Lecture 11: Capital Taxation and NDPF

Stefanie Stantcheva

Fall 2016

GOALS OF THIS LECTURE

(1) New model: A simpler framework for optimal capital tax theory.

(2) New Dynamic Public Finance: A toolbox.

The Need for a Simpler Model for Optimal Capital Taxation

1) Public debate centers around a **simple equity-efficiency tradeoff**:

Is the distribution of capital fair? How does capital react to taxation?

2) Econ literature: disparate models and results (individual preferences, shocks, govt objective, policy tools)

Connect 1) and 2) by deriving robust optimal capital tax formulas in terms of estimable elasticities and distributional parameters

$\Rightarrow$ optimal K tax theory looks like optimal L tax theory.

Centered around equity-efficiency trade-off.

Highlights main forces + policy implications for K tax (often obscured).

Goals and Contributions

1) Start with dynamic model with linear utility for consumption and concave utility for wealth.

$\Rightarrow$ Transitional dynamics instantaneous $\Rightarrow$ Simple, tractable theory.

Put simplicity to use: new formulas for policy-relevant cases (nonlinear tax, cross-effects, shifting, consumption tax, ..) and normative considerations.

2) Generalize to model with concave utility $\Rightarrow$ Same optimal K tax formulas apply, with *appropriately defined elasticity of the tax base.*

Qualitatively: Lessons and intuitions from simpler model still valid.

Quantitatively: Sluggish adjustments reflected in elasticity.

The faster K adjustments, the closer to simpler model.

3) Numerically explore optimal taxation using U.S. IRS data.

A Simpler Model of Capital Taxation

For exposition: Exogenous and uniform labor income z

Heterogeneous discount rate δ_i (assume $\delta_i > r$)

Exogenous and uniform rate of return r on wealth k, income: rk

Time invariant tax $T_K(rk)$

Initial wealth k_i^{init}, exogenous.

Individual i has instantaneous utility $u_i(c, k) = c + a_i(k)$

linear in consumption c and increasing and concave in wealth k.

Maximizes:

$$U_i = \delta_i \cdot \int_{t=0}^{\infty} [c_i(t) + a_i(k_i(t))] e^{-\delta_i t}$$

$$\text{s.t. } \frac{dk_i(t)}{dt} = rk_i(t) - T_K(rk_i(t)) + z_i(t) - c_i(t)$$

Solving the Individual's Maximization Problem

$$U_i = \delta_i \cdot \int_{t=0}^{\infty} [c_i(t) + a_i(k_i(t))] e^{-\delta_i t}$$

$$\text{s.t. } \frac{dk_i(t)}{dt} = rk_i(t) - T_K(rk_i(t)) + z_i(t) - c_i(t)$$

Hamiltonian: $c_i(t) + a_i(k_i(t)) + \lambda_i(t) \cdot [rk_i(t) - T_K(rk_i(t)) + z_i(t) - c_i(t)]$

$$\text{FOC in } c_i(t): \quad \lambda_i(t) = 1 \Rightarrow \text{ constant multiplier}$$

$$\text{FOC in } k_i(t): \quad a_i'(k_i(t)) + \lambda_i(t) \cdot r \cdot (1 - T_K') = -\frac{d\lambda_i(t)}{dt} + \delta_i \cdot \lambda_i(t)$$

$$\Rightarrow a_i'(k_i(t)) = \delta_i - \bar{r} \quad \text{where} \quad \bar{r} = r \cdot (1 - T_K')$$

Steady State

Utility for wealth puts limit on impatience to consume ($\delta_i > \bar{r}$)

MU for wealth $a_i'(k) = \delta_i - \bar{r}$ = value lost in delaying consumption

Wealth accumulation depends on heterogeneous preferences $a_i(\cdot), \delta_i$, and net-of-tax return $\bar{r}$ (substitution effects, no income effects)

$\Rightarrow$ Heterogeneity in (non-degenerate) steady-state wealth.

At time 0: jump from k_i^{init} to $k_i(t)$ (consumption quantum Dirac jump):

$$U_i = \underbrace{rk_i(t) - T_K(rk_i(t)) + z_i(t)}_{c_i(t)} + a_i(k_i(t)) + \delta_i \cdot (k_i^{init} - k_i(t))$$

Dynamic model equivalent to a static model:

$$U_i = c_i + a_i(k_i) + \delta_i \cdot (k_i^{init} - k_i) \quad \text{with} \quad c_i = rk_i - T_K(rk_i) + z_i$$

Announced vs. unannounced tax reforms have same effect.

Wealth in the Utility

Technical reason: to smooth otherwise degenerate steady state ($\delta_i = \delta = \bar{r}$)

Possible, but more complicated is uncertainty (in paper).

Entrepreneurship: "cost" of managing wealth, $-h_i(k)$ (return $r_i > \delta_i$).

Wealth brings non-consumption utility flows: Weber's *"spirit of capitalism."*

Keynes (1919, 1931) *"love of money as a possession"*, *"the virtue of the cake [savings] was that it was never to be consumed."*

Social status (measure of ability, performance, success)

Power and political influence.

Philanthropy and moral recognition, warm glow bequests.

Empirical evidence in favor of wealth in the utility:

Caroll (2000): helps explain top wealth holdings.

Isomorphism with Static Labor Taxation Model

$$U_i = c_i + a_i(k_i) + \delta_i \cdot (k_i^{init} - k_i) \quad \text{with} \quad c_i = rk_i - T_K(rk_i) + z_i$$

is mathematically isomorphic to static labor income model:

$$U_i = c_i - h_i(z_i) \quad \text{with} \quad c_i = z_i - T_L(z_i)$$

Optimal K tax analysis isomorphic to optimal L income tax theory.

Differences of degree rather than of kind, quantitative differences.

Key differences (e.g.: uncertainty, shocks to productivity vs. taste) reflected in estimable elasticities.

In general model, slow adjustment will be reflected in lower elasticity.

Bypasses transitional dynamics, greatly simplifies K tax analysis

Like labor supply decisions (not instantaneous, e.g. human capital investment).

Government Optimization

Government sets a time invariant budget balanced $T_K(\cdot)$ to maximize its social objective

$\int_i g_i \cdot U_i(c_i, k_i) di$ with $g_i \geq 0$ social marginal welfare weight

Optimal $T_K(\cdot)$ depends on three key ingredients:

(1) Social preferences: g_i = value of \$1 extra given to i ($\int_i g_i = 1$).

(2) Efficiency costs: Elasticitiy $e_K = (\bar{r}/k) \cdot (dk/d\bar{r})$ measures how wealth k responds to $\bar{r} = r \cdot (1 - T'_K)$

(3) Distribution of capital income: $H_K(rk)$ (for nonlinear tax).

Optimal Linear Capital Taxation at rate τ_K

$k^m(\bar{r}) \equiv \int_i k_i di$ average wealth (depends on $\bar{r}$ with elasticity e_K).

Revenues $\tau_K k^m(\bar{r})$ rebated lump-sum.

τ_K maximizes $SWF = \int_i g_i \cdot U_i(c_i, k_i) di$ with

$$U_i = \underbrace{rk_i \cdot (1-\tau_K) + \tau_K \cdot rk^m(\bar{r}) + z_i}_{c_i} + a_i(k_i) + \delta_i \cdot (k_i^{init} - k_i)$$

Standard optimal tax derivation (using envelope thm for k_i):

$$\frac{dSWF}{d\tau_K} = rk^m \cdot \underbrace{\int_i g_i \cdot \left(1 - \frac{k_i}{k^m}\right)}_{\text{Mechanical Revenue net of Welfare Effect}} - rk^m \cdot \underbrace{\frac{\tau_K}{1-\tau_K} \cdot e_K}_{\text{Behavioral Effect}}$$

Optimal τ_K such that $dSWF/d\tau_K = 0$.

Optimal Linear Capital Tax τ_K

$$\tau_K = \frac{1-\bar{g}_K}{1-\bar{g}_K + e_K} \quad \text{with} \quad \bar{g}_K = \frac{\int_i g_i \cdot k_i}{\int_i k_i} \quad \text{and} \quad e_K = \frac{\bar{r}}{k^m} \cdot \frac{dk^m}{d\bar{r}} > 0$$

Zero capital tax result: $\tau_K = 0$ only if:

$\bar{g}_K = 1$ (no inequality in rk, or no redistributive concerns $g_i \equiv 1$), or

$e_K = \infty$.

$\tau_K > 0$ as long as g_i decreasing in k_i, or wealth concentrated among low g_i agents.

$\tau_K = 1/(1+e_K)$ is revenue-maximizing in Rawlsian case: $g_i = 0$ if $k_i > 0$.

Top revenue maximizing rate: $\tau_K = 1/(1 + a_K^{top} \cdot e_K^{top})$ with a_K^{top} the Pareto tail parameter for top bracket.

