14.471: Public Economics

Capital Income Taxation

Emmanuel Saez

MIT: Fall 2009

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

$\Rightarrow$ 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.

$\Rightarrow$ Efficiency cost of capital taxation might be high.

MOTIVATION

3) Capital more mobile internationally than labor $\Rightarrow$ Incidence of capital taxation might fall on workers:

Open economy with fully mobile capital, net-of-tax rate of return is fixed by the international rate of return r^* so that $(1-\tau)f'(k) = r^*$ where k is capital stock per person

Wage $w = f(k) - r^*k$ falls with τ

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

MACRO FRAMEWORK

Constant return to scale aggregate production:

$Y = F(K, L) = rK + wL =$ output $=$ income

$K =$ capital stock (wealth), $L =$ labor input

$r =$ rate of return on capital, w is wage rate

$rK =$ capital income, $wL =$ labor income

$\alpha = rK/Y =$ capital income share (constant α when $F(K, L) = K^{\alpha}L^{1-\alpha}$ Cobb-Douglas), $\alpha \simeq 30\%$

$\beta = K/Y =$ wealth to annual income ratio, $\beta \simeq 5 - 6$

$r = (rK/Y) \cdot (Y/K) = \alpha/\beta$, $r = 5 - 6\%$

Infinite horizon model: $U = \sum_t u(c_t)/(1+\delta)^t \Rightarrow r = \delta$ (discount rate) and $\beta = \alpha/\delta$

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 $\uparrow$ 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 $\uparrow$ individual savings but $\downarrow$ govt saving [if govt spending stays constant]

FACTS ABOUT WEALTH AND CAPITAL INCOME

Definition: Capital Income = Returns from Wealth Holdings

Aggregate US **Personal** Wealth = 3.5*GDP = $ 50 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, Consumer credit debt

Substantial amount of financial wealth is held indirectly through: pension funds [DB+DC], mutual funds, insurance reserves

CAPITAL INCOME IN NATIONAL ACCOUNTS

Gross capital income (before depreciation) is about 40% of GDP

Net capital income (after depreciation) is about 25% of personal income

The capital income share in total income is relatively stable in the long-run (but with some short term fluctuations)

Average real rate of return of capital around 5-6%, varies greatly from year to year

FACTS ABOUT WEALTH AND CAPITAL INCOME

Wealth $= W$, Return $= r$, Capital Income $= rW$

$$W_t = W_{t-1} + r_t W_{t-1} + E_t + I_t - C_t$$

where W_t is wealth at age t, C_t is consumption, E_t labor income earnings (net of taxes), r_t is the average (net) rate of return on investments and I_t net inheritances (gifts received and bequests - gifts given).

Replacing W_{t-1} and so on, we obtain the following expression (assuming initial wealth W_0 is zero):

$$W_t = \sum_{k=1}^{t} (E_k - C_k + I_k) \prod_{j=k+1}^{t} (1 + r_j)$$

FACTS ABOUT WEALTH AND CAPITAL INCOME

$$W_t = \sum_{k=1}^{t} (E_k - C_k + I_k) \prod_{j=k+1}^{t} (1 + r_j)$$

Differences in Wealth and Capital income due to:

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

2) Net Inheritances received I_t [transfer wealth]

3) Rates of return r_t

[more details in Davies-Shorrocks '00, Handbook of Income Distribution]

WEALTH DISTRIBUTION

Wealth inequality is very large

US Household Wealth is divided 1/3,1/3,1/3 for the top 1%, the next 9%, and the bottom 90% [bottom 1/3 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 DISTRIBUTION

Wealth is more unequally distributed than income [true in all countries]

Top 1% income share in the US is around 20%

Top 1% labor income share in the US (among workers) is around 15%

US Income concentration has ↑ sharply since 1970:

top 1% income share was 9% in 1970 and 23.5% in 2007 [Piketty-Saez QJE'03 updated]

US Wealth concentration has only slightly increased:

Top 1% wealth share has grown "only" from 31% in 1962 to 34% in 2007 based on the Survey of Consumer Finances [Scholz '03, Kennickell '09]

FACTS OF US CAPITAL INCOME TAXATION

Good US references: Gravelle '94 book, Slemrod-Bakija '04 book

1) **Corporate Income Tax** (fed+state): 35% Federal tax rate on profits of corporations [complex rules with many industry specific provisions]

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 bonds interest are exempt

FACTS OF US CAPITAL INCOME TAXATION

3) **Estate and gift taxes:**

Fed taxes estates above $3.5M exemption (only .2% of deceased liable), top rate is 45%

