Capital Taxation

Stefanie Stantcheva

Fall 2019
MOTIVATION

1) Capital income is about 25% of national income (labor income is 75%) but distribution of capital income is much more unequal than labor income.

Capital income inequality is due to differences in savings behavior but also inheritances received.

⇒ Equity suggests it should be taxed more than labor.

2) Capital Accumulation correlated strongly with growth [although causality link is not obvious] and capital accumulation might be sensitive to the net-of-tax return.

⇒ Efficiency cost of capital taxation might be high.
3) Capital more mobile internationally than labor

Key distinction is **residence** vs. **source** base capital taxation:

**Residence:** Capital income tax based on residence of owner of capital.

Most individual income tax systems are residence based (with credits for taxes paid abroad)

Incidence falls on owner ⇒ can only escape tax through evasion (tax heavens) or changing residence (mobility of persons)

Tax evasion of capital income through tax heavens is a very serious concern (Zucman QJE’13, ’15)
Source: Capital income tax based on location of capital (most corporate income tax systems are source based)

Incidence is then partly shifted to labor if capital is mobile.

Example: Open economy with fully mobile capital and source taxation:
Local GDP: \( wL + rK = F(K, L) = L \cdot F(K/L, 1) = L \cdot f(k) \) where \( k = K/L \) is capital stock per worker

Net-of-tax rate of return is fixed by the international rate of return \( r^* \) so that \((1 - \tau_c)F_K(K, L) = (1 - \tau_c)f'(k) = r^* \) where \( k = K/L \) is capital stock per worker and \( \tau_c \) corp tax rate

As \( wL + r^*K = F(K, L) \), wage \( w = F_L(K, L) = f(k) - r^* \cdot k \) falls with \( \tau_c \)

4) Capital taxation is extremely complex and provides many tax avoidance opportunities
SAVING FLOWS

Saving is a flow and wealth or net worth is a stock

Three saving flows:

1) **Personal saving**: individual income less individual consumption [fell dramatically in the US since 1980s, recent ↑ since 2008]

2) **Corporate Saving**: retained earnings = after tax profits - distributions to shareholders

3) **Government Saving**: Taxes - Expenditures [federal, state and local]

Taxes on savings might affect different savings flows differently: savings subsidy through a tax credit can ↑ individual savings but ↓ govt saving [if govt spending stays constant]

Analyzes income, wealth, inheritance data over the long-run:

1) Growth rate $n + g = \text{population growth} + \text{growth per capita}$. Population growth will converge to zero, growth per capita for frontier economies is modest (1%) $\Rightarrow$ long-run $g \approx 1\%,\ n \approx 0\%$

2) Long-run steady-state Wealth to income ratio ($\beta$) = savings rate ($s$) / annual growth ($n + g$): $\beta = s / (n + g)$

Low growth $\Rightarrow$ high wealth-to-income ratio.

Proof: $K_{t+1} = (1 + n + g) \cdot K_t = K_t + s \cdot Y_t \Rightarrow K_t / Y_t = s / (n + g)$

With $s = 8\%$ and $n + g = 2\%$, $\beta = 400\%$ but with $s = 8\%$ and $n + g = 1\%$, $\beta = 800\%$ $\Rightarrow$ Wealth will become important
3) After-tax rate of return on wealth \( \bar{r} = r(1 - \tau_K) = 4 - 5\% \) significantly larger than \( n + g \) [except exceptional period of 1930–1970]

With \( \bar{r} > n + g \), role of inheritance in wealth and wealth concentration become large [past swallows the future]

Explanation: Rentier who saves all his return on wealth accumulates wealth at rate \( \bar{r} \) bigger than \( n + g \) and hence his wealth grows relative to the size of the economy. The bigger \( \bar{r} - (n + g) \), the easier it is for wealth to “snowball”

\[ \Rightarrow \text{Capital taxation reduces } r \text{ to } \bar{r} = r \cdot (1 - \tau_K) \Rightarrow \text{This can reduce wealth concentration} \]
The rate of return to capital (after tax and capital losses) fell below the growth rate during the 20th century, and may again surpass it in the 21st century. Sources and series: see piketty.pse.ens.fr/capital21c

Source: Piketty (2014)
WEALTH AND CAPITAL INCOME IN AGGREGATE

Definition: Capital Income = Returns from Wealth Holdings

Aggregate US Personal Wealth $\approx 4 \times \text{GDP} \approx 60 \text{ Tr}$

Tangible assets: residential real estate (land+buildings) [income = rents] and unincorporated business + farm assets [income = profits]

