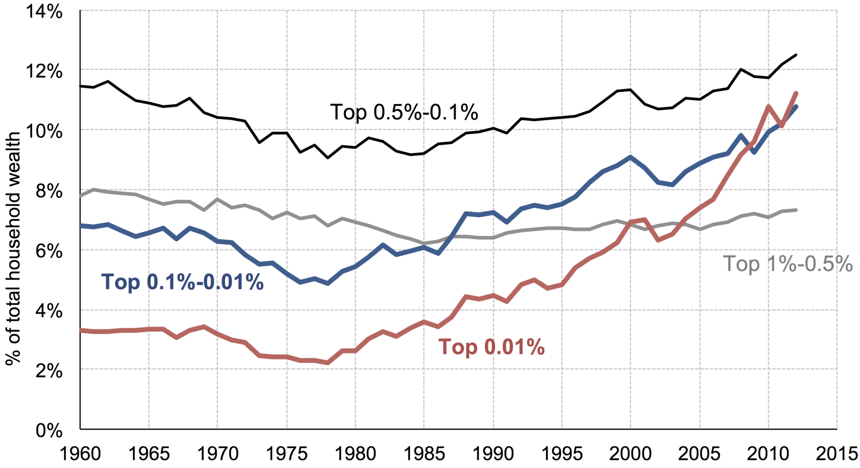
A Simpler Theory of Optimal Capital Taxation

Emmanuel Saez and Stefanie Stantcheva



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.

Related Literature

Key Results on K Taxation :

Neoclassical: Judd (1985), Chamley (1986), Straub and Werning (2015).

Incomplete markets: Aiyagari (1995), Farhi (2010).

NDPF/ Inverse Euler Equation: Kocherlakota (2005), Golosov, Tsyvinski, Werning (2006), Golosov and Tsyvinski (2006), Farhi and Werning (2012), Abraham, Koehne, Pavoni (2014), Scheuer and Wolitzky (2016).

Quantitative Models: Conesa, Kitao and Krueger (2008), Chen, Guvenen, Kambourov, Kuruscu, Ocampo (2016).

Estate Taxation: Farhi and Werning (2008), Piketty and Saez (2013), De Nardi and Yang (2015).

Isomorphism to labor taxation: Farhi and Werning (2013) applied to Estate taxation.

Sufficient Stats in Dynamic Models: Golosov, Tsyvinski, Werquin (2014), Badel and Huggett (2016).

Outline

A Simpler Model of Capital Taxation

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

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

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

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

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

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

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

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

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

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):

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

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

Putting the Model to Use: Topics

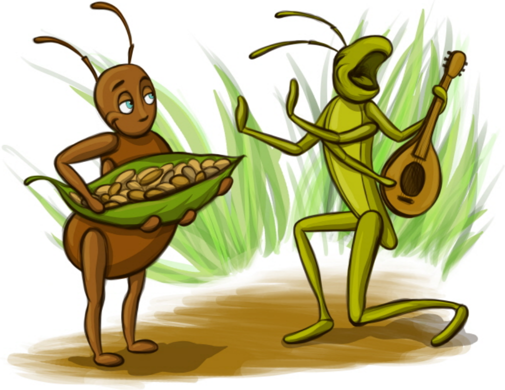
Equity Considerations: The Ant and the Grasshopper

Credit: Adelya Tumasyeva

Equity Considerations: The Ant and the Grasshopper

Credit: Adelya Tumasyeva

Equity Considerations: The Ant and the Grasshopper

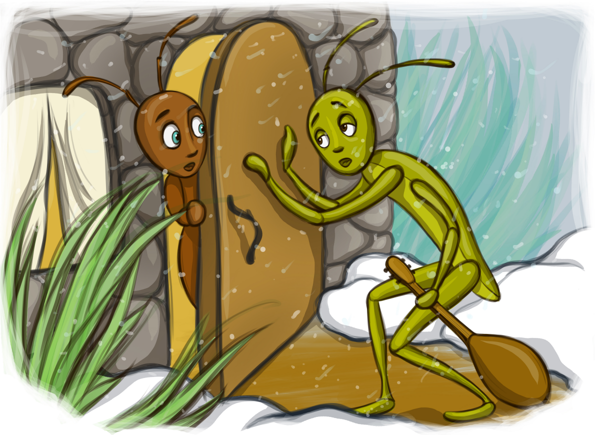


Credit: Adelya Tumasyeva

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.

