat{\sigma}_t dW_t,$$

2. *Allocation and wedges*

$$\frac{1}{\hat{u}^{t\prime}(c_t)} = \lambda_t \qquad \frac{1}{\hat{u}^{t\prime}(c_t)} - \frac{\theta_t}{h^{t\prime}(y_t/\theta_t)} = -\alpha\frac{\gamma_t}{\theta_t}.$$

$$\frac{\tau_{L,t}}{1-\tau_{L,t}} = -\alpha\frac{\gamma_t}{\lambda_t}\frac{1}{\theta_t} \qquad \tau_{K,t} = \sigma_{\lambda,t}^2\hat{\sigma}_t^2.$$

- Dual variables: $\lambda_t \equiv K_v(v_t, \Delta_t, \theta_t, t)$ and $\gamma_t \equiv K_\Delta(v_t, \Delta_t, \theta_t, t)$
- Sufficient control: $\sigma_\lambda(\lambda_t, \gamma_t, \theta_t, t)$

Labor Wedge: A Simple Formula

- continuous time counterpart...

$$d\left(\lambda_t \frac{\tau_{L,t}}{1-\tau_{L,t}}\right) = \left[\lambda_t \frac{\tau_{L,t}}{1-\tau_{L,t}} \theta_t \frac{d\hat{\mu}_t^{\log}}{d\theta_t} + \alpha \lambda_t \sigma_{\lambda,t} \hat{\sigma}_t^2\right] dt$$

- Drift: same discrete time
- Zero volatility: new!
  - realized paths $\Rightarrow$ bounded variation
  - innovations in $\tau_{L,t}/(1-\tau_{L,t})$ mirror λ_t
  - regressivity in short run

Labor Wedge: A Simple Formula

- continuous time counterpart...

$$d\left(\lambda_t \frac{\tau_{L,t}}{1-\tau_{L,t}}\right) = \left[\lambda_t \frac{\tau_{L,t}}{1-\tau_{L,t}} \theta_t \frac{d\hat{\mu}_t^{\text{log}}}{d\theta_t} + \alpha \lambda_t \sigma_{\lambda,t} \hat{\sigma}_t^2\right] dt$$

- Drift: same discrete time
- Zero volatility: new!
  - realized paths $\Rightarrow$ bounded variation
  - innovations in $\tau_{L,t}/(1-\tau_{L,t})$ mirror λ_t
  - regressivity in short run
- Intuiton: $\lambda_t \frac{\tau_{L,t}}{1-\tau_{L,t}} = \frac{1}{\hat{u}'(c_t)} - \frac{\theta_t}{\hat{h}'(\frac{y_t}{\theta_t})}$
  - $\frac{1}{\hat{u}'(c_t)} = \frac{\theta_t}{\hat{h}'(\frac{y_t}{\theta_t})}$ at first best
  - $\frac{1}{\hat{u}'(c_t)}$ tracks $\frac{\theta_t}{\hat{h}'(\frac{y_t}{\theta_t})}$ at second best

Labor Wedge: A Simple Formula

- Ito's Lemma....

$$d\left(\frac{\tau_{L,t}}{1-\tau_{L,t}}\right) = \left[\frac{\tau_{L,t}}{1-\tau_{L,t}}\theta_t \frac{d\hat{\mu}_t^{\log}}{d\theta} + \alpha\sigma_{\lambda,t}\hat{\sigma}_t^2\right] dt + \frac{\tau_{L,t}}{1-\tau_{L,t}}\frac{1}{\hat{u}_t'} d\left(\hat{u}_t'\right)$$