Optimal Nonlinear Capital Tax

$$T'_K(rk) = \frac{1 - \bar{G}_K(rk)}{1 - \bar{G}_K(rk) + \alpha_K(rk) \cdot e_K(rk)}$$

1) $\bar{G}_K(rk) \equiv \frac{\int_{\{i:rk_i \geq rk\}} g_i d_i}{P(rk_i \geq rk) \int_i g_i di}$ is the average g_i above capital income level rk

2) $\alpha_K(rK)$ the local Pareto parameter of capital income distribution

3) $e_K(rk)$ the local elasticity of k wrt to $1 - T'_K(rk)$ at income level rk

Capital income is very concentrated (top 1% capital income earners have 60%+ of total capital income)
$\Rightarrow$ Asymptotic formula:
$T'_K(\infty) = (1 - G_K(\infty))/(1 - G_K(\infty) + \alpha_K(\infty) \cdot e_K(\infty))$ relevant for most of the tax base

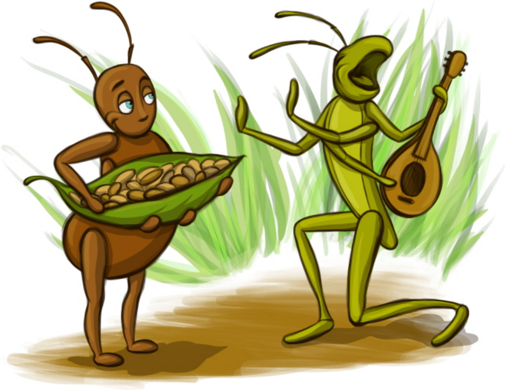
Equity Considerations: The Ant and the Grasshopper

Credit: Adelya Tumasyeva

Equity Considerations: The Ant and the Grasshopper

Credit: Adelya Tumasyeva

Equity Considerations: The Ant and the Grasshopper

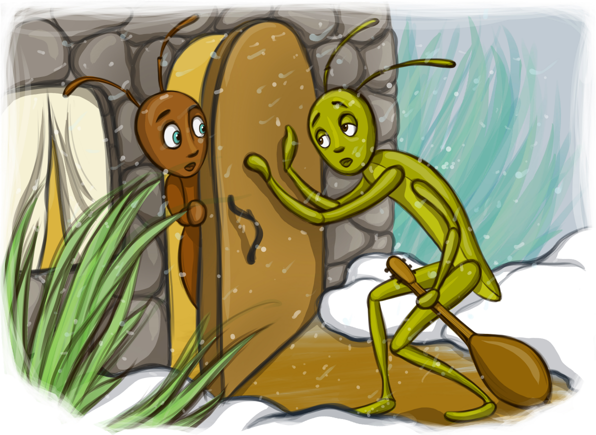


Credit: Adelya Tumasyeva

Equity Considerations: The Ant and the Grasshopper

Credit: Adelya Tumasyeva

Equity Considerations: The Ant and the Grasshopper



Credit: Adelya Tumasyeva

Equity Considerations: The Ant and the Grasshopper

Credit: Adelya Tumasyeva

Equity Considerations for Capital Taxation: Generalized Welfare Weights

(1) Inequality in wealth deemed fair and wealth is not a tag

Equality of opportunity argument: grasshopper had same savings opportunities as ant, conditional on labor earnings.

Capital accumulated by sacrificing consumption, why punish saving behavior?

What if ant had higher work (grain harvesting) ability? $\rightarrow$ role for nonlinear labor income tax.

$\rightarrow g_i$ independent of and uncorrelated with $k_i \rightarrow \tau_K = 0$.

Equity Considerations for Capital Taxation: Generalized Welfare Weights

(2) Inequality in wealth viewed as unfair

Even conditional on labor earnings, high wealth comes from higher patience δ_i or higher valuation of wealth a_i – unfair heterogeneity, like earnings ability.

or parental wealth (k_i^{init}) – ant's parents left extra grain.

or higher returns r_i (luck) – ant speculated on grain-forward derivatives.

$\rightarrow g_i$ decreasing in $k_i \rightarrow \tau_K > 0$.

Equity Considerations for Capital Taxation: Generalized Welfare Weights

(3) Wealth as a tag

May or may not care about k per se (g_i may not depend on k_i directly).

But wealth may be tag for aspects that enter g_i negatively: parental background (see Saez-Stantcheva), ability.

Having more grain means more likely to come from rich family.

$\bar{G}_K(rk)$ is representation index of agents from poor background at income rk.

$\rightarrow corr(g_i, k_i) < 0 \rightarrow \tau_K > 0$.

Adding in Labor Income Responses & Labor Taxation

Add in choice of labor income, with potentially arbitrary heterogeneity in disutility $h_i(z)$.

$$U_i = rk_i + z_i - T(rk_i + z_i) + a_i(k_i) + \delta_i \cdot (k_i^{init} - k_i) - h_i(z_i)$$

$$T_L'(z) = \frac{1 - \bar{G}_L(z)}{1 - \bar{G}_L(z) + \alpha_L(z) \cdot e_L(z)}$$

1) $\bar{G}_L(z) \equiv \frac{\int_{\{i:z_i \geq z\}} g_i d_i}{P(z_i \geq z) \int_i g_i di}$ is the average g_i above labor income level z

2) $\alpha_L(z)$ the local Pareto parameter of capital income distribution

3) $e_L(z)$ the local elasticity of k wrt to $\bar{r}$ at income level rk

Separable labor and capital taxes each set according to Mirrlees (1971) and Saez (2001) formulas.

Joint Preferences in Capital and Labor and Cross-Elasticities

Agent's dynamic problem is again equivalent to maximizing:

$$U_i = c_i + v_i(k_i, z_i) + \delta_i(k_i^{init} - k_i) \quad \text{with} \quad c_i = \bar{r}k_i + z_i - T_L(z_i)$$

Choice (c, k, z) is such that:

$$v_{iz}(k_i, z_i) = 1 - T_L'(z_i), \quad v_{ik}(k_i, z_i) = \delta_i - \bar{r}, \quad c_i = \bar{r}k_i + z_i - T_L(z_i)$$

Optimal capital tax (at any, possibly non-optimal τ_L):

$$\tau_K = \frac{1 - \bar{g}_K - \tau_L \frac{z^m}{k^m} e_{Z,(1-\tau_K)}}{1 - \bar{g}_K + e_K}$$

$$\text{with} \quad \bar{g}_K = \frac{\int_i k_i g_i}{k^m}, \quad e_{Z,(1-\tau_K)} = \frac{dz^m}{d(1-\tau_K)} \frac{(1-\tau_K)}{z^m}$$

Comprehensive nonlinear income taxation $T(rk+z)$

Govt uses solely comprehensive taxation $T(y)$ with $y_i \equiv rk_i + z_i$

$$U_i = rk_i + z_i - T(rk_i + z_i) + a_i(k_i) + \delta_i \cdot (k_i^{init} - k_i) - h_i(z_i)$$

Standard Mirrlees' formula applies to comprehensive income tax problem

$$T'(y) = \frac{1 - \bar{G}_Y(y)}{1 - \bar{G}_Y(y) + \alpha_Y(y) \cdot e_Y(y)}$$

with $\bar{G}_Y(y) \equiv \frac{\int_{\{i:y_i \geq y\}} g_i d_i}{P(y_i \geq y) \int_i g_i di}$

$\alpha_Y(y)$ local Pareto parameter for y distribution,

$e_Y(y)$ local elasticity of y with respect to $1 - T'$.

Tax shifting and Comprehensive Taxation

Suppose individual i can shift x dollars from labor income to capital income at utility cost $d_i(\underset{+}{x})$

Reported labor income z_L and capital income z_K are elastic to tax differential $\tau_L - \tau_K$

If shifting elasticity is infinite, then $\tau_L = \tau_K$ is optimal

If shifting elasticity is finite, then optimal τ_L, τ_K closer than they would be absent any shifting

If shifting elasticity is large then e_K can appear large, but wrong to set τ_K at $1/(1 + e_K)$ in that case

Heterogeneous Returns

Heterogeneous returns r_i important in practice:

Same sufficient stats formula, but replace:

$$\bar{g} = \frac{\int_i g_i \cdot r_i k_i}{\int_i r_i k_i} \quad \text{and} \quad e_K = \frac{(1-\tau_K)}{\int_i r_i k_i} \cdot \frac{d \int_i r_i k_i}{d(1-\tau_K)}$$

Values of e_K (responsiveness of k to taxes) and $\bar{g}_K$ (social judgement about capital income) could be affected.

Different Types of Capital Assets

Could have $\neq$ elasticities (housing vs. financial assets)

Different social judgments or distributional characteristics $\bar{g}_K^j$.