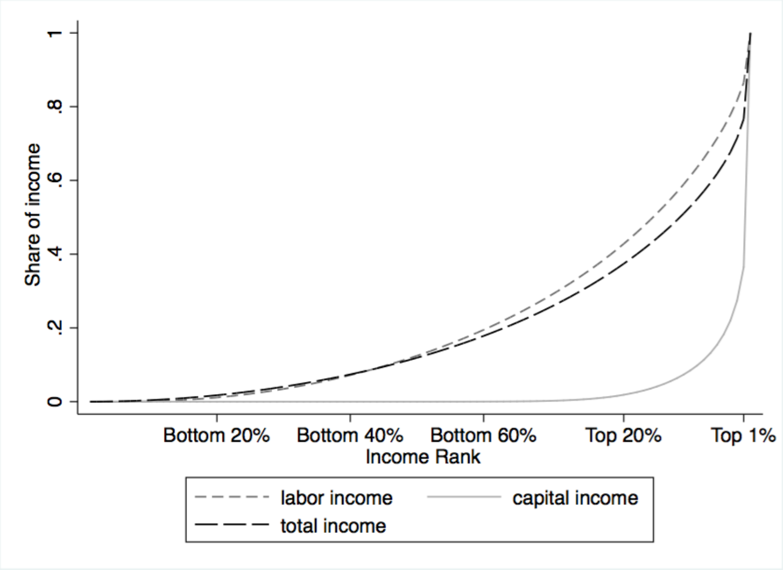
With τ_C, wealthy look like pay more taxes, but paid less when accumulated more nominal wealth. Real wealth inequality unaffected.

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

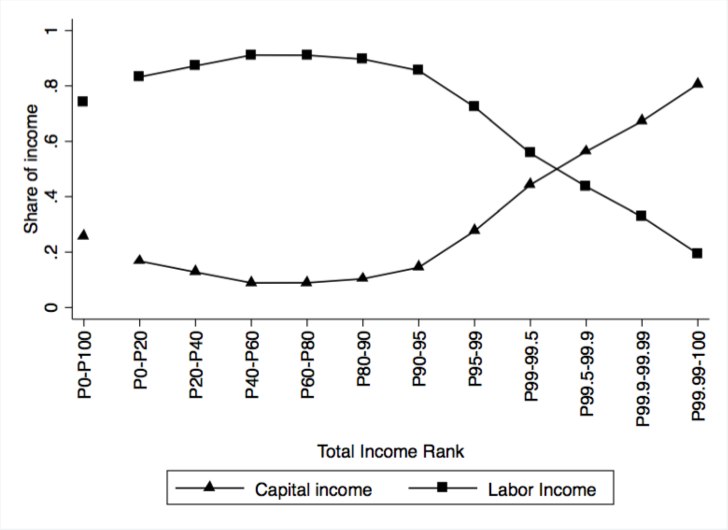
$\Rightarrow$ τ_C cannot address steady-state capital income inequality

Numerical Application to the U.S.

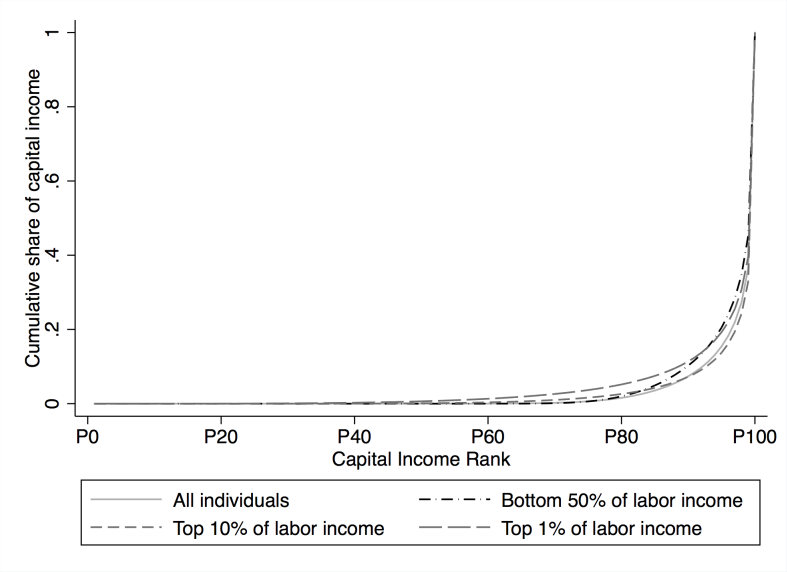
Fact 1: K income more unequally distributed than L income



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).