- Drift: same discrete time
- Zero volatility: new!
  - realized paths $\Rightarrow$ bounded variation
  - innovations in $\tau_{L,t}/(1-\tau_{L,t})$ mirror λ_t
  - regressivity in short run
- Intuiton: $\lambda_t \frac{\tau_{L,t}}{1-\tau_{L,t}} = \frac{1}{\hat{u}'(c_t)} - \frac{\theta_t}{\hat{h}'(\frac{y_t}{\theta_t})}$
  - $\frac{1}{\hat{u}'(c_t)} = \frac{\theta_t}{\hat{h}'(\frac{y_t}{\theta_t})}$ at first best
  - $\frac{1}{\hat{u}'(c_t)}$ tracks $\frac{\theta_t}{\hat{h}'(\frac{y_t}{\theta_t})}$ at second best

Labor Wedge: A Simple Formula

- Ito's Lemma...

$$d\left(\frac{\tau_{L,t}}{1-\tau_{L,t}}\right) = \left[\frac{\tau_{L,t}}{1-\tau_{L,t}}\theta_t \frac{d\hat{\mu}_t^{\log}}{d\theta} + \alpha\sigma_{\lambda,t}\hat{\sigma}_t^2 + \frac{\tau_{L,t}}{1-\tau_{L,t}}\sigma_{\lambda,t}^2\hat{\sigma}_t^2\right]dt - \frac{\tau_{L,t}}{1-\tau_{L,t}}\sigma_{\lambda,t}\hat{\sigma}_t dW_t$$

- Drift: same discrete time
- Zero volatility: new!
  - realized paths $\Rightarrow$ bounded variation
  - innovations in $\tau_{L,t}/(1-\tau_{L,t})$ mirror λ_t
  - regressivity in short run
- Intuiton: $\lambda_t \frac{\tau_{L,t}}{1-\tau_{L,t}} = \frac{1}{\hat{u}'(c_t)} - \frac{\theta_t}{\hat{h}'(\frac{y_t}{\theta_t})}$
  - $\frac{1}{\hat{u}'(c_t)} = \frac{\theta_t}{\hat{h}'(\frac{y_t}{\theta_t})}$ at first best
  - $\frac{1}{\hat{u}'(c_t)}$ tracks $\frac{\theta_t}{\hat{h}'(\frac{y_t}{\theta_t})}$ at second best

General Preferences

- Inverse Euler requires separability
- General utility $u^t(c, y, \theta)$
- Define:

$$\eta_t = \frac{\partial \log |\mathrm{MRS}_t|}{\partial \log \theta_t}$$

$|\mathrm{MRS}_t| = -\frac{u_y^t}{u_c^t}$

General Preferences

- Inverse Euler requires separability
- General utility $u^t(c, y, \theta)$
- Define:

$$\eta_t = \frac{\partial \log |\mathrm{MRS}_t|}{\partial \log \theta_t}$$

$$|\mathrm{MRS}_t| = -\frac{u_y^t}{u_c^t}$$

- Generalization

$$d\left(\frac{\tau_{L,t}}{1-\tau_{L,t}}\frac{1}{\eta_t}\frac{1}{u_c^t}\right) = \left[\lambda_t \sigma_{\lambda,t}\hat{\sigma}_t^2 + \frac{\tau_{L,t}}{1-\tau_{L,t}}\frac{1}{\eta_t}\frac{1}{u_c^t}\theta_t \frac{d\hat{\mu}_t^{\log}}{d\theta_t}\right] dt$$

- Interpretation:

$$\frac{\tau_{L,t}}{1-\tau_{L,t}}\frac{1}{\eta_t} = \text{discouragement}$$

General Preferences

- Interesting case

$$\tilde{u}^t\left(\hat{u}^t(c) - \frac{\kappa_t}{\alpha_t}\left(\frac{y}{\theta}\right)^{\alpha_t}\right)$$

- Then $\eta_t = \alpha_t$ deterministic and

$$d\left(\frac{\tau_{L,t}}{1-\tau_{L,t}}\right) = \left[\alpha_t \lambda_t \sigma_{\lambda,t} \hat{\sigma}_t^2 + \frac{\tau_{L,t}}{1-\tau_{L,t}} \frac{1}{u_c^t}\left(\theta_t \frac{d\hat{\mu}_t^{\log}}{d\theta_t} + \frac{1}{\alpha_t}\frac{d\alpha_t}{dt}\right)\right] dt + \frac{\tau_{L,t}}{1-\tau_{L,t}} \frac{1}{u_c^t} d\left(u_c^t\right)$$

- Life cycle pattern for elasticity?