Financial assets: corporate stock [income = dividends + retained earnings], fixed claim assets (corporate and govt bonds, bank accounts) [income = interest]

Liabilities: Mortgage debt, Student loans, Consumer credit debt

Substantial amount of financial wealth is held indirectly through: pension funds [DB+DC], mutual funds, insurance reserves
III.A. The Distribution of Taxable Capital Income

The starting point of our allocation is the capital income reported on individual tax returns. For the post-1962 period, we rely on the yearly public-use micro-files available at the NBER that provide information for a large sample of taxpayers, with detailed income categories. We supplement this dataset using the internal use Statistics of Income (SOI) individual tax return sample files from 1979 onward. For the pre-1962 period, no...

---

**Figure II**

Aggregate US Household Wealth, 1913–2013
The composition of capital income in the U.S., 1913-2013

Source: Saez and Zucman (2014)
INDIVIDUAL WEALTH AND CAPITAL INCOME

Differences in Wealth and Capital income due to:

1) Age

2) past earnings, and past saving behavior $E_t - C_t$ [life cycle wealth]

3) Net Inheritances received $l_t$ [transfer wealth]

4) Rates of return $r_t$

[details in Davies-Shorrocks ’00, Handbook chapter]
WEALTH DISTRIBUTION

Wealth inequality is very large (much larger than labor income)

US Household Wealth is divided 1/3,1/3,1/3 for the top 1%, the next 9%, and the bottom 90% [bottom 1/2 households hold almost no wealth]

Financial wealth is more unequally distributed than (net) real estate wealth

Share of real estate wealth falls at the top of the wealth distribution

Growth of private pensions [such as 401(k) plans] has “democratized” stock ownership in the US
WEALTH MEASUREMENT

In the US, wealth distribution much less well measured than income distribution because no systematic administrative source (no wealth tax). 3 methods to estimate wealth distribution:

1) **Surveys:** US Survey of Consumer Finances (SCF)

Top 10% wealth share has grown from 67% in 1989 to 75% in 2010

Top 1% wealth share has grown “only” from 30% in 1989 to 35% in 2010 [Kennickell ’09, ’12]

Problems: small sample size, measurement error, only every 3 years, starts in 1989
2) **Estate multiplier method:** use annual estate tax statistics and re-weights individual estates by inverse of death probability [based on age $\times$ gender $\times$ social class]

Kopczuk-Saez NTJ’04 create series 1916-2000 and find fairly small increases in wealth concentration in recent decades

Problems: social class effect on mortality not well known, significant estate tax avoidance, noisy measure of “young wealth”, estates cover only the super rich (top .1% in recent years)

3) **Capitalization method:** use capital income from individuals tax statistics and estimates rates of returns by asset class to infer wealth: shows big increase in wealth concentration [Saez-Zucman ’16]
CAPITAL TAXATION IN THE US


1) **Corporate Income Tax** (fed+state): 21% Federal tax rate on profits of corporations [complex rules with many industry specific provisions]: effective tax rate much lower and incidence depends on mobility of capital

2) **Individual Income Tax** (fed+state): taxes many forms of capital income

   - Realized capital gains and dividends (dividends since ’03 only) receive preferential treatment
   - Imputed rent of home owners, returns on pension funds, state+local government bonds interest are exempt
FACTS OF US CAPITAL INCOME TAXATION

3) Estate and gift taxes:
Fed taxes estates above $11.2M exemption for singles and $22.4M for married, tax rate is 40% above exemption.
Charitable and spousal giving is exempt
Substantial tax avoidance activity through tax accountants
Step-up of realized capital gains at death (lock-in effect)

4) Property taxes (local) on real estate (old tax):
Tax varies across jurisdictions. About 0.5% of market value on average, like a 10% tax on imputed rent if return is 5%
Lock-in effect in states that use purchase price base such as California
Heterogeneous individuals and government uses nonlinear tax on earnings. Should the govt also use tax on savings?

\[ V^h = \max U^h(v(c_1, c_2), l) \text{ st } c_1 + c_2 / (1 + r(1 - \tau_K)) = wl - T_L(wl) \]

If utility is weakly separable and \( v(c_1, c_2) \) is the same for all individuals, then the government should use only labor income tax and should not use tax on savings.