Formulas hold asset by asset, determined by: $\bar{g}_K^j$, e_K^j, and cross-elasticities $e_{K^s,(1-\tau_K^j)}$.

$$\tau_K^j = \frac{1 - \bar{g}_K^j}{1 - \bar{g}_K^j + e_K^j}$$

$$\bar{g}_K^j = \frac{\int_i g_i \cdot k_i^j}{\int_i k_i^j}, \quad e_K^j = \frac{\bar{r}^j}{k^{m,j}} \cdot \frac{dk^{m,j}}{d\bar{r}^j} > 0, \quad e_{K^s,(1-\tau_K^j)} = \frac{\bar{r}^j}{k^{m,s}} \cdot \frac{dk^{m,s}}{d\bar{r}^j}$$

Different Types of Capital Assets

Could have $\neq$ elasticities (housing vs. financial assets)

Different social judgments or distributional characteristics $\bar{g}_K^j$.

Formulas hold asset by asset, determined by: $\bar{g}_K^j$, e_K^j, and cross-elasticities $e_{K^s,(1-\tau_K^j)}$.

$$\tau_K^j = \frac{1 - \bar{g}_K^j - \sum_{s\neq j} \tau_K^s \frac{k^{m,s}}{k^{m,j}} e_{K^s,(1-\tau_K^j)}}{1 - \bar{g}_K^j + e_K^j}$$

$$\bar{g}_K^j = \frac{\int_i g_i \cdot k_i^j}{\int_i k_i^j}, \quad e_K^j = \frac{\bar{r}^j}{k^{m,j}} \cdot \frac{dk^{m,j}}{d\bar{r}^j} > 0, \quad e_{K^s,(1-\tau_K^j)} = \frac{\bar{r}^j}{k^{m,s}} \cdot \frac{dk^{m,s}}{d\bar{r}^j}$$

Consumption taxation: The Policy Debate

Can a consumption tax be better than a wealth tax and more progressive than a tax on labor income?

Bill Gates: *"Imagine three types of wealthy people. One guy is putting his capital into building his business. Then there's a woman who's giving most of her wealth to charity. A third person is mostly consuming, spending a lot of money on things like a yacht and plane. While it's true that the wealth of all three people is contributing to inequality, I would argue that the first two are delivering more value to society than the third. I wish Piketty had made this distinction, because it has important policy implications."*

Consumption Taxation in our Model

Consider linear consumption tax at (inclusive) tax rate τ_C so that:

$$\frac{dk_i(t)}{dt} = r(1-\tau_K)k_i(t) + z_i(t) - T_L(z_i(t)) - c_i(t)/(1-\tau_C)$$

Agents care about real wealth $k^r = k \cdot (1-\tau_C)$.

Even with wealth-in-utility, τ_C equivalent labor tax + tax on initial wealth (Kaplow, 1994, Auerbach, 2009).

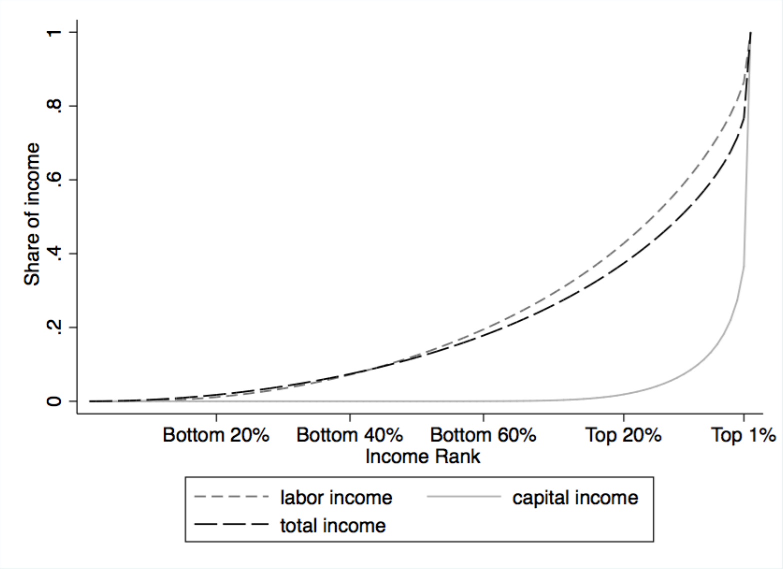
Thought experiment: equal labor income.

With τ_C, wealthy look like pay more taxes, but paid less when accumulated more nominal wealth. Real wealth inequality unaffected.

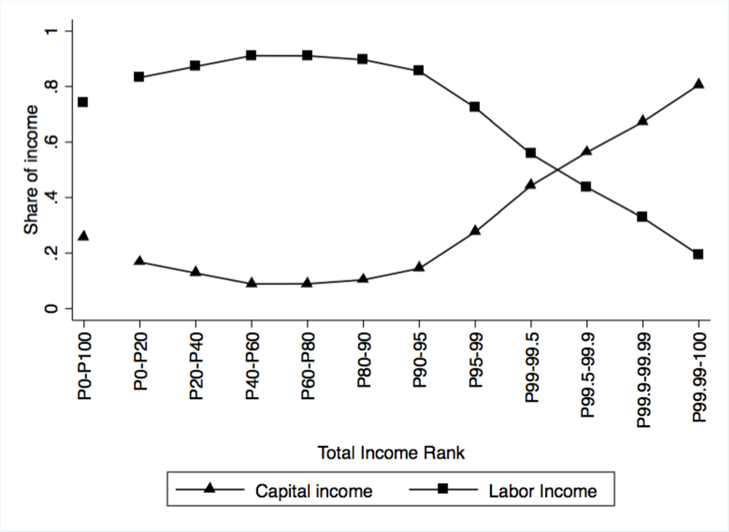
With 2-dim heterogeneity: labor tax not sufficient (Atkinson-Stiglitz).

$\Rightarrow$ τ_C cannot address steady-state capital income inequality

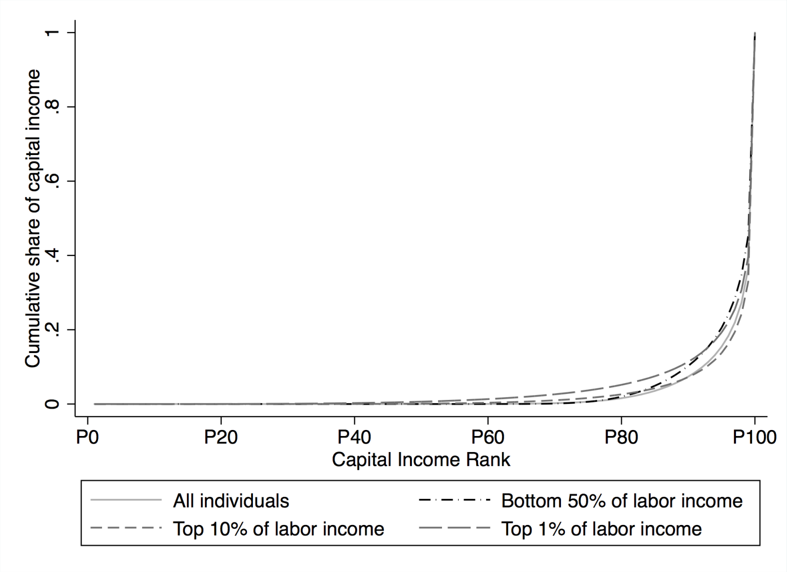
Fact 1: K income more unequally distributed than L income



Fact 2: At the top, total income is mostly capital income



Fact 3: Two-dimensional heterogeneity, inequality in K income even conditional on L income



Methodology for Computing Optimal Tax Rates

Suppose constant elasticity of labor, capital, and total income (e_L, e_K, e_Y) and that choice at zero tax represents preference type: (θ_i, η_i).

Based on the IRS micro data, use pairs (z_i, rk_i) to invert individual choices to obtain (θ_i, η_i).

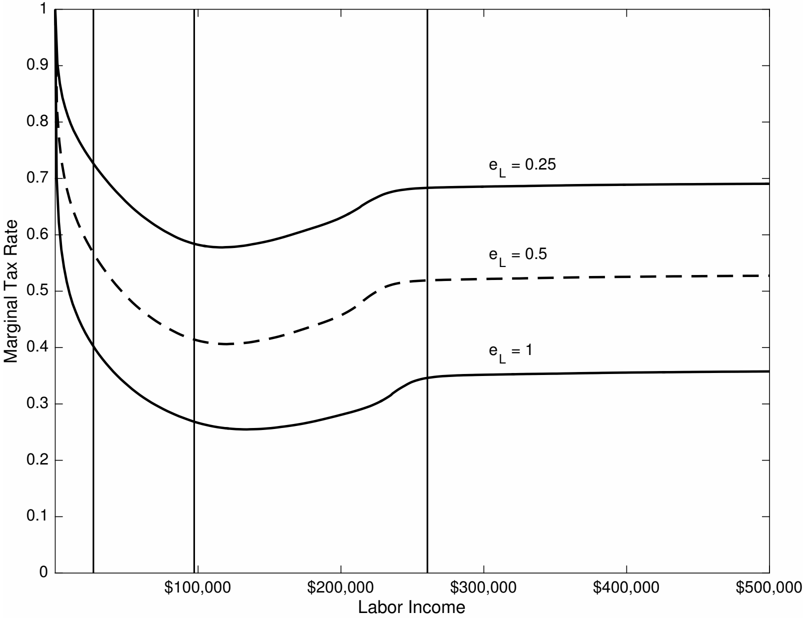
Non-parametrically fit type distributions and empirical Pareto parameters.