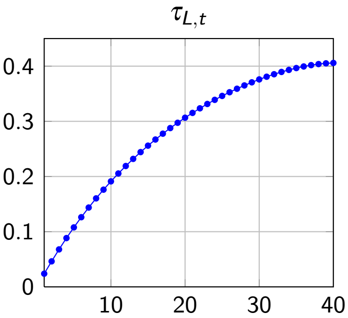
Numerical Simulation

- Agents live for $T = 60$ years, work for 40 years and retire for 20 years
- Utility function:
$$\log(c_t) - \frac{\kappa}{\alpha}\left(\frac{y_t}{\theta}\right)^{\alpha}$$
with $\alpha = 3$; $\kappa = 1$; $q = \beta = 0.95$.
- Storesletten, Telmer and Yaron (2004)
$$\theta_t = \varepsilon_t \theta_{t-1},$$
ε lognormal with $\text{Var}(\log \varepsilon) = 0.0161$.

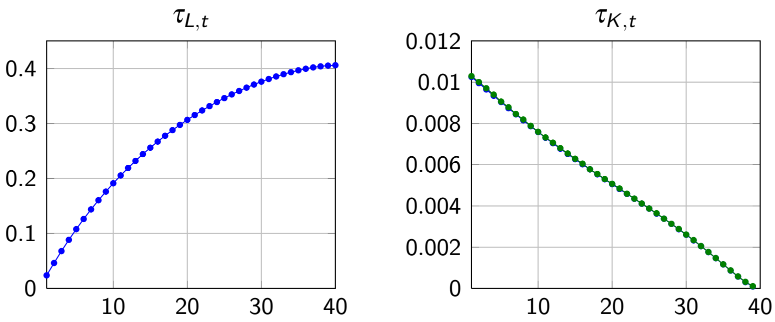
Insurance and Redistribution

- Initial heterogeneity:
  - $f^0(\theta_0)$
  - Pareto weights $\int \Lambda(\theta_0)[\mathbb{E}_0 \sum_{t=1}^{T} \beta^{t-1} u^t(c_t, y_t; \theta_t))] f^0(\theta_0) d\theta_0$ ...
  - ...or SWF $\int W(\mathbb{E}_0 \sum_{t=1}^{T} \beta^{t-1} u^t(c_t, y_t; \theta_t)) f^0(\theta_0) d\theta_0$
  - initial tax rate $\tau_{L,0}(\theta_0)$
- Focus on social insurance:
  - no initial heterogeneity
  - independent of Pareto weights or SWF
  - easy to extend

Wedges

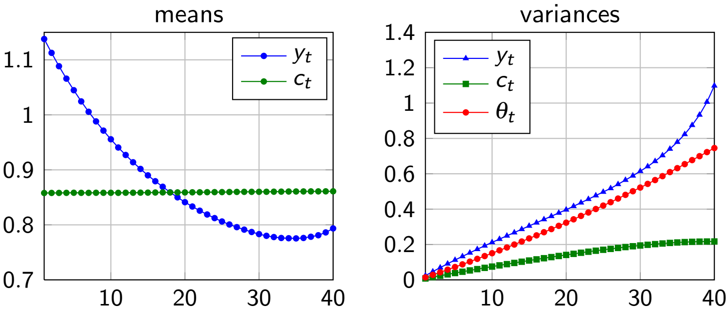

$\tau_{L,t}$
0.4
0.3
0.2
0.1
0
10
20
30
40
$\tau_{K,t}$
0.012
0.01
0.008
0.006
0.004
0.002
0
10
20
30
40

