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 (MiriIees 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

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

\[ t \leq T_E \quad u^t(c, y; \theta) = \tilde{u}(c, y/\theta) \]

▶ retirement

\[ T_E < t \leq T \quad u^t(c, y; \theta) = \begin{cases} \tilde{u}(c, 0) & y = 0 \\ -\infty & y > 0 \end{cases} \]
Uncertainty and Information

- $\theta_t$ private info

- $\{\theta\}$ Markov:
  - Support: $[\theta_t(\theta_{t-1}), \overline{\theta}_t(\theta_{t-1})]$
  - Differentiable density: $f^t(\theta_t|\theta_{t-1})$

- Start with:
  - Fixed support $[\theta, \overline{\theta}]$
  - Relax later...
Planning Problem

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

s.t. \[ U(\{c,y\}) \geq \nu \]
\[ U(\{c,y\}) \geq U(\{c^\sigma, y^\sigma\}) \quad \forall \sigma \in \Sigma \]

- 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.

\[ v = \int \dot{w}(\theta) f_t(\theta | \theta_-) d\theta \quad \text{where} \quad \dot{w}(\theta) = u(c(\theta), y(\theta); \theta) + \beta v(\theta) \]

\[ \Delta = \int \dot{w}(\theta) f_{\theta_-(\theta | \theta_-)}(\dot{\theta} | \theta_-) d\theta \]

▶ Kapicka (2009), Williams (2009), Pavan, Segal and Toikka (2009)
Fernandes-Phelan

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

- Fernandes-Phelan:

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

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

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

\[
v_k = \sum_{\theta} w(\theta) f(\theta|\theta^k) \quad k = 1, 2, \ldots, N
\]

\[
w(\theta^n) = u(c(\theta^n), y(\theta^n); \theta^n) + \beta v_n(\theta^n)
\]
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. $\forall n$:

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

\[v_k = \sum_{\theta} w(\theta)f(\theta|\theta^k) \quad k = l, l+1\]

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

- Intertemporal wedge

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

- 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:
\[
\frac{1}{\hat{u}^{t-1'}(c_{t-1})} = \frac{1}{\beta R^{E_t}t-1} \left[ \frac{1}{\hat{u}'(c_t)} \right]
\]

- Intertemporal wedge
\( \tau_{K,t-1} \geq 0 \)
Labor Wedge: A Simple Formula

Assumption

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

Assumption

*AR(1) productivity*

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

Proposition

\[
\mathbb{E}_{t-1} \left[ \frac{\tau_{L,t}}{1 - \tau_{L,t}} \frac{1}{\beta R} \frac{\hat{u}^{t-1'}(c_{t-1})}{\hat{u}^{t'}(c_t)} \right] = \rho \frac{\tau_{L,t-1}}{1 - \tau_{L,t-1}} + \alpha \text{Cov}_{t-1} \left( \log(\theta_t), \frac{1}{\beta R} \frac{\hat{u}^{t-1'}(c_{t-1})}{\hat{u}^{t'}(c_t)} \right)
\]
Labor Wedge: A Simple Formula

- Labor wedge formula:

\[
\mathbb{E}_{t-1} \left[ \frac{\tau_{L,t}}{1 - \tau_{L,t}} \frac{1}{\beta R} \frac{\hat{u}^{t-1'}(c_{t-1})}{\hat{u}'(c_t)} \right] \\
= \rho \frac{\tau_{L,t-1}}{1 - \tau_{L,t-1}} + \alpha \text{Cov}_{t-1} \left( \log(\theta_t), \frac{1}{\beta R} \frac{\hat{u}^{t-1'}(c_{t-1})}{\hat{u}'(c_t)} \right)
\]
Labor Wedge: A Simple Formula

- **Labor wedge formula:**

\[
\mathbb{E}_{t-1} \left[ \frac{\tau_{L,t}}{1 - \tau_{L,t}} \frac{1}{\beta R} \frac{\hat{u}^{t-1'}(c_{t-1})}{\hat{u}'(c_t)} \right] = \rho \frac{\tau_{L,t}}{1 - \tau_{L,t}} + \alpha \text{Cov}_{t-1} \left( \log(\theta_t), \frac{1}{\beta R} \frac{\hat{u}^{t-1'}(c_{t-1})}{\hat{u}'(c_t)} \right)
\]

- **LHS: risk-adjusted conditional expectation of \( \tau_{L,t}/(1 - \tau_{L,t}) \)**
Labor Wedge: A Simple Formula