Solve for optimal T'_K, T'_L, and T'_Y using sufficient stats formulas.

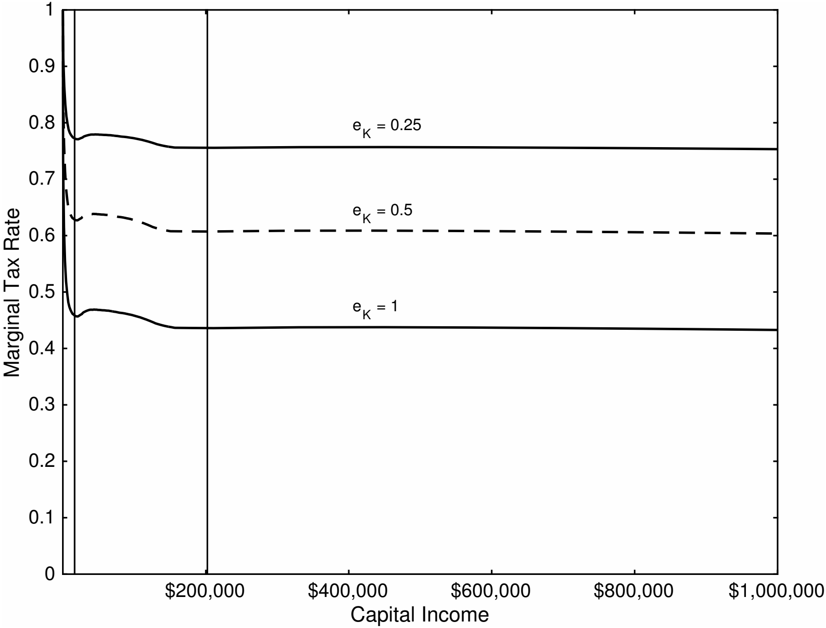
For capital – our simpler theory provides a much easier way to compute optimal tax rates based on the data.

Simulations set $g_i = \frac{1}{\text{disposable income}_i}$ and use several values for elasticities.

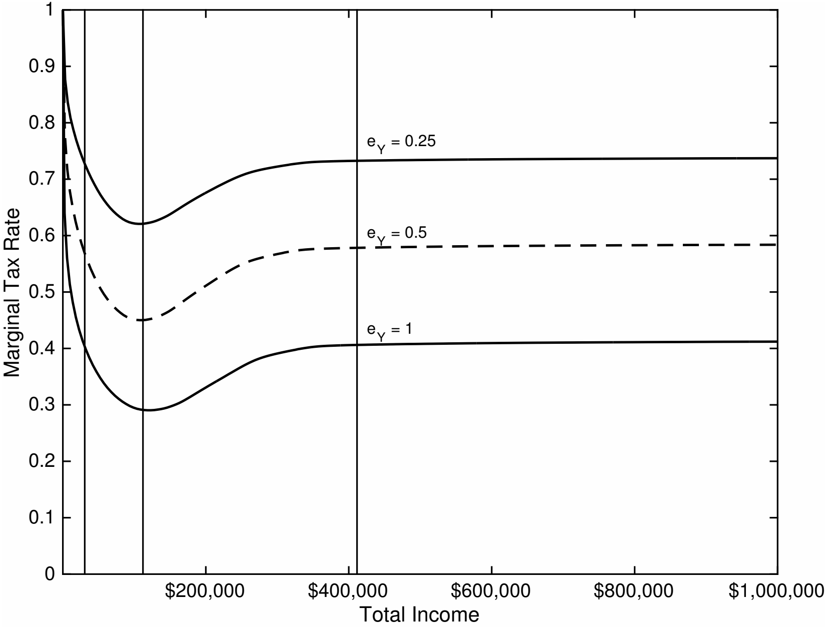
Optimal Labor Income Tax Rate $T'_L(z)$



Optimal Capital Income Tax Rate $T'_K(rk)$



Optimal Tax Rate on Comprehensive Income $T'_Y(y)$



The generalized model

Utility is

$$V_i(\{c_i(t), k_i(t), z_i(t)\}_{t \geq 0}) = \delta_i \cdot \int_{t=0}^{\infty} u_i(c_i(t), k_i(t), z_i(t)) e^{-\delta_i t} dt$$

with $u_i(\underset{+}{c}, \underset{+}{k}, \underset{-}{z})$ **concave in** $\mathbf{c}$, concave in k, concave in z

$\Rightarrow$ consumption smoothing $\Rightarrow$ sluggish transitional dynamics (a sum of anticipatory and build-up effects).

Convergence to steady state no longer instantaneous:
$u_{ik} / u_{ic} = \delta_i - \bar{r}$, $u_{ic} \cdot (1 - T'_L) = -u_{iz}$ and $c = rk + z - T(rk, z)$.

Social welfare:

$$SWF = \int_i \omega_i V_i(\{c_i(t), k_i(t), z_i(t)\}_{t \geq 0})$$

Optimal Linear Capital Tax in the Steady State

Given τ_K and τ_L, rebated lump-sum $\rightarrow$ convergence to steady state.

At time 0, start from steady state, consider unanticipated small reform $d\tau_K$, with elasticities:

$e_K(t) = dk^m(t)/d\bar{r}(\bar{r}/k^m(t)) \rightarrow e_K$.

$e_{L,(1-\tau_K)} = dz^m/d\bar{r}(\bar{r}/z^m)$.

Optimal linear capital income tax in steady state:

$$\tau_K = \frac{1 - \bar{g}_K - \tau_L \frac{z^m}{k^m} e_{L,1-\tau_K}}{1 - \bar{g}_K + \bar{e}_K}$$

If fast responses $\bar{e}_K \approx e_K$, quantitative results of simpler model hold.

Slow adjustment: $\bar{e}_K < e_K$.
But is it reasonable to exploit short-run sluggishness?

Optimal Linear Capital Tax in the Steady State

Given τ_K and τ_L, rebated lump-sum $\rightarrow$ convergence to steady state.

At time 0, start from steady state, consider unanticipated small reform $d\tau_K$, with elasticities:

$e_K(t) = dk^m(t)/d\bar{r}(\bar{r}/k^m(t)) \rightarrow e_K$.

$e_{L,(1-\tau_K)} = dz^m/d\bar{r}(\bar{r}/z^m)$.

Optimal linear capital income tax in steady state:

$$\tau_K = \frac{1 - \bar{g}_K - \tau_L \frac{z^m}{k^m} e_{L,1-\tau_K}}{1 - \bar{g}_K + \bar{e}_K}$$

If fast responses $\bar{e}_K \approx e_K$, quantitative results of simpler model hold.

Slow adjustment: $\bar{e}_K < e_K$.
But is it reasonable to exploit short-run sluggishness?

Optimal Linear Capital Tax in the Steady State

Given τ_K and τ_L, rebated lump-sum $\rightarrow$ convergence to steady state.

At time 0, start from steady state, consider unanticipated small reform $d\tau_K$, with elasticities:

$$e_K(t) = dk^m(t)/d\bar{r}(\bar{r}/k^m(t)) \rightarrow e_K.$$

$$e_{L,(1-\tau_K)} = dz^m/d\bar{r}(\bar{r}/z^m).$$

Optimal linear capital income tax in steady state:

$$\tau_K = \frac{1 - \bar{g}_K - \tau_L \frac{z^m}{k^m} e_{L,1-\tau_K}}{1 - \bar{g}_K + \bar{e}_K}$$

If fast responses $\bar{e}_K \approx e_K$, quantitative results of simpler model hold.

Slow adjustment: $\bar{e}_K < e_K$.
But is it reasonable to exploit short-run sluggishness?

Optimal Linear Capital Tax in the Steady State

Given τ_K and τ_L, rebated lump-sum $\rightarrow$ convergence to steady state.

At time 0, start from steady state, consider unanticipated small reform $d\tau_K$, with elasticities:

$$e_K(t) = dk^m(t)/d\bar{r}(\bar{r}/k^m(t)) \rightarrow e_K.$$

$$e_{L,(1-\tau_K)} = dz^m/d\bar{r}(\bar{r}/z^m).$$

Optimal linear capital income tax in steady state:

$$\tau_K = \frac{1 - \bar{g}_K - \tau_L \frac{z^m}{k^m} e_{L,1-\tau_K}}{1 - \bar{g}_K + \bar{e}_K} \quad \text{with} \quad \bar{e}_K = \int_i g_i \delta_i \int_0^\infty e_K(t) \cdot e^{-\delta_i t} dt$$

If fast responses $\bar{e}_K \approx e_K$, quantitative results of simpler model hold.