Wedges



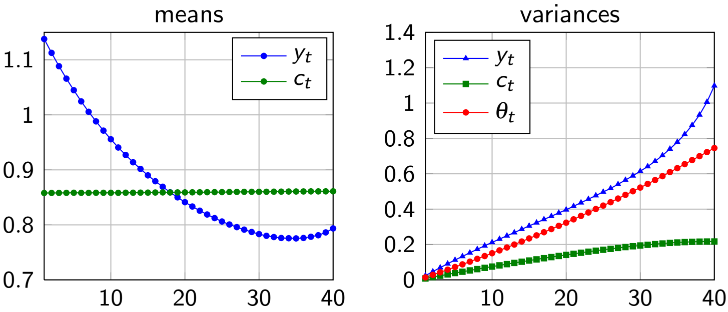
- Recall from continuous time:
  - key statistic $\{\sigma_\lambda\}$
  - with log utility $\lambda_t = c_t$
  - $\text{var}_t(c_{t+1}/c_t)$=$\sigma^2_{\lambda,t}\hat{\sigma}^2$ decreases to 0
    - as retirement nears uncertainty goes to 0
    - labor wedge increasing over time⇒increased insurance
  - explains patterns for labor and intertemporal wedges.

Allocation



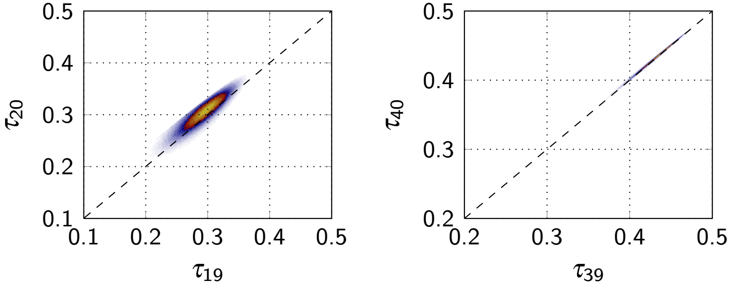
means
1.1
1
0.9
0.8
0.7
10
20
30
40
y_t
c_t
variances
1.4
1.2
1
0.8
0.6
0.4
0.2
0
10
20
30
40
y_t
c_t
θ_t

Allocation

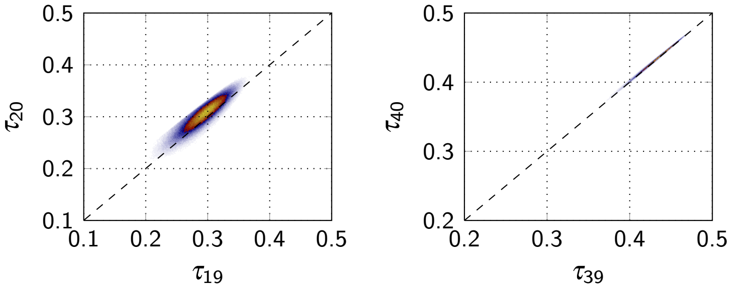


- $\mathbb{E}[c_t]$ constant: Inverse Euler
- $\mathbb{E}[y_t]$ decreasing: increasing labor wedge
- $\text{var}(y_t) > \text{var}(\theta_t)$: income and substitution effects
- $\text{var}(c_t) < \text{var}(y_t)$: insurance