- Labor wedge formula:

\[
E_{t-1} \left[ \frac{\tau_{L,t}}{1 - \tau_{L,t}} \frac{1}{\beta R} \frac{\hat{u}^{t-1'}(c_{t-1})}{\hat{u}^{t'}(c_t)} \right]
\]

\[
= \rho \frac{\tau_{L,t-1}}{1 - \tau_{L,t-1}} + \alpha \text{Cov}_{t-1} \left( \log(\theta_t), \frac{1}{\beta R} \frac{\hat{u}^{t-1'}(c_{t-1})}{\hat{u}^{t'}(c_t)} \right)
\]

- 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:**
  \[
  \mathbb{E}_{t-1} \left[ \frac{\tau_{L,t}}{1 - \tau_{L,t}} \frac{1}{\beta R} \frac{\hat{u}^{t-1'}(c_{t-1})}{\hat{u}^{t'}(c_t)} \right] \\
  = \rho \frac{\tau_{L,t-1}}{1 - \tau_{L,t-1}} + \alpha \text{Cov}_{t-1} \left( \log(\theta_t), \frac{1}{\beta R} \frac{\hat{u}^{t-1'}(c_{t-1})}{\hat{u}^{t'}(c_t)} \right)
  \]

- **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: \( \text{Cov}_{t-1} \left( \log(\theta_t), \frac{1}{\beta R} \frac{\hat{u}^{t-1'}(c_{t-1})}{\hat{u}^{t'}(c_t)} \right) \)
  - incentive cost increases with Frisch elasticity \( 1/(\alpha - 1) \)
Labor Wedge: A Simple Formula

- If $\theta_t$ 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 $\rho$
General Stochastic Process

Define

\[
\phi_t^{\text{log}}(\theta_{t-1}) \equiv \int \log(\theta_t) f_t^{\text{t}}(\theta_t|\theta_{t-1}) d\theta_t
\]

Proposition

\[
\mathbb{E}_{t-1}\left[ \frac{\tau_{L,t}}{1-\tau_{L,t}} \frac{1}{\beta R} \frac{\hat{u}^{t-1'}(c_{t-1})}{\hat{u}^{t'}(c_t)} \right] = \\
\theta_{t-1} \frac{d \phi_t^{\text{log}}}{d \theta_{t-1}} \frac{\tau_{L,t-1}}{1-\tau_{L,t-1}} + \alpha \text{Cov}_{t-1} \left( \log(\theta_t), \frac{1}{\beta R} \frac{\hat{u}^{t-1'}(c_{t-1})}{\hat{u}^{t'}(c_t)} \right)
\]
Moving Support

Moving support:

\[\theta_t(\theta_{t-1}), \tilde{\theta}_t(\theta_{t-1})\]

Only difference

\[
\Delta(\theta^t) = \int w(\theta^{t+1}) f^{t+1}_\theta(\theta_{t+1}|\theta_t) d\theta_{t+1} \\
+ \frac{d\tilde{\theta}_t^{t+1}}{d\theta_t} w(\tilde{\theta}_{t+1}) f^{t+1}(\tilde{\theta}_{t+1}|\theta_t) \\
- \frac{d\theta_t^{t+1}}{d\theta_t} w(\theta_{t+1}) f^{t+1}(\theta_{t+1}|\theta_t)
\]
Labor Wedge at Top and Bottom

Proposition

\[
\frac{\bar{\tau}_{L,t}}{1 - \bar{\tau}_{L,t}} = \frac{\tau_{L,t-1}}{1 - \tau_{L,t-1}} \beta R \frac{\hat{u}'^t}{\hat{u}'^{t-1}} \frac{d \log \bar{\theta}_t}{d \log \theta_{t-1}} \\
\frac{\tau_L}{1 - \tau_{L,t}} = \frac{\tau_{L,t-1}}{1 - \tau_{L,t-1}} \beta R \frac{\hat{u}'^t}{\hat{u}'^{t-1}} \frac{d \log \theta_t}{d \log \theta_{t-1}}
\]

- Generalizes Mirrlees (1971):
  - fixed support...
    \[
    \tau_L (\theta^{t-1}, \bar{\theta}_t) = \tau_L (\theta^{t-1}, \theta_t) = 0
    \]
  - \( \theta_t = \varepsilon_t \theta_{t-1} \) and \( \varepsilon_t \in [\varepsilon, \bar{\varepsilon}] \)...
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 $\nu_t$ and $\Delta_t$...
  - ...HJB equation for cost function $K(\nu_t, \Delta_t, \theta_t, t)$
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

\[ \frac{1}{\hat{u}'(c_t)} = \lambda_t \quad \frac{1}{\hat{u}'(c_t)} - \frac{\theta_t}{h'(y_t/\theta_t)} = -\alpha \frac{\gamma_t}{\theta_t}. \]

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

- 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...

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

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

- continuous time counterpart...