Slow adjustment: $\bar{e}_K < e_K$.
But is it reasonable to exploit short-run sluggishness?

General analysis of reforms

Comparison to standard dynamic objective:
$SWF_d = \int_i \omega_i \cdot V_i(\{c_i(t), k_i(t), z_i(t)\}_{t \geq 0})$

Any reform can be summarized by:

$$e_K^{total} = e_K^{ante} + e_K^{post}$$

Simpler model: $e_K^{total} = e_K$.

Generalized model: $\bar{e}_K = e_K^{total} = e_K^{ante} + e_K^{post}$ (if anticipated),
$\bar{e}_K = e_K^{post}$ if not anticipated.

In every model: difference between primitives vs. reform considered.

Comparison with Previous Dynamic Models

e_K steady state: Chamley-Judd model:

Infinite (degenerate) steady state elasticity $e_K = \infty$.

Aiyagari and wealth-in-utility have $e_K < \infty$.

e_K^{ante} anticipation elasticity:

If reform announced infinitely in advance, $e^{ante} = \infty$, always, with full certainty.

Reasonable?

$e^{ante} < \infty$ if uncertainty (Aiyagari).

e_K^{post} adjustment to reform: sluggish in all models, except with no transitional dynamics (linear utility).

NEW DYNAMIC PUBLIC FINANCE: REFERENCES

Dynamic taxation in the presence of future earnings uncertainty

Recent series of papers following upon on Golosov, Kocherlakota, Tsyvinski REStud '03 (GKT)

Principle can be understood in 2 period model: Diamond-Mirrlees JpubE '78 and Cremer-Gahvari EJ '95

Generalized to many periods by GKT and subsequent papers

Simple exposition is Kocherlakota AER-PP '04

Two comprehensive surveys: Golosov-Tsyvinski-Werning '06 and Kocherlakota '10 book

SET UP: UNCERTAINTY ON EARNINGS

- So far: representative agents, ex ante heterogeneity, aggregate uncertainty
- We now consider idiosyncratic uncertainty that is not only ex ante, but unfolds over time
- Skill shocks or preference shocks
- Start with finite horizon: $t = 0, 1$
- Preferences $U(c_0, c_1(s), y(s)/s)$
- interpretation: skill shock s realized in period 1. Consumption decision c_0 in period 0 is made before the shock is realized

SET UP: FAILURE OF A-S and RESOURCE CONSTRAINT

- Note difference to time-0 shock (ex ante heterogeneity) as considered so far. Preferences would be $U(c_0(s), c_1(s), y(s)/s)$
- Under separability + homogeneity, the Atkinson-Stiglitz (1976) theorem would rule out the optimality of a capital tax.
- With the period-1 shock, we will find a downward distortion of saving to be optimal (positive capital tax).
- Technology: linear storage with rate of return $R^* = 1/q$, so that the aggregate resource constraint is

$$c_0 + q \sum_s c_1(s)p(s) \leq q \sum_s y(s)p(s) \tag{1}$$

FIRST BEST

- $$\max_{c_0, c_1(s), y(s)} \sum_s U(c_0, c_1(s), y(s)/s)p(s)$$

  s.t. (1)
- FOCs for $[c_0]$
  $$\mathbb{E}[U_{c_0}, c_1(s), y(s)/s] = \lambda$$
  and for $[c_1(s)]$
  $$U_{c_1(s)}(c_0, c_1(s), y(s)/s) = \lambda q$$

FIRST BEST – FULL INSURANCE

- Hence,

$$\mathbb{E}[U_{c_0}, c_1(s), y(s)/s] = R^* U_{c_1(s)}(c_0, c_1(s), y(s)/s) \quad \forall s$$

$\Rightarrow$ Full Insurance

- Taking expectations on both sides

$$\mathbb{E}[U_{c_0}, c_1(s), y(s)/s] = R^* \mathbb{E}[U_{c_1(s)}(c_0, c_1(s), y(s)/s)] \tag{2}$$

- For instance, if $U(c_0, c_1(s), y(s)/s) = u(c_0) + \beta u(c_1(s)) - h(y(s)/s)$, then we obtain the usual Euler equation

$$u'(c_0) = \beta R^* \mathbb{E}[u'(c_1(s))]$$

and $c_1(s) = c_1$ for all s.

FREE SAVING GIVEN INCOME TAX

- Free saving with non-linear income tax $T(Y)$:

$$\max_{c_0, c_1(s), y(s)} \sum_s U(c_0, c_1(s), y(s)/s) p(s)$$

s.t.

$$c_0 + k_1 \leq e$$

$$c_1(s) \leq y(s) - T(y(s)) + Rk_1 \quad \forall s$$

- FOCs and $R = R^*$ yields the Euler equation

$$u'(c_0) = \beta R^* \mathbb{E}[u'(c_1(s))]$$

- If agents can freely decide how much to save in a risk-free asset with return $R = R^*$, we obtain the Euler equation as in the first best

PRIVATE INFORMATION AND INCENTIVE CONSTRAINTS

- Suppose s is private information and agents make reports $r = \sigma(s)$, where σ denotes the reporting strategy
- Truth-telling: $\sigma^*(s) = s \quad \forall s$
- Denote

$$c_1^\sigma(s) = c_1(\sigma(s))$$

  and

$$y^\sigma(s) = y(\sigma(s))$$

- The set of incentive constraints is

$$\mathbb{E}[U(c_0, c_1(s), y(s)/s)] \geq \mathbb{E}[U(c_0, c_1^\sigma(s), y^\sigma(s)/s] \quad \forall \sigma, s$$

- This is equivalent to

$$U(c_0, c_1(s), y(s)/s) \geq U(c_0, c_1(r), y(r)/s) \quad \forall r, s \tag{3}$$

SECOND BEST DYNAMIC MIRRLEES PROBLEM

- The second best (dynamic Mirrlees) problem is

$$\max_{c_0, c_1(s), y(s)/s} \mathbb{E}[U(c_0, c_1(s), y(s)/s)] \tag{4}$$

s.t.

$$c_0 + q \sum_s c_1(s) p(s) \leq q \sum_s y(s) p(s) \quad \text{(RC)}$$

and

$$U(c_0, c_1(s), y(s)/s) \geq U(c_0, c_1(r), y(r)/s) \quad \forall r, s \quad \text{(IC)}$$

FEASIBLE VARIATIONS

- (RC) and (IC) define the set F of *feasible* allocations, i.e.

$$F \equiv \{(c_0, c_1(s), y(s)) \mid (c_0, c_1(s), y(s)) \quad \text{satisfies (RC) and (IC)}$$

- Key question: Is free saving feasible? Formally, if $(c_0, c_1(s), y(s)) \in F$, does this imply that $(c_0 - \triangle, c_1(s) + R^*\triangle, y(s)) \in F$ as well, for some $\triangle \in \mathbb{R}$?
- In other words, if the agent saves a little in period 0 ($\triangle$) is she stil willing to supply the same output (i.e. not lie about s)?
- Depends on income effects in general
- For instance, suppose

$$U(c_0, c_1(s), y(s)/s) = \hat{U}(c_0, c_1(s) - h(y(s)/s))$$

Then, given c_0, just maximize $c_1(s) - h(y(s)/s)$. There are no income effects due to quasilinearity, and the above variation is feasible.

FEASIBLE VARIATIONS

- Easier to see using (IC):

$$\hat{U}(c_0 - \triangle, c_1(s) + R^*\triangle - h(y(s)/s)) \geq$$

$$\hat{U}(c_0 - \triangle, c_1(r) + R^*\triangle - h(y(r)/s))$$

  if and only if

$$c_1(s) + R^*\triangle - h(y(s)/s) \geq c_1(r) + R^*\triangle - h(y(r)/s)$$

  which is implied by the original allocation being feasible, i.e. $(c_0, c_1(s), y(s)) \in F$
- But in general, saving in period 0 has a negative income effect on labor supply in period 1 (if leisure is a normal good)
- e.g. consider preferences

$$U(c_0, c_1(s), y(s)/s) = u(c_0) + \beta u(c_1(s)) - h(y(s)/s)$$

  Additive separability + concavity of $u(.)$ mean leisure is normal (output is “inferior“) and so the variation above is no longer feasible.

CAN WE FIND A FEASIBLE VARIATION?

- Free saving is not feasible with these preferences due to negative income effect on labor supply
- Consider

$$(c_0 - \triangle, c_1(s) + \delta(\triangle, s), y(s)) \tag{5}$$

with $\delta(\triangle, s)$ chosen such that (IC) is satisfied:

$$u(c_0 - \triangle) + \beta u(c_1(s) + \delta(\triangle, s)) = u(c_0) + \beta u(c_1(s)) + A(\triangle) \quad \forall s, \triangle \tag{6}$$

for some $A(\triangle)$, and such that it is resource neutral:

$$-\triangle + q \sum_m \delta(\triangle, s) p(s) = 0 \quad \forall \triangle \tag{7}$$

- With the "free saving" variation, we had $\delta(\triangle, s) = -R^* \triangle$. What is key difference?