Tax Smoothing and Drift

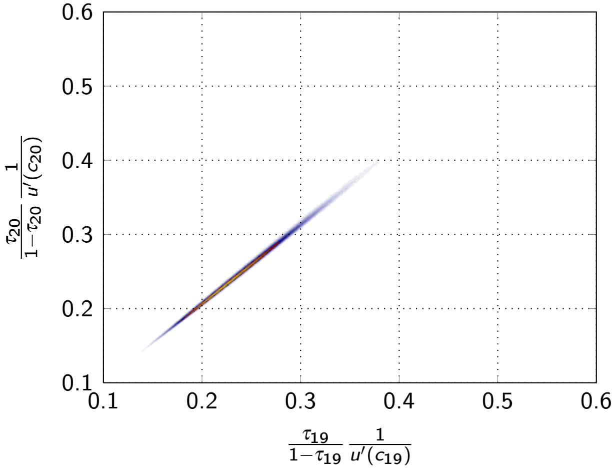


Tax Smoothing and Drift

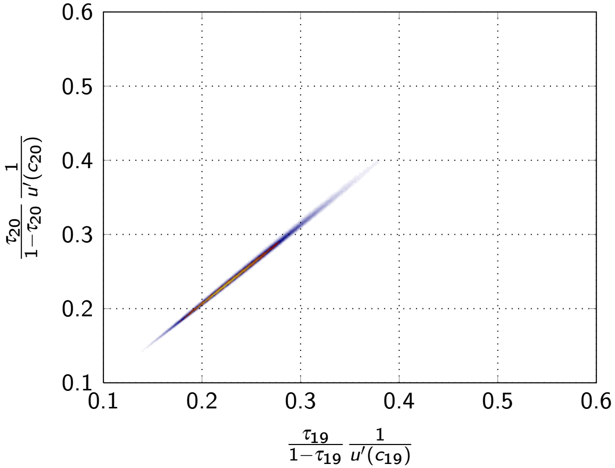


- Tax smoothing: slope close to one
- dispersion: innovations in c_t
- Drift: above 45 degree line
- Late in life:
  - lower dispersion
  - smaller drift
  - key to both: $\lim_{t \to T_E} \sigma_{\lambda,t} = 0$

Near Zero Volatility

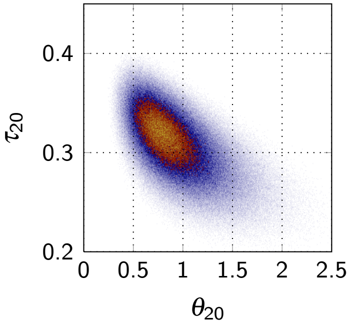


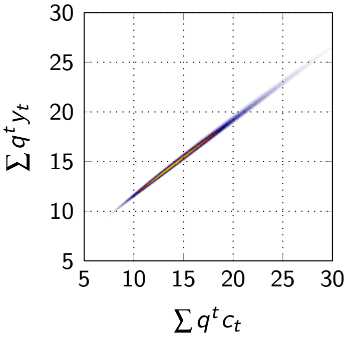
Near Zero Volatility



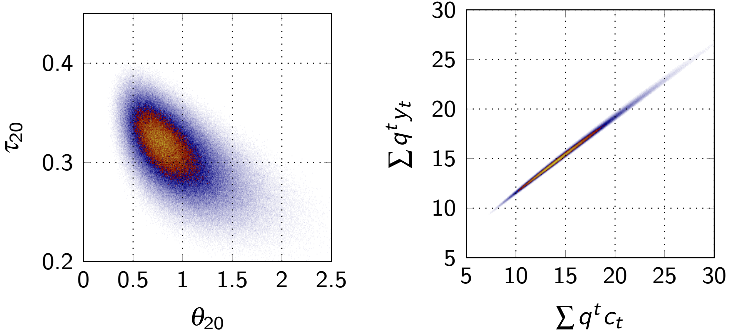
- Little dispersion
- Illustrates zero volatility result

History Dependence and Insurance





History Dependence and Insurance



- Regressive tax on average: short-term regressivity
- History dependence: dispersion
- Insurance: slope of 0.67