E.g.: \( v(c_1, c_2) = u(c_1) + \frac{u(c_2)}{1+\delta} \)

Tax on savings justified within Saez (2002) framework if:

1. High skill people have higher taste for saving (e.g., high skill people have lower discount rate, better education)
2. \( c_2 \) is complementary with leisure.
3. Inheritances (won’t have same consumption patterns conditional on earned income).
A Simpler Theory of Optimal Capital Taxation

Emmanuel Saez and Stefanie Stantcheva

% of total household wealth

- Top 0.5%-0.1%
- Top 0.1%-0.01%
- Top 0.01%
- Top 1%-0.5%

Year:
- 1960
- 1965
- 1970
- 1975
- 1980
- 1985
- 1990
- 1995
- 2000
- 2005
- 2010
- 2015
A Simpler Model of Capital Taxation

For exposition: Exogenous and uniform labor income \( z \)

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

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

Time invariant tax \( T_K(rk) \)

Initial wealth \( k_{i\text{init}} \), 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:

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

subject to

\[
\frac{dk_i(t)}{dt} = rk_i(t) - T_K(rk_i(t)) + z_i(t) - c_i(t)
\]
Solving the Individual’s Maximization Problem

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

s.t. \[ \frac{d k_i(t)}{d t} = r k_i(t) - T_K(r k_i(t)) + z_i(t) - c_i(t) \]

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

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

FOC in \( k_i(t) \): \[ a'_i(k_i(t)) + \lambda_i(t) \cdot r \cdot (1 - T'_K) = - \frac{d \lambda_i(t)}{d t} + \delta_i \cdot \lambda_i(t) \]

\[ \Rightarrow a'_i(k_i(t)) = \delta_i - \bar{r} \quad \text{where} \quad \bar{r} = r \cdot (1 - T'_K) \]
Steady State

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

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

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

\[ \Rightarrow \text{Heterogeneity in (non-degenerate) steady-state wealth.} \]

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

\[
U_i = r k_i(t) - T_K(r k_i(t)) + z_i(t) + a_i(k_i(t)) + \delta_i \cdot (k_i^{init} - k_i(t))
\]

Dynamic model equivalent to a static model:

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

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

\[ U_i = c_i + a_i(k_i) + \delta_i \cdot (k_i^{\text{init}} - k_i) \]  \quad \text{with} \quad c_i = r_k i - T_K(r_k 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 $i$ ($\int g_i = 1$).

2. **Efficiency costs**: Elasticity $e_K = (\bar{r} / k) \cdot (dk / d\bar{r})$ 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 $\tau_K$

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

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

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

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

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

$$\frac{dSWF}{d\tau_K} = rk^m \cdot \int_i g_i \cdot \left(1 - \frac{k_i}{k^m}\right) - rk^m \cdot \frac{\tau_K}{1 - \tau_K} \cdot e_K$$

- Mechanical Revenue net of Welfare Effect
- Behavioral Effect

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

13 41
Optimal Linear Capital Tax $\tau_K$

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

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

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

$$e_K = \infty.$$

$\tau_K > 0$ 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: $\tau_K = 1/(1 + a_{K}^{top} \cdot e^{top}_K)$ 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) \( \bar{G}_K(rk) \equiv \frac{\int_{\{i: rki \geq rk\}} g_i di}{P(rki \geq rk) \int_i g_i di} \) 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:

\[
T'_K(\infty) = \frac{1 - G_K(\infty)}{1 - G_K(\infty) + \alpha_K(\infty) \cdot e_K(\infty)}
\]
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 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? → role for nonlinear labor income tax.

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

(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 $rk$.

$\rightarrow \text{corr}(g_i, k_i) < 0 \rightarrow \tau_K > 0$. 
Adding in Labor Income Responses & Labor Taxation

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

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

$$T'_L(z) = \frac{1 - \tilde{G}_L(z)}{1 - \tilde{G}_L(z) + \alpha_L(z) \cdot e_L(z)}$$

1) $\tilde{G}_L(z) \equiv \frac{\int_{\{i: z_i \geq z\}} g_i d_i}{P(z_i \geq z) \int g_i d_i}$ 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:

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

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 \(\tau_L\)):

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

with \(\bar{g}_K = \int \frac{k_i g_i}{k^m}, \quad e_{Z,(1-\tau_K)} = \frac{dz^m}{d(1-\tau_K)} \frac{(1 - \tau_K)}{z^m}\)
Comprehensive nonlinear income taxation $T(rk + z)$

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

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

Standard Mirrlees' formula applies to comprehensive income tax problem

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

with $\G(y) \equiv \int_{\{i:y_i \geq y\}} g_i d_i$

$\alpha_Y(y)$ 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 \( d_i(x) \)

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 \( \tau_L, \tau_K \) closer than they would be absent any shifting

If shifting elasticity is large then \( e_K \) can appear large, but wrong to set \( \tau_K \) at \( 1/(1 + e_K) \) in that case
Heterogeneous returns $r_i$ important in practice:
Same sufficient stats formula, but replace:

$$\bar{g} = \frac{\int f_i g_i \cdot r_i k_i}{\int f_i r_i k_i}$$
and

$$e_K = \frac{(1 - \tau_K)}{\int f_i r_i k_i} \cdot \frac{d \int f_i r_i k_i}{d(1 - \tau_K)}$$

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 \( \bar{g}_K \).