VERIFY INCENTIVE COMPATIBILITY OF VARIATION

- Verify that the variation maintains incentive compatibility:

$$u(c_0 - \triangle) + \beta u(c_1(s) + \delta(\triangle, s)) - h(y(s)/s) \geq$$

$$u(c_0 - \triangle) + \beta u(c_1(r) + \delta(\triangle, r)) - h(y(r)/s)$$

if and only if (6)

$$u(c_0) + \beta u(c_1(s)) + A(\triangle) - h(y(s)/s) \geq$$

$$u(c_0) + \beta u(c_1(r)) + A(\triangle) - h(y(r)/s)$$

if and only if

$$u(c_0) + \beta u(c_1(s)) - h(y(s)/s) \geq u(c_0) + \beta u(c_1(r)) - h(y(r)/s)$$

- Is this true?
- Key: Given separability, all that matters for incentive compatibility is the total utility from consumption.

INVERSE EULER EQUATION

- Suppose the original allocation $(c_0, c_1(s), y(s))$ solves the second best problem. Then, since the variation $\delta(\triangle, s)$ is feasible as just shown, it cannot improve the objective.
- Formally,

$$0 = argmax_{\triangle} \sum_s p(s)[u(c_0 - \triangle) + \beta u(c_1(s) + \delta(\triangle, s)) - h(y(s)/s)]$$

$$= argmax_{\triangle} \sum_s p(s)[u(c_0) + \beta u(c_1(s) + A(\triangle)) - h(y(s)/s)]$$

$$= argmax_{\triangle} A(\triangle),$$

where we used

$$u(c_0 - \triangle) + \beta u(c_1(s) + \delta(\triangle, s)) = u(c_0) + \beta u(c_1(s)) + A(\triangle) \quad \forall s, \triangle \tag{8}$$

- FOC $A'(0) = 0$

INVERSE EULER EQUATION

- $\delta(\triangle, s)$ satisfies (IC); differentiate w.r.t $\triangle$:

$$-u'(c_0) + \beta u'(c_1(s)) \frac{\partial \delta(\triangle, s)}{\partial \triangle}\Big|_{\triangle=0} = A'(0)$$

rearrange (at optimum):

$$\frac{\partial \delta(\triangle, s)}{\partial \triangle}\Big|_{\triangle=0} = \frac{u'(c_0) + A'(0)}{\beta u'(c_1(s))} = \frac{u'(c_0)}{\beta u'(c_1(s))} \quad \forall s \tag{9}$$

- Condition for resource neutrality of the variation implies:

$$-1 + q \sum_s p(s) \frac{\partial \delta(\triangle, s)}{\partial \triangle}\Big|_{\triangle=0} = 0$$

INVERSE EULER EQUATION (II)

- Using expression for $\left.\frac{\partial\delta(\triangle,s)}{\partial\triangle}\right|_{\triangle=0}$:

$$\frac{1}{u'(c_0)} = \frac{1}{\beta R^*}\sum_s \frac{p(s)}{u'(c_1(s))} \tag{10}$$

- With seperable preferences, optimal allocation has to satisfy this inverse Euler equatoin (Diamond/Mirrlees 1978, Rogerson 1985, Golosov et al. 2003)
- Is this necessary and sufficient? (Think of optimality of $y(s)$).
- Implies that the Euler equation is violated.

$$u'(c_0) = \beta R^* \sum_s p(s)u'(c_1(s)) \tag{11}$$

Is it always violated?

POSITIVE SAVINGS WEDGE

- Inverse Euler implies that, at the optimum,

$$u'(c_0) = \left[\frac{1}{\beta R^*}\sum_s \frac{p(s)}{u'(c_1(s))}\right]^{-1} = \beta R^* \left(\mathbb{E}\left[\frac{1}{u'(c_1(s)}\right]\right)^{-1}$$

  By Jensen's inequality and convexity of the function $f(x) = 1/x$,

$$u'(c_0) < \beta R^* \mathbb{E}[u'(c_1(s))]$$

- The optimal allocation is incompatible with free saving. Is saving is discouraged or encouraged?
- Intuition: saving in period 0 increases income in period 1 across all shocks $s \rightarrow$ negative income effect on $y \rightarrow$ is this good or bad for Planner?
- Implications for capital taxation, but study distinction between the wedge derived here and actual implementations later

TECHNICAL POINT: DUAL APPROACH

- Consider allocation in terms of utils

$$u_0 \equiv u(c_0), \quad u_1(s) \equiv u(c_1(s))$$

- Move from original allocation $(u_0, u_1(s))$ to variation $(\tilde{u}_0, \tilde{u}_1(s))$ such that

$$u_0 + \beta u_1(s) = \tilde{u}_0 + \beta \tilde{u}_1(s) \quad \forall s$$

- In particular, set

$$\tilde{u}_0 = u_0 - \beta \triangle$$

and

$$\tilde{u}_1(s) = u_1(s) + \triangle \quad \forall s$$

TECHNICAL POINT: DUAL APPROACH (II)

- Are incentive constraints affected?

$$\tilde{u}_0 + \beta \tilde{u}_1(s) - h(y(s)/s) \geq \tilde{u}_1(r) + \beta \tilde{u}_1(r) - h(y(r)/s) \quad \forall r, s$$

  if and only if

$$u_0 - \beta\triangle + \beta(u_1(s) + \triangle) - h(y(s)/s) \geq$$

$$u_1(r) - \beta\triangle + \beta(u_1(r) + \triangle) - h(y(r)/s) \quad \forall r, s$$

  if and only if

$$u_0 + \beta u_1(s) - h(y(s)/s) \geq u_1(r) + \beta u_1(r) - h(y(r)/s) \quad \forall r, s$$

  is this true?
- Variation by construction keeps total expected utility unchanged
- Dual problem: minimize total resource cost of allocation

$$min_{\triangle}\left\{C(u_0 - \beta\triangle) + q\sum_{s} p(s)C(u_1(s)) + \triangle)\right\}$$

  where $C(u)$ is the inverse function of $u(c)$

INVERSE EULER AGAIN

- If the original allocation $(u_0, u_1(s))$ is optimal, $\triangle = 0$ must solve this problem
- The FOC evaluated at $\triangle = 0$ is

$$-C'(u_0)\beta + q\sum_s p(s)C'(u_1(s)) = 0$$

- Use $C'(u) = 1/u'(c)$

$$\frac{1}{u'(c_0)} = \frac{q}{\beta}\sum_s \frac{p(s)}{u'(c_1(s))}$$

  which is the inverse Euler equation again (recall $q = 1/R^*$)
- Alternative interpretation: $1/u'(c)$ is resource cost of providing some given incentives
- IEE requires the equalization of the expected resource cost of providing incentives across both periods

INFINITE HORIZON

- General model with separable preferences

$$\sum_{t,s^t} \beta^t [u(c(s^t)) - h(y(s^t)/s_t)] Pr(s^t)$$

and $s^t = (s_0, s_1, ..., s_t)$

- Agents have reporting strategies such that (why does it depend on s^t not s_t?)

$$r_t = \sigma_t(s^t)$$

where the truth telling strategy is such that

$$\sigma_t^*(s^t) = s_t \quad \forall s^t, t$$

- $\sigma^t(s^t)$ denotes the history of reports induced by the strategy $\sigma_t(s^t)$, i.e.

$$\sigma^t(s^t) = (r_0, r_1, ..., r_t) = (\sigma_0(s_0), \sigma_1(s^1), ..., \sigma_t(s^t))$$

DYNAMIC INCENTIVE CONSTRAINTS

- Dynamic incentive constraints

$$\sum_{t,s^t} \beta^t [u(c(s^t)) - h(y(s^t)/s_t)/Pr(s^t)$$

$$\geq \sum_{t,s^t} \beta^t [u(c(\sigma^t(s^t))) - h(y(\sigma^t(s^t))/s_t] Pr(s^t) \quad \forall \sigma$$

- Pick some node s^t. Then set

$$\tilde{u}(s^\tau) = u(s^\tau)$$

for any $s^\tau \neq s^t$ and $s^\tau \neq (s^t, s_{t+1})$. i.e. leave consumption utilities unchanged at any node that is not s^t or any of its direct successors
- At s^t, set

$$\tilde{u}(s^t) = u(s^t) - \beta \triangle$$

and

$$\tilde{u}(s^t, s_{t+1}) = u(s^t, s_{t+1}) + \triangle \quad \forall s_{t+1}$$

DYNAMIC INCENTIVE CONSTRAINTS

- Key: if initial allocation was incentive compatible, perturbed one is as well.
- Moreover, perturbed allocation does not change total expected utility (from any reporting strategy $\sigma_t(s^t)$, thus also from truth-telling)
- Minimize expected resource cost of the perturbed allocation by choosing $\triangle$

$$min_{\triangle}\Big\{C(u(s^t)-\beta\triangle)+q\sum_{s^{t+1}|s^t}Pr(s^{t+1}|s^t)C\big(u_1(s^{t+1})+\triangle\big)\Big\}$$