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

LIFE CYCLE MODEL OF WEALTH (MODIGLIANI)

Individuals work for R years and live for T years: $T - R$ is retirement duration

Individuals earn income w_t from period 0 to R and earn zero afterwards

Individuals have additive separable utility $\sum_{t=1}^{T} u(c_t)/(1+\delta)^t$ with concave $u(.)$

subject to inter-temporal budget constraint: $\sum_{t=1}^{T} c_t/(1+r)^t \leq \sum_{t=1}^{R} w_t/(1+r)^t$ (multiplier λ)

FOC: $u'(c_t)/(1+\delta)^t = \lambda/(1+r)^t$

Euler equation: $u'(c_{t+1})/u'(c_t) = (1+\delta)/(1+r)$

LIFE CYCLE MODEL OF WEALTH (MODIGLIANI)

Euler equation: $u'(c_{t+1})/u'(c_t) = (1+\delta)/(1+r)$

If $\delta < r$, $c_{t+1} > c_t \Rightarrow$ Individuals save to consume more later on

If $\delta > r$, $c_{t+1} < c_t \Rightarrow$ Individuals want to consume more earlier on

If $\delta = r$ then c_t is constant with t:

$\Rightarrow$ Individuals want to smooth consumption by saving while working and consuming saving while retired $\Rightarrow$ Wealth W_t is inversely U-shaped during life-cycle

$\Rightarrow$ Wealth inequality only slightly higher than labor income inequality [does not fit facts]

OTHER FACTORS AFFECTING WEALTH DISPERSION

1) Heterogeneity in tastes for saving:

- traditional discount rate

- self-control problems [hyperbolic discount rate] and financial education

2) Rates of returns received on assets: traditional risk aversion, luck, but also financial education

3) Net inheritances and gifts received [in general from parents]

LIFE CYCLE VS. INHERITED WEALTH

Old view: Tobin and Modigliani: life cycle wealth accounts for the bulk of the wealth hold in the US. Kotlikoff-Summers JPE'81 challenged the old view.

Debate: Kotlikoff vs. Modigliani in JEP'88.

Why is this question important?

1) Economic Modelling: what accounts for wealth accumulation and inequality? Is widely used life-cycle model with no bequests a good approximation? [Causality between growth and savings]

2) Policy Implications: taxation of capital income and estates. Role of pay-as-you-go vs. funded retirement programs

LIFE CYCLE VS. INHERITED WEALTH

W total wealth in the economy, LCW is life cycle wealth, and T wealth due to transfers

Two components in the individual wealth equation W_t:

$$LCW_t = \sum_{k=1}^{t} (E_k - C_k)(1+r)^{t-k}$$

$$T_t = \sum_{k=1}^{t} I_k(1+r)^{t-k}$$

Aggregate this over all individuals or households in the economy to estimate T/W and LCW/W. Two methods:

(1) Compute T_t (flow of bequests method)

(2) Compute LCW_t (comparison of earnings and consumption)