Impulse Response

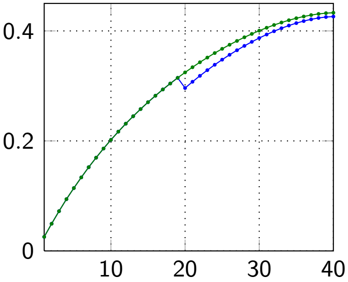
- Baseline:

$$\varepsilon_t = F(0.5) \qquad t = 1, 2, \ldots, 60$$

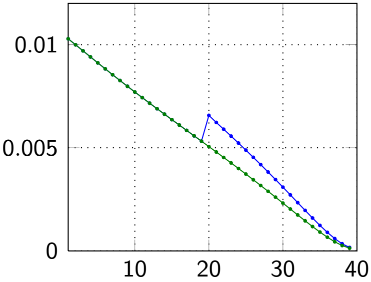
- Shock:

$$\varepsilon_{20} = F(0.95)$$

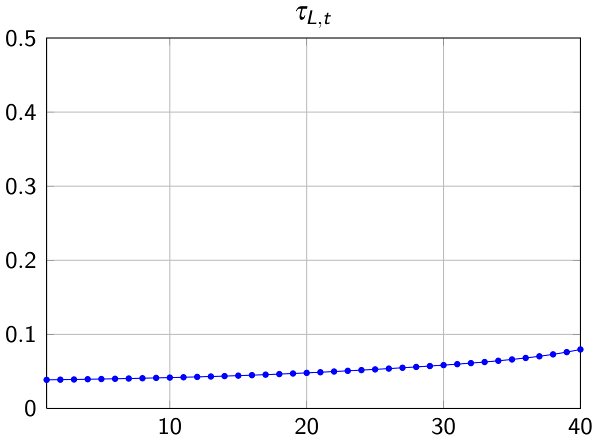
$\tau_{L,t}$



$\tau_{K,t}$



I.i.d. Case



- Normalize so that same cross sectional variance
- Level: smaller shocks in NPV
- Dynamics: easier to smooth incentives early in life

$$\mathbb{E}_{t-1}\left[\frac{\tau_{L,t}}{1-\tau_{L,t}}c_t\right] = \alpha \mathrm{Cov}_{t-1}\left(\log(\theta_t), c_t\right)$$

Welfare Gains and Simple Tax Systems

- Welfare gains relative to no taxes from...

	$\hat{\sigma}^2 = 0.0161$
second best	3.43%
first best	13.04%

Welfare Gains and Simple Tax Systems

- Welfare gains relative to no taxes from...

	$\hat{\sigma}^2 = 0.0161$
second best	3.43%
first best	13.04%

- simple policies:
  - history independent
  - age dependent
  - linear taxes = average of wedges from simulation

Welfare Gains and Simple Tax Systems

- Welfare gains relative to no taxes from...

	$\hat{\sigma}^2 = 0.0161$
second best	3.43%
first best	13.04%

- simple policies:
  - history independent
  - age dependent
  - linear taxes = average of wedges from simulation
- Simple policies capture most of the gains...

$\tau_{L,t},\ \tau_{K,t}$	3.30%
$\tau_{L,t},\ \tau_{K,t} = 0$	3.16%
$\tau_{L,t},\ \tau_{K,t} = \bar{\tau}_K$	3.29%
$\tau_{L,t} = \bar{\tau}_L,\ \tau_{K,t} = \bar{\tau}_K$	2.71%

Welfare Gains and Simple Tax Systems

- age indendent taxes $\tau_{L,t} = \bar{\tau}_L$, $\tau_{K,t} = \bar{\tau}_K$

$$\Rightarrow \bar{\tau}_K \approx 0$$

- intuition
  - mimick effects of missing age dependent taxes...
  - ... subisdy on savings $\approx$ increasing taxes on labor
  - ... encourages earlier rather than later work
    (e.g. work in earlier periods buys more goods at retirement)
  - cancels force for positive tax on savings (Inverse Euler)

Welfare Gains and Simple Tax Systems

- what is the benefit of sophisticated savings distortions?
- Exercise (Farhi-Werning, 2008)
  - take allocation from simple tax system
  - perturbation:
    - Inverse Euler holds
    - labor allocation unchanged

Welfare Gains and Simple Tax Systems

- what is the benefit of sophisticated savings distortions?
- Exercise (Farhi-Werning, 2008)
  - take allocation from simple tax system
  - perturbation:
    - Inverse Euler holds
    - labor allocation unchanged
- Welfare gains from Inverse Euler relative to...