- If the initial allocation is optimal, this program must be solved at $\triangle=0$ with FOC

$$\frac{1}{u'(c(s^t))}=\frac{1}{\beta R^*}\mathbb{E}\Big[\frac{1}{u'(c(s^{t+1}))}\Big|s^t\Big]$$

General inverse Euler equation has to hold for all nodes s^t

GENERAL INVERSE EULER EQUATION

- Implies

$$u'(c(s^t)) < \beta R^* \mathbb{E}\Big[u'(c(s^{t+1}))\Big| s^t\Big] \quad \forall s^t$$

i.e. savings need to be distorted downwards compared to the Euler equation from free saving

- May require individualized capital taxes that keep track of the entire history of skill shocks s^t such that

$$u'(c(s^t)) = \beta \mathbb{E}\Big[(1 + r^*(1 - \tau^k(s^{t+1})))u'(c(s^{t+1}))\Big| s^t\Big] \quad \forall s^t$$

where $r^* \equiv R^* - 1$

- However, simple linear capital tax may not work
- Farhi and Werning (2011) show how to use this framework to evaluate the welfare gains from optimal saving distortions starting from some baseline allocation, e.g. the free saving allocation (Aiyagari 1994)

IMPLEMENTATION

- STEP BACK: What is implementation? Why was it not discussed before?!
- Back to 2 period model, 2 shocks $s \in \{H, L\}$
- Suppose have found the optimal allocation with consumption $\{c_0^*, c_1^*(L), c_1^*(H)\}$
- It satisfies the inverse Euler equation

$$\frac{1}{u'(c_0^*)} = \frac{1}{\beta R^*}\left[\frac{p_L}{u'(c_1^*(L))} + \frac{p_H}{u'(c_1^*(H))}\right]$$

- Suppose we introduce a linear capital tax τ^k such that the Euler equation is satisfied

$$u'(c_0^*) = \beta R^*(1 - \tau^k)[p_H u'(c_1^*(L)) + p_H u'(c_1^*(H))] \qquad (12)$$

IMPLEMENTATION: LINEAR CAPITAL TAX, NONLINEAR INCOME TAX

- Introduce a non-linear income tax system $T_0, T_1(y)$ so that the individuals' budget constraints become

$$c_0 + k_1 \leq e_0 - T_0$$

in period 0 and

$$c_1(s) \leq y(s) - T_1(y(s)) + (1 - \tau^k)R^* k_1$$

in period 1
- Note: Very restrictive tax system where the capital tax is linear and separable from the labor income tax
- Can we find a tax system T_0, T_1, τ^k such that $\{c_0^*, c_1^*(L), c_1^*(H)\}$ is incentive compatible? [What do you think?]
- If we could force the agent to choose c_0^* and thus k_1^*? We'd be back to a standard Mirlees problem in period 1, so we can always find $T_1(y)$ that implements $c_1^*(s), y^*(s)$

PROBLEM WITH LINEAR CAPITAL TAX

- Suppose H's incentive constraint is the binding one at the optimum (which means?)

$$u(c_1^*(H)) - h(y^*(H)/H) = u(c_1^*(L)) - h(y^*(L)/H) \tag{13}$$

  i.e. *if* the agent saves optimally k_1^*, truth-telling is optimum
- Moreover, *given* truth-telling in period 1, Euler equation holds, so the agent finds it optimal to choose optimal savings k_1^*
- But: double-deviation $\sigma_1(s) = L$ for all $s \in \{H, L\}$ *and* $\tilde{k}_1 = k_1^* + \epsilon$
- If $\sigma_1(s)$ for all s and $k_1 = k_1^*$, then

$$u'(c_0^*) < \beta R^*(1 - \tau^k)u'(c_1^*(L)) \tag{14}$$

  Why?
- It is optimal to deviate to $\tilde{k}_1 = k_1^* + \epsilon$ with $\epsilon > 0$
- What is agent tempted to do here? Explain in words.

DOUBLE DEVIATION

- Profitable deviation: save more in period 0 and always claim to be low type in period 1 period

$$\tilde{U} = u(c_0^* - \epsilon) + \beta\Big[u(c_1^*(L) + R^*(1-\tau^k)\epsilon) - p_L h(y^*(L)/L) - p_H h(y^*(L)/H)\Big]$$

$$\approx \epsilon\Big[-u'(c_0^*) + \beta R^*(1-\tau^k)u'(c_1^*(L))\Big]$$

$$+u(c_0^*) + \beta\Big[u(c_1^*(L)) - p_L h(y^*(L)/L) - p_H h(y^*(L)/(H)\Big]$$

$$> u(c_0^*) + \beta\Big[p_H(u(c_1^*(H)) - h(y^*(H)/H)) + p_L(u(c_1^*(L)) - h(y^*(L)/L))\Big]$$

- Where is first approximation coming from? Where did the second equality come from?
- Hence, the double deviation makes the agent better off than truth telling under the optimal allocation that we wanted to implement.

LINEAR CAPITAL TAX DOES NOT WORK – SOLUTIONS?

- State-dependent linear capital tax $\tau^k(s)$ so that

  $$u'(c_0^*) = \beta R^*(1 - \tau^k(s))u'(c_1^*(s)) \quad \forall s$$

  state by state (Kocherlakota 2005). Prevents profitable double-deviations.

  $$\tau^k(s) = 1 - \frac{u'(c_0^*)}{\beta R^* u'(c_1^*(s))}$$

  is high whenever $c_1^*(s)$ is low.

- What does this mean? Returns to saving are made risky so as to make savings unattractive.

- However: $\tau^k(s)$ is zero in expectation so that the government does not raise revenue with the capital tax.

- In general, capital taxes must be contingent on the entire history of shocks.

LINEAR CAPITAL TAX DOES NOT WORK – SOLUTIONS? (II)

- Albanesi and Sleet (2006): joint tax function on wealth and income $T(y, k)$ with iid shocks (wealth is a sufficient statistic for history of shocks with iid shocks)
- Werning (2010): non-linear capital tax works more generally , i.e. $T^k(R^*k_t, s^t)$ rather than $(1-\tau^k(s^t))R^*k_t$. Moreover, with such a non-linear capital tax, one can make it history-independent, i.e. $T^k(R^*k_t)$, (in contrast to Kocherlakota's implementation).

REFERENCES

Altonji, J., F. Hayashi, and L. Kotlikoff "Is the Extended Family Altruistically Linked? Direct Tests Using Micro Data", American Economic Review, Vol. 82, 1992, 1177-98. (web)

Altonji, J., F. Hayashi and L. Kotlikoff "Parental Altruism and Inter Vivos Transfers: Theory and Evidence", Journal of Political Economy, Vol. 105, 1997, 1121-66. (web)

Atkinson, A.B. and A. Sandmo "Welfare Implications of the Taxation of Savings", Economic Journal, Vol. 90, 1980, 529-49. (web)

Atkinson, A.B. and J. Stiglitz "The design of tax structure: Direct versus indirect taxation", Journal of Public Economics, Vol. 6, 1976, 55-75. (web)

Atkinson, A.B. and J. Stiglitz Lectures on Public Economics, Chap 14-4 New York: McGraw Hill, 1980. (web)

Auerbach, A. and L. Kotlikoff "Evaluating Fiscal Policy with a Dynamic Simulation Model", American Economic Review, May 1987, 49?55. (web)

Aura, S. "Does the Balance of Power Within a Family Matter? The Case of the Retirement Equity Act", Journal of Public Economics, Vol. 89, 2005, 1699-1717. (web)

Banks J. and P. Diamond "The Base for Direct Taxation", IFS Working Paper, The Mirrlees Review: Reforming the Tax System for the 21st Century, Oxford University Press, 2009. (web)

Bernheim, B. D., A. Shleifer, and L. Summers "The Strategic Bequest Motive", Journal of Political Economy, Vol. 93, 1985, 1045-76. (web)

Carroll, C. "Why Do the Rich Save So Much?", NBER Working Paper No. 6549, 1998. (web)

Chamley, C. "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives", Econometrica, Vol. 54, 1986, 607-622. (web)

Christiansen, Vidar and Matti Tuomala "On taxing capital income with income shifting", International Tax and Public Finance, Vol. 15, 2008, 527-545. (web)

Cremer, H. and F. Gahvari "Uncertainty and Optimal Taxation: In defense of Commodity Taxes", Journal of Public Economics, Vol. 56, 1995, 291-310. (web)

Davies, J. and A. Shorrocks, Chapter 11 The distribution of wealth, In: Anthony B. Atkinson and Francois Bourguignon, Editor(s), Handbook of Income Distribution, Elsevier, 2000, Vol. 1, 605-675. (web)

DeLong, J.B. "Bequests: An Historical Perspective," in A. Munnell, ed., *The Role and Impact of Gifts and Estates*, Brookings Institution, 2003 (web)