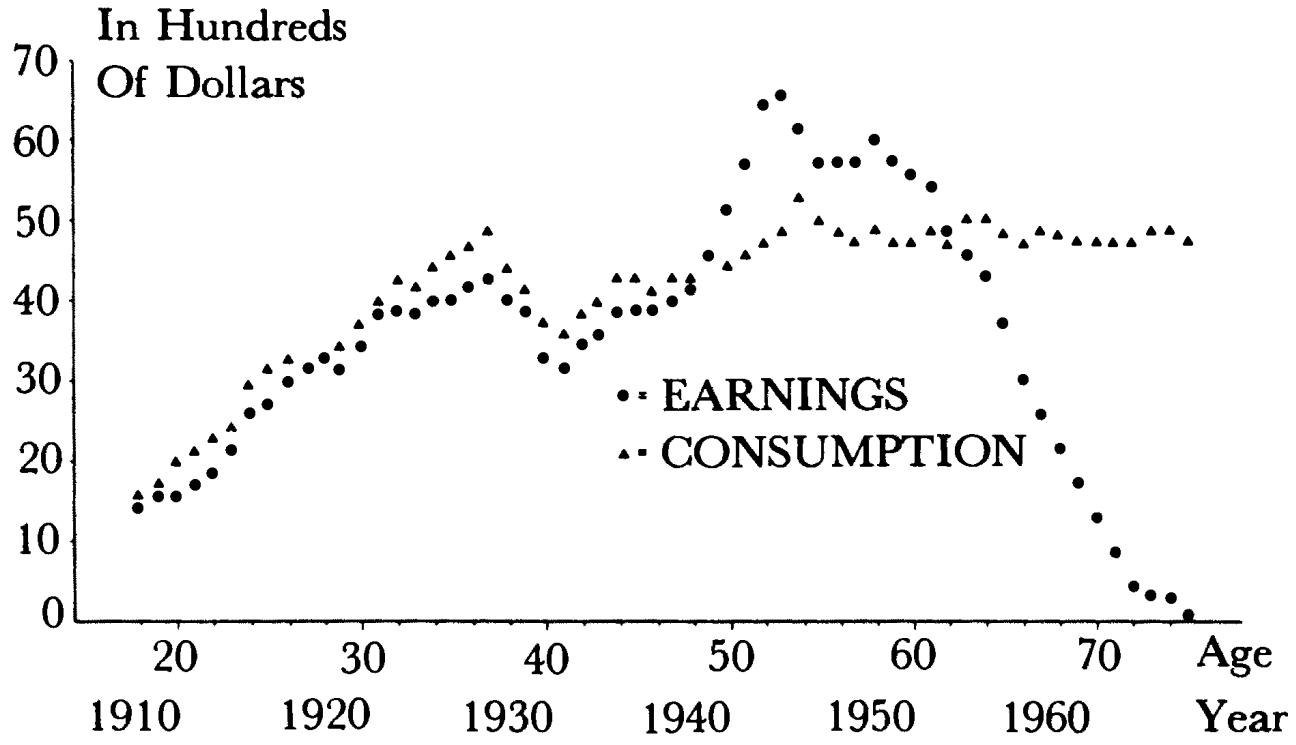

Fig. 1. Sum of male and female longitudinal average earnings and average consumption profiles, age 18 in 1910–age 82 in 1974

LIFE CYCLE VS. INHERITED WEALTH

(a) Modigliani JEP'88 claims that over 2/3 of wealth is due to life-cycle

(b) Kotlikoff-Summers JPE'81, JEP'88 claim that over 2/3 of wealth is due to transfers

Differences due primarily in methodology (Gale and Scholtz JEP'94):

(a) how to capitalize past transfers

(b) whether to count college tuition paid by parents as transfers

Transfer wealth is probably quite important, especially at the top of the wealth distribution

Table 4

Intergenerational Transfers as a Source of Capital Accumulation, 1986

Transfer Category	*Annual Flow ($ billions)*	*Stock of Transfer Wealth ($ billions) ($r - n = 0.01$)*
Support Given to:		
Children	32.69	1346.7
Parents	3.37	−104.3
Grandparents	0.07	−4.0
Grandchildren	5.05	416.2
Trusts	14.17	576.1
Life Insurance	7.84	258.3
Totals		
Intended Transfers	63.19	2489.3
College Payments	35.29	1441.5
Bequests	105.00	3708.1
As a % of net worth[a]		
Intended Transfers	0.53	20.8
College Expenses	0.29	12.0
Bequests	0.88	31.0

Source: Authors' calculations from the Survey of Consumer Finances.
[a]Aggregate net worth in the SCF in 1986 is $11,976 billion.

LIFE CYCLE VS. INHERITED WEALTH

More interesting question: how do the shares of inheritance vs. life-cycle evolve over time?

Inheritance share likely huge in the distant past: class society with rentiers vs. workers [Delong '03]

Inheritance share likely $\downarrow$ in 20th century but might have $\uparrow$ recently (Piketty '10 for France)

Post-war period was a time of fast population growth and fast economic growth $\Rightarrow$ If n (growth) large relative to r (rate of return on wealth) $\Rightarrow$ Inheritances play a minor role in life-time wealth

Could be an exceptional episode and Western countries are going back to earlier situation where inheritances were important

KEY ELEMENTS OF DEBATE ON CAPITAL INCOME TAXATION

Economic debate:

1) Distributional concerns: capital income accrues disproportionately to higher income families

2) Efficiency concerns: capital tax distorts savings, business creation, capital mobility across countries

Public policy debate:

3) Should we tax income vs. consumption? [Fundamental tax reform debate]

4) Should we encourage savings by cutting tax on capital income or with tax favored savings vehicles?