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

\[
\tau^j_K = \frac{1 - \bar{g}^j_K}{1 - \bar{g}^j_K + e^j_K}
\]

\[
\bar{g}^j_K = \frac{\int_i g_i \cdot k^j_i}{\int_i k^j_i}, \quad e^j_K = \frac{\bar{r}^j}{k^{m,j}} \cdot \frac{dk^{m,j}}{d\bar{r}^j} > 0, \quad e_{Ks,(1-\tau^j)} = \frac{\bar{r}^j}{k^{m,s}} \cdot \frac{dk^{m,s}}{d\bar{r}^j}
\]
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 $e_{K^s,(1-\tau_K^j)}$.

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

$$\bar{g}_K^j = \frac{\int_i g_i \cdot k_i^j}{\int_i k_i^j}, \quad e_K^j = \frac{\bar{r}_j}{k_{m,j}} \cdot \frac{dk_{m,j}^j}{d\bar{r}_j} > 0, \quad e_{K^s,(1-\tau_K^j)} = \frac{\bar{r}_j}{k_{m,s}} \cdot \frac{dk_{m,s}^j}{d\bar{r}_j}$$
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 $\tau_C$ 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, $\tau_C$ equivalent labor tax + tax on initial wealth (Kaplow, 1994, Auerbach, 2009).

Thought experiment: equal labor income.

With $\tau_C$, 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).

$\Rightarrow \tau_C$ 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: \((\theta_i, \eta_i)\).

Based on the IRS micro data, use pairs \((z_i, r_{ki})\) to invert individual choices to obtain \((\theta_i, \eta_i)\).

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 $T'_K(rk)$

![Graph showing optimal capital income tax rate $T'_K(rk)$ with different curves for $e_K = 1$, $e_K = 0.5$, and $e_K = 0.25$.](image)

- $e_K = 1$
- $e_K = 0.5$
- $e_K = 0.25$

The graph plots the marginal tax rate against capital income, illustrating how the optimal tax rate varies with income and specific parameters $e_K$. The x-axis represents capital income ranging from $200,000$ to $1,000,000$, while the y-axis shows the marginal tax rate ranging from $0$ to $1$. Each curve corresponds to a different fiscal parameter $e_K$, demonstrating the impact of these parameters on the optimal tax structure.
Optimal Tax Rate on Comprehensive Income $T'_Y(y)$

<table>
<thead>
<tr>
<th>Total Income</th>
<th>Marginal Tax Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>$200,000$</td>
<td>0</td>
</tr>
<tr>
<td>$400,000$</td>
<td>0.1</td>
</tr>
<tr>
<td>$600,000$</td>
<td>0.2</td>
</tr>
<tr>
<td>$800,000$</td>
<td>0.3</td>
</tr>
<tr>
<td>$1,000,000$</td>
<td>0.4</td>
</tr>
</tbody>
</table>

$e_Y = 1$

$e_Y = 0.25$

$e_Y = 0.5$

$e_Y = 1$
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 \( u_i(c, k, z) \) concave in \( c \), 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:

\[ SWF = \int \omega_i V_i(\{c_i(t), k_i(t), z_i(t)\}_{t \geq 0}) \]
Optimal Linear Capital Tax in the Steady State

Given $\tau_K$ and $\tau_L$, 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:

$$\tau_K = \frac{1 - \bar{g}_K - \tau_L \bar{z}_m / k^m \cdot e_L,1-\tau_K}{1 - \bar{g}_K + \bar{e}_K}$$

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 $\tau_K$ and $\tau_L$, rebated lump-sum $\rightarrow$ convergence to steady state.

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

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

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

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 \( \tau_K \) and \( \tau_L \), rebated lump-sum \( \rightarrow \) convergence to steady state.

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

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

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

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 $\tau_K$ and $\tau_L$, 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:

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

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?
REFERENCES


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