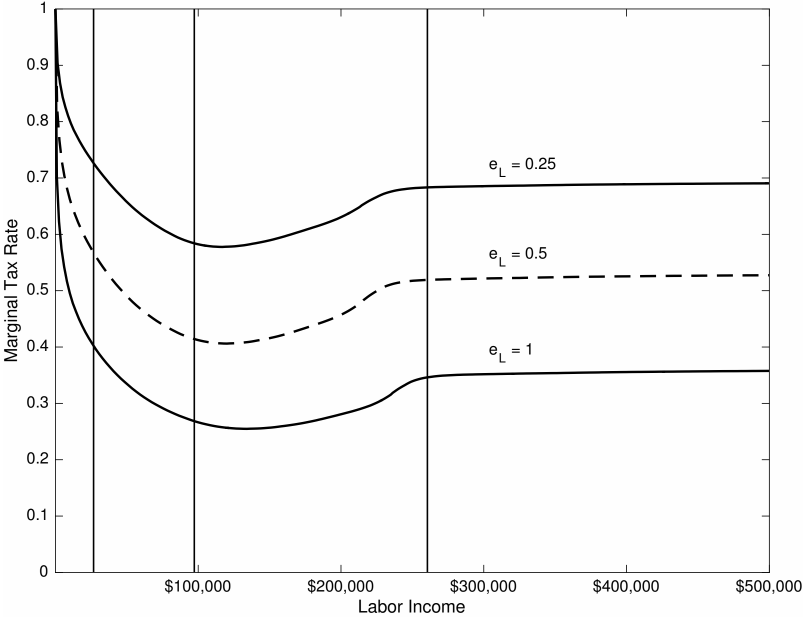
Non-parametrically fit type distributions and empirical Pareto parameters.

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

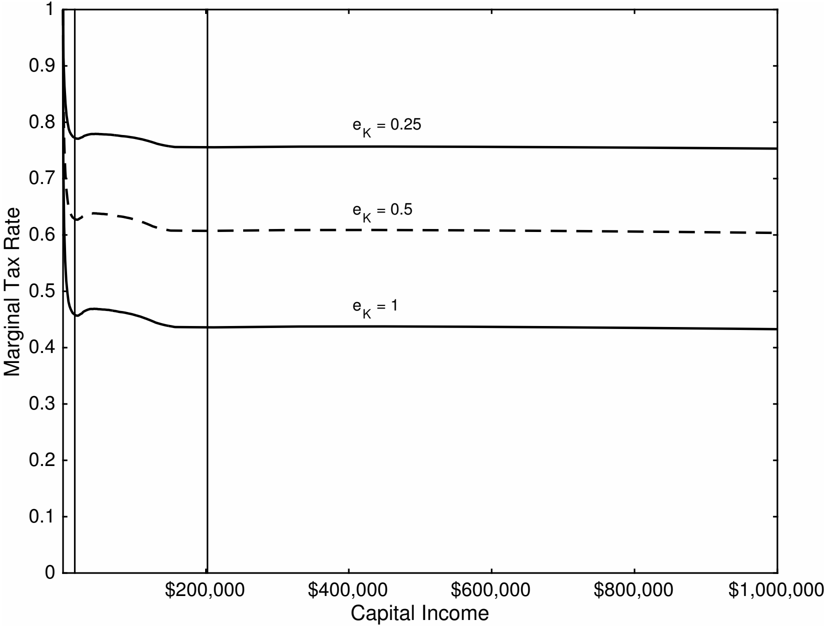
For capital – our simpler theory provides a much easier way to compute optimal tax rates based on the data.