TAXES IN OLG LIFE-CYCLE MODEL

max $U = u(c_1, l_1) + \delta u(c_2, l_2)$

No tax situation: earn $w_1 l_1$ in period 1, $w_2 l_2$ in period 2

Savings $s = w_1 l_1 - c_1$, $c_2 = w_2 l_2 + (1 + r)s$

Capital income rs

Intertemporal budget with no taxes:

$c_1 + c_2/(1 + r) \leq w_1 l_1 + w_2 l_2/(1 + r)$

This model has uniform rate of return and does not capture excess returns

TAXES IN OLG MODEL

Budget with consumption tax t_c:

$$(1 + t_c)[c_1 + c_2/(1 + r)] \leq w_1 l_1 + w_2 l_2/(1 + r)$$

Budget with labor income tax τ_L:

$$c_1 + c_2/(1 + r) \leq (1 - \tau_L)[w_1 l_1 + w_2 l_2/(1 + r)]$$

Consumption and labor income tax are equivalent if

$$1 + t_c = 1/(1 - \tau_L)$$

Both taxes distort only labor-leisure choice

TAXES IN OLG MODEL

Budget with capital income tax τ_K:

$$c_1 + c_2/(1 + r(1 - \tau_K)) \leq w_1 l_1 + w_2 l_1/(1 + r(1 - \tau_K))$$

τ_K distorts only inter-temporal consumption choice

Budget with comprehensive income tax τ:

$$c_1 + c_2/(1 + r(1 - \tau)) \leq (1 - \tau)[w_1 l_1 + w_2 l_2/(1 + r(1 - \tau))]$$

τ distorts both labor-leisure and inter-temporal consumption choices

τ imposes "double" tax: (1) tax on earnings, (2) tax on savings

EFFECT OF r ON SAVINGS

Assume that labor supply is fixed. Suppose $r \uparrow$:

1) Substitution effect: price of $c_2 \downarrow \Rightarrow c_2 \uparrow$, $c_1 \downarrow \Rightarrow$ savings $s = w_1 l_1 - c_1 \uparrow$.

2) Wealth effect: Price of $c_2 \downarrow \Rightarrow$ both c_1 and $c_2 \uparrow \Rightarrow$ save less

3) Human wealth effect: present discounted value of labor income $\downarrow \Rightarrow$ both c_1 and $c_2 \downarrow \Rightarrow$ save more

Note: If $w_2 l_2 < c_2$ (ie $s > 0$), 2)+3) $\Rightarrow$ save less

Total net effect is theoretically ambiguous $\Rightarrow \tau_K$ has ambiguous effects on s

SHIFT FROM LABOR TO CONSUMPTION TAX

Labor and consumption are equivalent for the individual if $1 + t_c = 1/(1 - \tau_L)$ but savings pattern is different

Assume $w_2 = 0$ and $l_1 = 1$

$(1 + t_c)[c_1 + c_2/(1 + r)] = w_1$ with consumption tax

$c_1 + c_2/(1 + r) = (1 - t_L)w_1$ with labor tax

1) Consumption tax t_c: $c_1^c = (w_1 - s_c)/(1 + t_c)$, $c_2^c = (1 + r)s_c/(1 + t_c)$

2) Labor income tax τ_L: $c_1^L = w_1(1 - \tau_L) - s_L$, $c_2^L = (1 + r)s_L$

Same consumption in both cases so $s_L = s_c/(1 + t_c) \Rightarrow$ Save more with a consumption tax

TRANSITION FROM LABOR TO C TAX

In OLG model and closed economy, capital stock is due to life-cycle savings s

Start a labor tax τ_L and you decide to switch to a consumption tax t_c

The old [at time of transition] would have paid nothing in labor tax regime but now have to pay tax on c_2

For the young [and future generations], the two regimes look equivalent so they now save more and increase the capital stock

However, this increase in capital stock comes at the price of hurting the old who are taxed twice

TRANSITION FROM LABOR TO C TAX

Suppose the government keeps the old as well off as in previous system by exempting them from consumption tax

This creates a deficit in government budget equal to

$$d = \tau_L w_1 - t_c c_1 = t_c w_1/(1+t_c) - t_c c_1 = t_c s_L$$

Extra saving by the young is $s_c - s_L = t_c s_L$ exactly equal to government deficit.

Full neutrality result: Extra savings of young is equal to old capital stock + new government deficit $\Rightarrow$ no change in the aggregate capital stock

Full neutrality depends crucially on same r for govt debt and aggregate r [in practice: equity premium puzzle]

[Same result for Social Security privatization]

AUERBACH-KOTLIKOFF '87 MODEL

Develop an inter-temporal Computational General Equilibrium (CGE) model:

1) Life cycle model, no bequests, people live for 55 years (born at age 21). Work for 45 years, and retire for 10 years.

2) Only one individual per cohort, representative agent model [Useful for redistribution analysis across cohorts but not within cohorts]

3) Stock of wealth = life cycle savings [Classical Modigliani graph]

4) Labor income tax distorts labor supply, capital income tax distorts savings choice.