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

- Drift: same discrete time
- Zero volatility: new!
  - realized paths \( \Rightarrow \) bounded variation
  - innovations in \( \tau_{L,t}/(1 - \tau_{L,t}) \) mirror \( \lambda_t \)
  - regressivity in short run
- Intuition: \( \lambda_t \frac{\tau_{L,t}}{1 - \tau_{L,t}} = \frac{1}{\hat{u}'(c_t)} - \frac{\theta_t}{\hat{h}'(\frac{\gamma_t}{\hat{\theta}_t})} \)
  - \( \frac{1}{\hat{u}'(c_t)} = \frac{\theta_t}{\hat{h}'(\frac{\gamma_t}{\hat{\theta}_t})} \) at first best
  - \( \frac{1}{\hat{u}'(c_t)} \) tracks \( \frac{\theta_t}{\hat{h}'(\frac{\gamma_t}{\hat{\theta}_t})} \) at second best
Ito’s Lemma:

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

Drift: same discrete time

Zero volatility: new!

- realized paths \( \Rightarrow \) bounded variation
- innovations in \( \tau_{L,t}/(1-\tau_{L,t}) \) mirror \( \lambda_t \)
- regressivity in short run

Intuition:

\[
\lambda_t \frac{\tau_{L,t}}{1-\tau_{L,t}} = \frac{1}{\hat{\nu}'(c_t)} - \frac{\theta_t}{\hat{h}'(\frac{\gamma t}{\theta_t})}
\]

- \( \frac{1}{\hat{\nu}'(c_t)} = \frac{\theta_t}{\hat{h}'(\frac{\gamma t}{\theta_t})} \) at first best
- \( \frac{1}{\hat{\nu}'(c_t)} \) tracks \( \frac{\theta_t}{\hat{h}'(\frac{\gamma t}{\theta_t})} \) at second best
Labor Wedge: A Simple Formula

- **Ito’s Lemma...**

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

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

- Inverse Euler requires separability
- General utility $u^t(c, y, \theta)$
- Define:
  \[ \eta_t = \frac{\partial \log |MRS_t|}{\partial \log \theta_t} \]
  \[ |MRS_t| = -\frac{u^t_y}{u^t_c} \]
General Preferences

- Inverse Euler requires separability
- General utility \( u^t(c, y, \theta) \)
- Define:
  \[
  \eta_t = \frac{\partial \log |MRS_t|}{\partial \log \theta_t}
  \]
  \[
  |MRS_t| = -\frac{u_y^t}{u_c^t}
  \]
- 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

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

- Then \( \eta_t = \alpha_t \) deterministic and

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

+ \[ \frac{\tau_{L,t}}{1 - \tau_{L,t}} \frac{1}{\mu_c^t} d \left( \mu_c^t \right) \]

- Life cycle pattern for elasticity?
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 $\alpha = 3; \kappa = 1; \beta = 0.95$.

Storesletten, Telmer and Yaron (2004)

$$\theta_t = \varepsilon_t \theta_{t-1},$$

$\varepsilon$ lognormal with $\text{Var}(\log \varepsilon) = 0.0161$. 
Insurance and Redistribution

- **Initial heterogeneity:**
  - $f^0(\theta_0)$
  - Pareto weights $\int \Lambda(\theta_0)\left[\mathbb{E}_0 \sum_{t=1}^{T} \beta^{t-1} u^t(c_t, y_t; \theta_t)\right]f^0(\theta_0)d\theta_0$...
  - ...or SWF $\int \mathcal{W}\left(\mathbb{E}_0 \sum_{t=1}^{T} \beta^{t-1} u^t(c_t, y_t; \theta_t)\right)f^0(\theta_0)d\theta_0$
  - initial tax rate $\tau_{L,0}(\theta_0)$

- **Focus on social insurance:**
  - no initial heterogeneity
  - independent of Pareto weights or SWF
  - easy to extend
Recall from continuous time:

- key statistic \( \{ \sigma, \lambda \} \)
- with log utility

\[ \lambda_t = c_t \]

\[ \text{var} t (c_t + 1/c_t) = \sigma_t^2 \]

\( \hat{\sigma}_t^2 \) decreases to 0 as retirement nears uncertainty goes to 0

- labor wedge increasing over time ⇒ increased insurance

explains patterns for labor and intertemporal wedges.
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 \( \Rightarrow \) increased insurance
- explains patterns for labor and intertemporal wedges.
Allocation

Means

Variances

\[ E[c_t] \] constant: Inverse Euler

\[ E[y_t] \] 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
Allocation

\[ E[c_t] \text{ constant: Inverse Euler} \]
\[ E[y_t] \text{ decreasing: increasing labor wedge} \]
\[ \text{var}(y_t) > \text{var}(\theta_t): \text{income and substitution effects} \]
\[ \text{var}(c_t) < \text{var}(y_t): \text{insurance} \]
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: $\lim_{t \to T} E \sigma \lambda_t = 0$
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