$\tau_{L,t} = 0,\ \tau_{K,t} = 0$	0.449%
$\tau_{L,t},\ \tau_{K,t}$	0.011%
$\tau_{L,t},\ \tau_{K,t} = 0$	0.095%
$\tau_{L,t}, \tau_{K,t} = \bar{\tau}_K$	0.003%
$\tau_{L,t} = \bar{\tau}_L,\ \tau_{K,t} = \bar{\tau}_K$	0.180%

- Gains
  - overall: relatively modest
  - no taxes: higher gains (higher variance in consumption)
  - small gains from sophisticated capital taxes
  - gains from better mimicking?

Summary

- Methodology:
  - first order approach
  - discrete and continuous time
- Characterization of second best:
  - formula for the labor wedge
  - labor wedge at top and bottom
  - zero volatility result (short term regressivity)
- Second best informs us of simple policies:
  - labor taxes increasing with age
  - capital taxes decreasing with age
- Age dependent taxes important

Extensions

- Other productivity processes
- Human capital accumulation
- Extensive margin for retirement
- Occupational choice

Verification

- Solve using FOA...

$$
\begin{aligned}
c_t &= g^c(v_t, \Delta_t, r_{t-1}, r_t, t)\\
y_t &= g^y(v_t, \Delta_t, r_{t-1}, r_t, t)\\
v_{t+1} &= g^v(v_t, \Delta_t, r_{t-1}, r_t, t)\\
\Delta_{t+1} &= g^\Delta(v_t, \Delta_t, r_{t-1}, r_t, t)\\
w_t &= g^w(v_t, \Delta_t, r_{t-1}, r_t, t)
\end{aligned}
$$

- agent's problem

$$
V(v,\Delta,r_-,\theta,t) = \max_r\{u^t(g^c(v,\Delta,r_-,r,t), g^y(v,\Delta,r_-,r,t),\theta)
$$

$$
+\beta\int V(g^v(v,\Delta,r_-,r,t), g^\Delta(v,\Delta,r_-,r,t), r, \theta', t+1) f^{t+1}(\theta'|\theta)d\theta'\}
$$

- IC = verify that

$$
V(v,\Delta,r_-,\theta,t) = g^w(v,\Delta,r_-,\theta,t)
$$

General Weighting Function

- For any function $\pi(\theta)$, let $\Pi(\theta)$ be a primitive of $\pi(\theta)/\theta$
- Define

$$\phi_t^{\Pi}(\theta_{t-1}) \equiv \int \Pi(\theta_t) f^t(\theta_t|\theta_{t-1}) d\theta_t$$

Proposition

$$\mathbb{E}_{t-1}\left[\frac{\tau_{L,t}}{1-\tau_{L,t}}\frac{q}{\beta}\frac{\hat{u}^{t-1\prime}(c_{t-1})}{\hat{u}^{t\prime}(c_t)}\pi(\theta_t)\right]$$
$$= \frac{\tau_{L,t-1}}{1-\tau_{L,t-1}}\theta_{t-1}\frac{d\phi_t^{\Pi}}{d\theta_{t-1}} + \alpha Cov_{t-1}\left(\Pi(\theta_t), \frac{q}{\beta}\frac{\hat{u}^{t-1\prime}(c_{t-1})}{\hat{u}^{t\prime}(c_t)}\right)$$

- Generalizes previous formula (i.e. $\pi(\theta) = 1$)
- Characterizes process $\{\tau_L/(1-\tau_L)\}$

Formula from Global IC

- Formula:
  - local ICs $\rightarrow$ any π
  - global ICs $\rightarrow$ some π

Proposition

If $\{c, y\}$ is optimal then the labor wedge satisfies the formula above for some $\pi(\theta)$.

- e.g. $\{\theta\}$ geometric random walk $\rightarrow \pi(\theta) = \theta^{-\alpha}$