CES utility, discount rate, path of earnings with life cycle.

AUERBACH-KOTLIKOFF '87: RESULTS

Tax reform experiments: shift from comprehensive income tax to either

(a) pure consumption tax

(b) pure wage income tax

(c) pure capital income tax

[budget neutral but no transitional compensation]

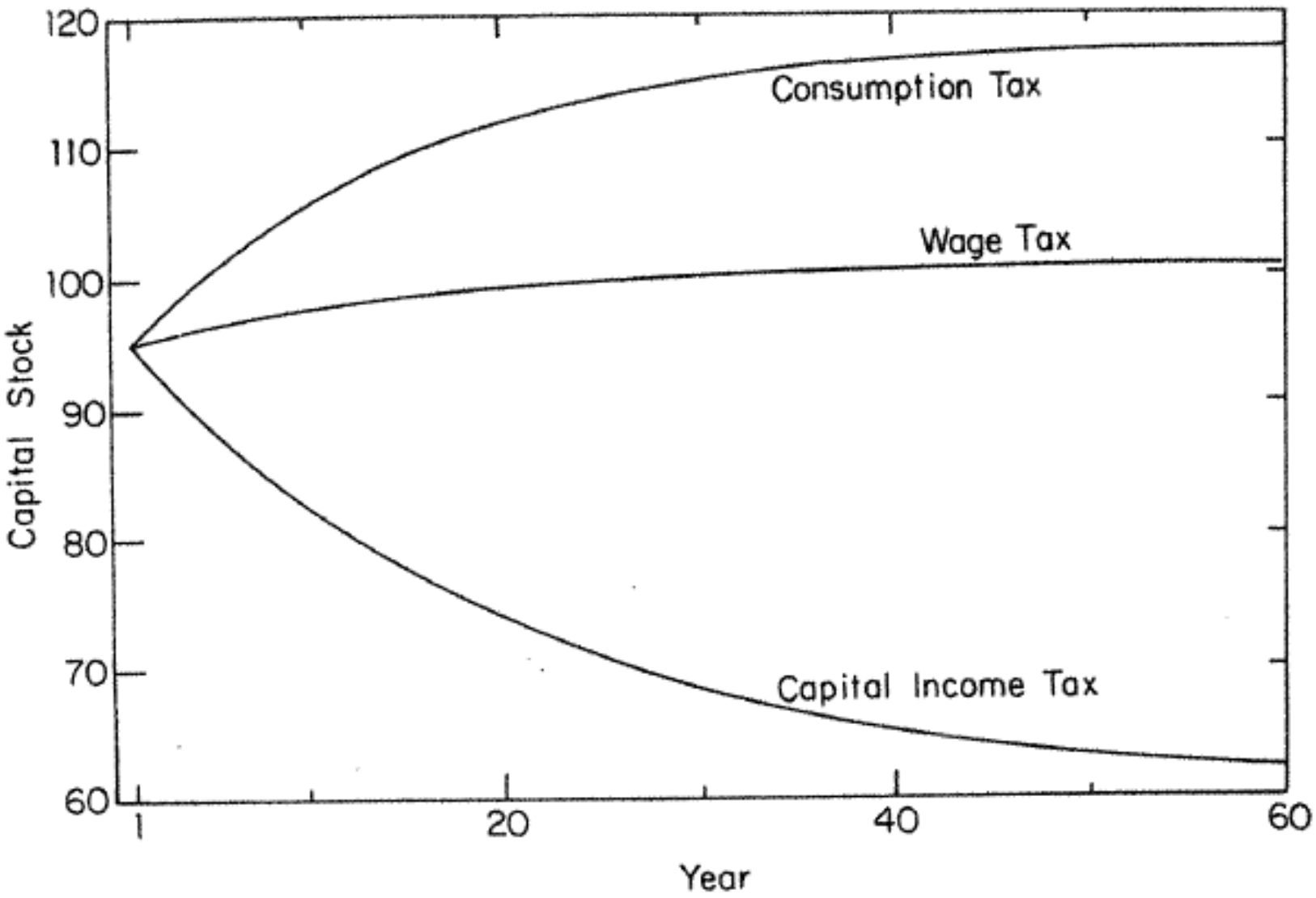

Figure 5.3. The impact on capital formation of tax reform.