The graph illustrates the zero volatility result with the equation:

\[ \frac{\tau_{20}}{1 - \tau_{20}} \frac{1}{u'(c_{20})} \]

\[ \frac{\tau_{19}}{1 - \tau_{19}} \frac{1}{u'(c_{19})} \]
Near Zero Volatility

Little dispersion

Illustrates zero volatility result
History Dependence and Insurance

\[ \tau_{20} \]
\[ \theta_{20} \]
\[ \sum q_t y_t \]
\[ \sum q_t c_t \]

Regressive tax on average: short-term regressivity

History dependence: dispersion

Insurance: slope of 0.67
History Dependence and Insurance

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

- **Baseline:**
  \[ \varepsilon_t = F(0.5) \quad t = 1, 2, \ldots, 60 \]

- **Shock:**
  \[ \varepsilon_{20} = F(0.95) \]
l.i.d. Case

Normalize so that same cross sectional variance

Level: smaller shocks in NPV

Dynamics: easier to smooth incentives early in life

\[ \mathbb{E}_{t-1} \left[ \frac{\tau_{L,t} c_t}{1 - \tau_{L,t}} \right] = \alpha \text{Cov}_{t-1} (\log(\theta_t), c_t) \]
Welfare Gains and Simple Tax Systems

- Welfare gains relative to no taxes from...

\[ \hat{\sigma}^2 = 0.0161 \]

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>second best</td>
<td>3.43%</td>
</tr>
<tr>
<td>first best</td>
<td>13.04%</td>
</tr>
</tbody>
</table>

Simple policies capture most of the gains...
## Welfare Gains and Simple Tax Systems

- Welfare gains relative to no taxes from...\[ \hat{\sigma}^2 = 0.0161 \]

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>second best</td>
<td>3.43%</td>
</tr>
<tr>
<td>first best</td>
<td>13.04%</td>
</tr>
</tbody>
</table>

- 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...
  \[ \tau_{L,t}, \tau_{K,t} \quad 3.30\% \]
  \[ \tau_{L,t}, \tau_{K,t} = 0 \quad 3.16\% \]
  \[ \tau_{L,t}, \tau_{K,t} = \bar{\tau}_{K} \quad 3.29\% \]
  \[ \tau_{L,t} = \bar{\tau}_{L}, \tau_{K,t} = \bar{\tau}_{K} \quad 2.71\% \]
Welfare Gains and Simple Tax Systems

- age independent 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...
  - ... subsidy 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

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

Gains from better mimicking?
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...
  \[
  \tau_{L,t} = 0, \; \tau_{K,t} = 0 \quad 0.449\%
  \]
  \[
  \tau_{L,t}, \; \tau_{K,t} \quad 0.011\%
  \]
  \[
  \tau_{L,t}, \; \tau_{K,t} = 0 \quad 0.095\%
  \]
  \[
  \tau_{L,t}, \tau_{K,t} = \bar{\tau}_K \quad 0.003\%
  \]
  \[
  \tau_{L,t} = \bar{\tau}_L, \; \tau_{K,t} = \bar{\tau}_K \quad 0.180\%
  \]
- 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...

\[ c_t = g^c(v_t, \Delta_t, r_{t-1}, r_t, t) \]
\[ y_t = g^y(v_t, \Delta_t, r_{t-1}, r_t, t) \]
\[ v_{t+1} = g^v(v_t, \Delta_t, r_{t-1}, r_t, t) \]
\[ \Delta_{t+1} = g^\Delta(v_t, \Delta_t, r_{t-1}, r_t, t) \]
\[ w_t = g^w(v_t, \Delta_t, r_{t-1}, r_t, t) \]

- 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

$$\mathbb{E}_{t-1} \left[ \frac{\tau_{L,t}}{1 - \tau_{L,t}} \frac{q}{\beta} \frac{\hat{u}^{t-1'}(c_{t-1})}{\hat{u}'(c_t)} \pi(\theta_t) \right]$$

$$= \frac{\tau_{L,t-1}}{1 - \tau_{L,t-1}} \theta_{t-1} \frac{d\phi_t^\Pi}{d\theta_{t-1}} + \alpha \text{Cov}_{t-1} \left( \Pi(\theta_t), \frac{q}{\beta} \frac{\hat{u}^{t-1'}(c_{t-1})}{\hat{u}'(c_t)} \right)$$

- Generalizes previous formula (i.e. $\pi(\theta) = 1$)

- Characterizes process $\{\tau_L/(1 - \tau_L)\}$
Formula from Global IC

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

Proposition

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

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