Diamond, P. and J. Mirrlees "A Model of Social Insurance with Variable Retirement", Journal of Public Economics, Vol. 10, 1978, 295-336. (web)

Diamond, Peter and Emmanuel Saez "The Case for a Progressive Tax: From Basic Research to Policy Recommendations", Journal of Economic Perspectives, 25(4), Fall 2011, 165-190. (web)

Diamond, P. and J. Spinnewijn "Capital Income Taxes with Heterogeneous Discount Rates", NBER Working Paper, No. 15115, 2009. (web)

Feldstein, M. "The Welfare Cost of Capital Income Taxation", Journal of Political Economy, Vol. 86, 1978, 29-52. (web)

Finkelstein A. and J. Poterba, "Adverse Selection in Insurance Markets: Policyholder Evidence from the U.K. Annuity Market", Journal of Political Economy, Vol. 112, 2004, 183-208. (web)

Finkelstein A. and J. Poterba, "Selection Effects in the United Kingdom Individual Annuities Market", The Economic Journal, Vol. 112, 2002, 28-50. (web)

Gale, William G. and John Karl Scholz, "Intergenerational Transfers and the Accumulation of Wealth", Journal of Economic Perspectives, Vol. 8(4), 1994 145-160. (web)

Gordon, R.H. and J. Slemrod "Are "Real" Responses to Taxes Simply Income Shifting Between Corporate and Personal Tax Bases?," NBER Working Paper, No. 6576, 1998. (web)

Holtz-Eakin, D., D. Joulfaian and H.S. Rosen "The Carnegie Conjecture: Some Empirical Evidence", Quarterly Journal of Economics Vol. 108, May 1993, pp.288-307. (web)

Judd, K. "Redistributive Taxation in a Simple Perfect Foresight Model", Journal of Public Economics, Vol. 28, 1985, 59-83. (web)

Kaplow, L. "A Framework for Assessing Estate and Gift Taxation", in Gale, William G., James R. Hines Jr., and Joel Slemrod (eds.) *Rethinking estate and gift taxation* Washington, D.C.: Brookings Institution Press, 2001. (web)

Kaplow, L. "On the undesirability of commodity taxation even when income taxation is not optimal", Journal of Public Economics, Vol.90, 2006, 1235-1260. (web)

Kennickell, A. "Ponds and streams: wealth and income in the U.S., 1989 to 2007", Board of Governors of the Federal Reserve System (U.S.), Finance and Economics Discussion Series: 2009-13, 2009. (web)

King, M. "Savings and Taxation", in G. Hughes and G. Heal, eds., Public Policy and the Tax System (London: George Allen Unwin, 1980), 1-36. (web)

Kopczuk, W. "The Trick Is to Live: Is the Estate Tax Social Security for the Rich?", The Journal of Political Economy, Vol. 111, 2003, 1318-1341. (web)

Kopczuk, Wojciech "Taxation of Intergenerational Transfers and Wealth", in A. Auerbach, R. Chetty, M. Feldstein, and E. Saez (eds.), Handbook of Public Economics, Vol. 5 (Amsterdam: North-Holland, 2013). (web)

Kopczuk, Wojciech and Joseph Lupton 2007. "To Leave or Not to Leave: The Distribution of Bequest Motives," Review of Economic Studies, 74(1), 207-235. (web)

Kopczuk, Wojciech and Emmanuel Saez "Top Wealth Shares in the United States, 1916-2000: Evidence from Estate Tax Returns", National Tax Journal, 57(2), Part 2, June 2004, 445-487. (web)

Kopczuk, Wojciech and Joel Slemrod, "The Impact of the Estate Tax on the Wealth Accumulation and Avoidance Behavior of Donors", in William G. Gale, James R. Hines Jr., and Joel B. Slemrod (eds.), *Rethinking Estate and Gift Taxation,* Washington, DC: Brookings Institution Press, 2001, 299-343. (web)

Kotlikoff, L. "Intergenerational Transfers and Savings", Journal of Economic Perspectives, Vol. 2, 1988, 41-58. (web)

Kotlikoff, L. and L. Summers "The Role of Intergenerational Transfers in Aggregate Capital Accumulation", Journal of Political Economy, Vol. 89, 1981, 706-732. (web)

Kuziemko, Ilyana, Michael I. Norton, Emmanuel Saez, and Stefanie Stantcheva "How Elastic are Preferences for Redistribution? Evidence from Randomized Survey Experiments," NBER Working Paper No. 18865, 2013. (web)

Laroque, G. "Indirect Taxation is Superfluous under Separability and Taste Homogeneity: A Simple Proof", Economic Letters, Vol. 87, 2005, 141-144. (web)

Light, Audrey and Kathleen McGarry. "Why Parents Play Favorites: Explanations For Unequal Bequests," American Economic Review, 2004, v94(5,Dec), 1669-1681. (web)

Modigliani, F. "The Role of Intergenerational Transfers and Lifecyle Savings in the Accumulation of Wealth", Journal of Economic Perspectives, Vol. 2, 1988, 15-40. (web)

Norton, M. and D. Ariely "Building a Better America–One Wealth Quintile at a Time", Perspectives on Psychological Science 2011 6(9). (web)

Park, N. "Steady-state solutions of optimal tax mixes in an overlapping-generations model", Journal of Public Economics, Vol. 46, 1991, 227-246. (web)

Piketty, Thomas 2000. "Theories of Persistent Inequality and Intergenerational Mobility", In Atkinson A.B., Bourguignon F., eds. *Handbook of Income Distribution*, (North-Holland). (web)

Piketty, T. "Income Inequality in France, 1901-1998", CEPR Discussion Paper No. 2876, 2001 (appendix with optimal capital tax point). (web)

Piketty, T. "Income Inequality in France, 1901-1998", Journal of Political Economy, Vol. 111, 2003, 1004-1042. (web)

Piketty, T. "On the Long-Run Evolution of Inheritance: France 1820-2050", Quarterly Journal of Economics, 126(3), 2011, 1071-1131. (web)

Piketty, Thomas, *Capital in the 21st Century*, Cambridge: Harvard University Press, 2014, (web)

Piketty, Thomas, Gilles Postel-Vinay and Jean-Laurent Rosenthal, "Inherited vs. Self-Made Wealth: Theory and Evidence from a Rentier Society (1872-1927)," Explorations in Economic History, 2014. (web)

Piketty, T. and E. Saez "Income Inequality in the United States, 1913-1998", Quarterly Journal of Economics, Vol. 118, 2003, 1-39. (web)

Piketty, T. and E. Saez "A Theory of Optimal Capital Taxation", NBER Working Paper No. 17989, 2012. (web)

Piketty, T. and E. Saez "A Theory of Optimal Inheritance Taxation", Econometrica, 81(5), 2013, 1851-1886. (web)

Piketty, T. and G. Zucman "Capital is Back: Wealth-Income Ratios in Rich Countries, 1700-2010", Quarterly Journal of Economics, 2014 (web)

Piketty, T. and G. Zucman "Wealth and Inheritance in the Long-Run", Handbook of Income Distribution, Volume 2, Elsevier-North Holland, 2014 (web)

Pirttila, Jukka and Hakan Selin, "Income shifting within a dual income tax system: evidence from the Finnish tax reform," Scandinavian Journal of Economics, 113(1), 120-144, 2011. (web)

Saez, E. "Optimal Capital Income Taxes in the Infinite Horizon Model", Journal of Public Economics, 97(1), 2013, 61-74. (web)

Saez, E. "The Desirability of Commodity Taxation under Nonlinear Income Taxation and Heterogeneous Tastes", Journal of Public Economics, Vol. 83, 2002, 217-230. (web)

Saez, E. and S. Stantcheva "A Simpler Theory of Optimal Capital Taxation", NBER Working Paper 22664, 2016. (web)

Saez, E. and G. Zucman "Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data', Quarterly Journal of Economics, 2016. (web)

Sandmo, A. "The Effects of Taxation on Savings and Risk-Taking", in A. Auerbach and M. Feldstein, Handbook of Public Economics vol. 1, chapter 5, 1985, 265-311. (web)

Scholz, J. “Wealth Inequality and the Wealth of Cohorts”, University of Wisconsin, mimeo,2003. (web)

Slemrod, J. and J. Bakija, Taxing Ourselves: A Citizen’s Guide to the Debate over Taxes. The MIT Press, 2004.

Wilhelm, Mark O. “Bequest Behavior and the Effect of Heirs’ Earnings: Testing the Altruistic Model of Bequests,” American Economic Review, 86(4), 1996, 874-892. (web)

Zucman, G. “The Missing Wealth of Nations: Are Europe and the US Net Debtors or Net Creditors”, Quarterly Journal of Economics, 2013, 1321-1364. (web)

Zucman, Gabriel, “Taxing Offshore Wealth and Profits,” Journal of Economic Perspectives 28(3), 2014.