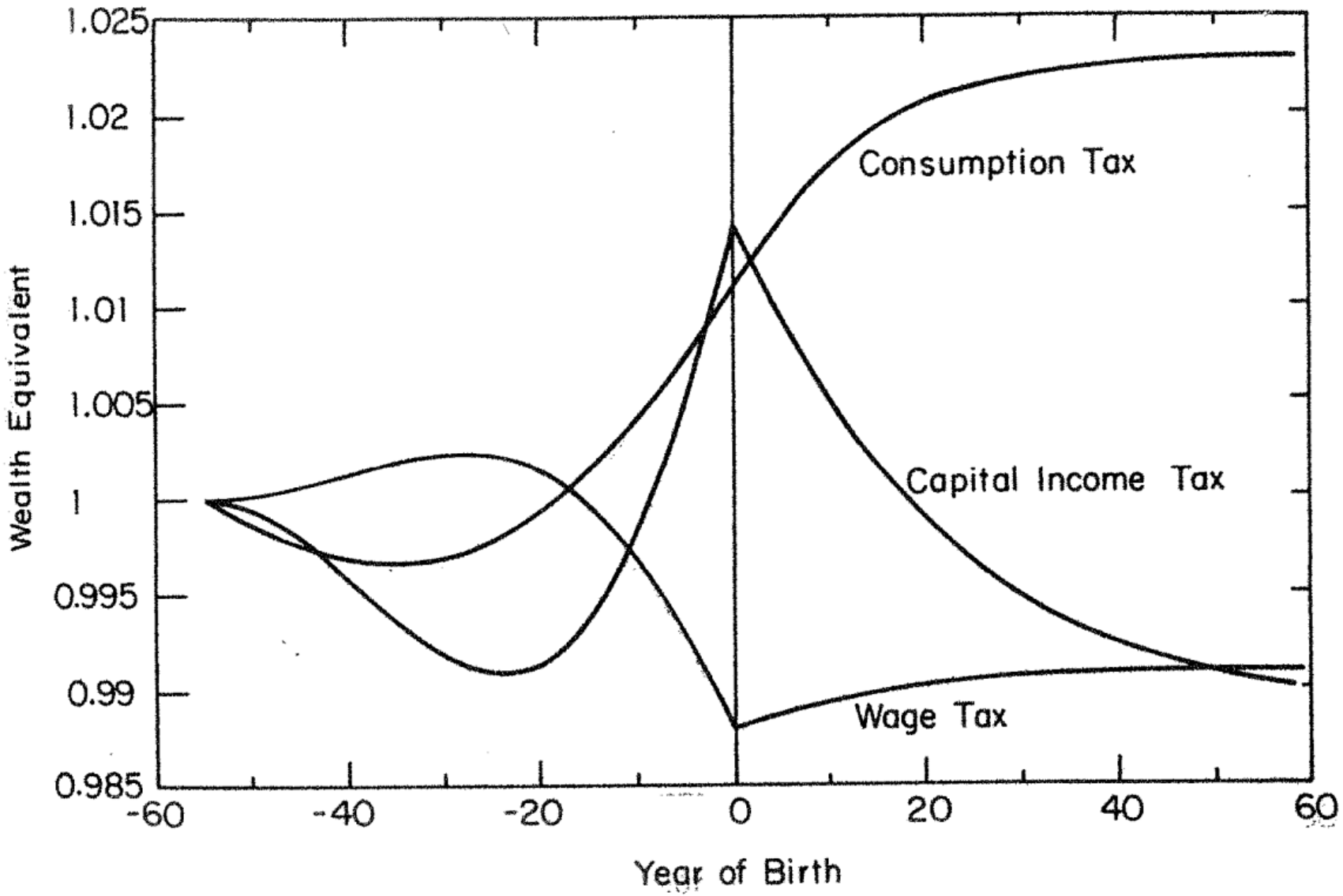

Figure 5.4. The welfare effects of tax reform.

AUERBACH-KOTLIKOFF '87: RESULTS

1) Effect on capital stock:

(a) Consumption tax is best (because no compensation of the old)

(b) Wage income tax has limited impact on capital stock

(c) Capital income tax is worst (significant elasticity of savings wrt to r).

AUERBACH-KOTLIKOFF '87: RESULTS

2) Effect on welfare measured in percentage increase of consumption for each generation:

Consumption tax hurts current old and benefits the young and future generations [no transitional relief]

Wage income tax benefits the old but hurts the young

Capital income tax hurts current generation (double tax), benefits next generation (implicit levy of previously accumulated capital) but hurts future generations (inefficient)

Key lessons: Transitional reliefs rules and anticipated vs. not tax changes has large impact on results

OPTIMAL CAPITAL INCOME TAXATION

Complex problem with many sub-literatures: Banks and Diamond Mirrlees Review '09 provide excellent recent survey

1) Life-cycle models [linear and non-linear earnings tax]

2) Models with bequests [many models including the infinite horizon model]

3) Models with future earnings uncertainty: New Dynamic Public Finance [Kocherlakota '09 book]

Bigger gap between theory and policy practice than in the case of static labor income taxation

RAMSEY TAX IN LIFE-CYCLE MODEL

Best reference is King (1980) [volume Heal-Hughes]. Also Atkinson-Sandmo EJ '80, Sandmo '85 PE Handbook Chapter, Atkinson-Stiglitz '80, Chap 14-4.