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

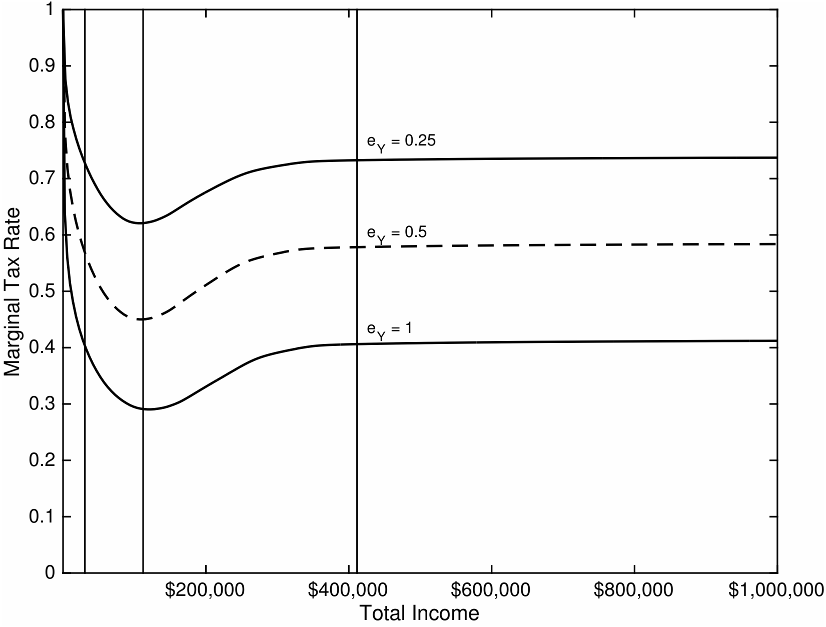
Optimal Labor Income Tax Rate $T_L'(z)$



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



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



Generalized Model

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).

Conclusion

Tractable model for K taxation centered on efficiency-equity tradeoff.

Step 1: Linear utility model with wealth in the utility.

Simplicity allows us to consider various policy relevant issues: shifting, consumption taxation, cross-elasticities, ...

Step 2: Extend results to general model.

Qualitative intuitions and results still apply if define elasticity $\bar{e}_K$ properly.

Quantitative difference: sluggish adjustments, reflected in elasticity.

Sufficient stats map easily to the data to simulate optimal tax rates.

Asymptotic optimal capital tax rate relevant for most of capital distribution, given that capital highly concentrated.

Appendix

Exogenous Economic Growth at rate g

Same theory replacing δ_i by $\delta_i - g$ and r by $r - g$.

Proof: $a_{it}(k(t)) = e^{gt} \cdot a_i(\tilde{k}(t))$ (needed for BGP), with $\tilde{c}(t) = c(t) \cdot e^{-g \cdot t}$, $\tilde{k}(t) = k(t) \cdot e^{-g \cdot t}$, then maximization equivalent to: $U_i = \int_{t=0}^{\infty} [\tilde{c}(t) + a_i(\tilde{k}(t))] \cdot e^{-[\delta_i - g] \cdot t}$
$\dot{\tilde{k}}(t) = (\bar{r} - g)\tilde{k}(t) + r \cdot \tau_K \cdot \tilde{k}^m - \tilde{c}_i(t)$

With growth, maintaining wealth per capita requires higher savings, but those are less costly: acts as if discount rate were reduced from δ_i to $\delta_i - g$.

If $\bar{r} < g$, wealth lovers hold more wealth, but have lower consumption.

If we care about consumption: $\Rightarrow \bar{\tau}_K = 1 - g/r$ (i.e., $\bar{r} = g$) may be natural **upper** bound for τ_K.

Link with Chamley-Judd model of perfect certainty

Same formulas apply.

Chamley-Judd zero tax result relies critically on anticipation effects to long-distance capital tax reforms (Piketty and Saez, 2013)
$\rightarrow e_K^{ante} = \infty$ for long distance reforms.

> Without uncertainty, wealth in the utility model, or endogenous discount rate (Judd, 1985) also have infinite anticipation elasticities to reforms far in the future.

Steady state elasticity $e_K = \infty$ in contrast to our model (impossible to incorporate heterogeneous discount rates).

$\Rightarrow$ Not very useful for policy recommendations (absent compelling empirical evidence that anticipatory effects are large) given that reforms not announced far in advance in practice.

In paper: Formulas also robust to Judd (1985) endogenous discount rate $\delta_i(c_i)$ model.