Ramsey model with **representative** agent and **linear** taxes on labor and savings to raise exogenous amount of revenue

Individual maximization problem:

$$V(q, w(1-\tau_L) = max_{c_1,c_2,l}\ u(c_1, c_2, l)$$

st $c_1 + c_2/(1 + r(1-\tau_K)) = wl(1-\tau_L)$

where $q = 1/(1 + r(1-\tau_K))$ and $p = 1/(1+r)$

RAMSEY CAPITAL INCOME TAX

Optimal tax rates can be obtained by solving the standard Ramsey problem:

$max\ V(q, w(1-\tau_L))$

st $wl\tau_L + (q-p)c_2 \geq g\ (\lambda)$

where g is exogenous tax revenue requirement

Can apply the results from the 3 good Ramsey model

Derive FOC for τ_K and τ_L

Can express them in terms of compensated elasticities

RAMSEY CAPITAL INCOME TAX

Combining the two FOC to get rid of λ, you get:

$$\frac{r\tau_K}{1+r}(\sigma_{L2}-\sigma_{22}) = \frac{\tau_L}{1-\tau_L}(\sigma_{LL}-\sigma_{2L})$$

where $\sigma_{LL} = (w(1-\tau_L)/l)\partial l^c/\partial(w(1-\tau_L)) > 0$ is the compensated elasticity of labor supply with to wage rate.

$\sigma_{22} = (q/c_2)\partial c_2^c/\partial q < 0$

$\sigma_{L2} = (q/l)\partial l^c/\partial q$

$\sigma_{2L} = (w(1-\tau_L)/c_2)\partial c_2^c/\partial(w(1-\tau_L))$

Formula defines relative optimal rates of taxation on labor and capital (absolute levels depend on g)

RAMSEY CAPITAL INCOME TAX: DISCUSSION

Little known about cross elasticities so we might as well assume that they are zero [symmetric by Slutsky] $\Rightarrow$ Optimal formula simplifies to:

$$-\frac{r\tau_K}{1+r}\sigma_{22} = \frac{\tau_L}{1-\tau_L}\sigma_{LL}$$

Inverse elasticity rule as in standard Ramsey model: If $\sigma_{LL} << |\sigma_{22}|$ then τ_K should be small relative to τ_L

Key lesson: What matters is the relative size of elasticities, not the number of distortions

FELDSTEIN JPE'78

Feldstein JPE'78 makes famous theoretical argument why σ_{22} can be large even if $\varepsilon^u_{sq} = (q/s)\partial s/\partial q$ [uncompensated savings elasticity] is zero: Budget $c_1 + qc_2 = w(1-\tau_L)l + y$

Slutsky equation [y is endowment =0 in equilibrium]: $\partial c_2^c/\partial q = \partial c_2/\partial q + c_2 \partial c_2/\partial y \Rightarrow$

$\sigma_{22} = \varepsilon^u_{2q} + q\partial c_2/\partial y$

$c_2 = s/q$ so $\varepsilon^u_{2q} = (q/c_2)\partial c_2/\partial q = \varepsilon^u_{sq} - 1 \Rightarrow$

$\sigma_{22} = \varepsilon^u_{sq} - 1 + q\partial c_2/\partial y$

$c_1 + qc_2 = w(1-\tau_L)l + y \Rightarrow \partial c_1/\partial y + q\partial c_2/\partial y = w(1-\tau_L)\partial l/\partial y + 1 \simeq 1$ (small income effects on labor supply)

$\sigma_{22} \simeq \varepsilon^u_{sq} + \partial c_1/\partial y \simeq \partial c_1/\partial y \geq 0.75$ [as saving rate modest]

RAMSEY TAX: ENDOGENOUS CAPITAL STOCK

Full dynamic model:

Govt maximizes $SW = \sum_t V_t/(1+\delta)^t$

subject to $\sum_t Tax_t/(1+r)^t \geq \sum_t g_t/(1+r)^t$

$\Rightarrow$ Optimal dynamic capital stock k is given by Modified Golden rule $r = f'(k) = \delta$

Optimal k can be attained in steady state using debt policy [implicit in budget constraint]

RAMSEY TAX: ENDOGENOUS CAPITAL STOCK

If the govt cannot use debt policy then optimal dynamic capital level may not be attained because savings equal capital $s_t = K_t$
$\Rightarrow$ tax formulas need to be modified: optimal tax rate reflect

(a) the trade-off between conventional [intra-generational] efficiency losses [static Ramsey]

(b) the failure to achieve the dynamic optimality condition on capital stock [inter-generational efficiency trade-off]

$\Rightarrow$ Effect on capital tax rate level is actually ambiguous

RAMSEY CAPITAL INCOME TAX: REMARKS

1) No redistributive concerns: Can extend model to the multi-person case $\Rightarrow$ Higher rate τ_K if capital income concentrated among the rich (Park JPubE, 1991).

2) No bequests so this model does not capture an important aspect of wealth accumulation and justification for redistribution.

3) Only a two period model, if more periods are introduced (as in the Auerbach-Kotlikoff simulation model), then optimal tax formula would be more complex.

ATKINSON-STIGLITZ JpubE '76

Heterogeneous individuals and government uses nonlinear tax on earnings. Should the govt also use tax on savings?

$V^h = maxU^h(v(c_1, c_2), l)$ 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

Recent proof by Laroque EJ '05 or Kaplow JpubE '06.

Tax on savings justified if: (1) High skill people have higher taste for saving [Saez, JpubE '02 with calibration Golosov-Tsyvinski-Weinzerl '09], (2) c_2 is complementary with leisure

RECONCILING RAMSEY AND ATKINSON-STIGLITZ

1) Ramsey model: use relative elasticities rule

$$\frac{r\tau_K}{1+r}(\sigma_{L2}-\sigma_{22}) = \frac{\tau_L}{1-\tau_L}(\sigma_{LL}-\sigma_{2L})$$

2) Atkinson-Stiglitz: tax only labor income when $U(v(c_1,c_2),l)$

Why are results so different across the two models?

Atkinson-Stiglitz imposes strong implicit assumptions on cross elasticities: $\max U(V((1-\tau_L)wl+y,q),l) \Rightarrow \partial l^c/\partial q \neq 0$ and loosely speaking $\sigma_{2L} \simeq \sigma_{LL}$

Difficult to know whether $\sigma_{2L} \simeq 0$ is better assumption than $\sigma_{2L} \simeq \sigma_{LL}$

DIAMOND-SPINNEWIJN '09

Heterogeneity of individuals in ability (wage rate) and discount rate. Discrete earnings choice model (high vs. low) and discrete discount (high vs. low) [Four type model]

Govt can tax both earnings and savings non-linearly: bi-dimensional tax function with bi-dimensional heterogeneity

Start from no savings tax and optimal earnings tax

Result: introducing a small savings tax on high earners or a small savings subsidy on low earners increases welfare

Intuition: Those valuing the future more (relative to the disutility of work) are more willing to work than those valuing the future less $\Rightarrow$ work IC constraint binds for high wage/low savers but not for high wage/high saver $\Rightarrow$ Scope for taxing savings

LIMITS OF LIFE-CYCLE MODEL

Atkinson-Stiglitz shows that life-time savings should not be taxed, tax only labor income

From justice view: seems fair to not discriminate against savers if labor earnings is the only source of inequality and is taxed non-linearly

In reality, capital income inequality also due

(1) difference in rates of returns

(2) shifting of labor income into capital income

(3) inheritances

(1) is not relevant if individuals handle risky assets rationally (as in CAPM model), probably not a very good assumption $\Rightarrow$ Tax on lucky returns might be desirable

SHIFTING OF LABOR / CAPITAL INCOME

In practice, difficult to distinguish between capital and labor income [e.g., small business profits, professional traders]

Differential tax treatment can induce shifting:

(1) US C-corporations vs S-corporations: shift from corporate income and realized capital gains toward individual business income [Gordon and Slemrod '00]

(2) Carried interest in the US: hedge fund and private equity fund managers receive fraction of profits of assets they manage for clients. Those profits are really labor income but are taxed as realized capital gains

(3) Finnish Dual income tax system: taxes separately capital income at preferred rates since 1993: Pirtila and Selin (2007) show that it induced shifting from labor to capital income especially among self-employed

THEORY: SHIFTING OF LABOR / CAPITAL INCOME

Extreme case where government cannot distinguish at all between labor and capital income $\Rightarrow$ Govt observes only $wl + rk$ $\Rightarrow$ Only option is to have identical MTRs at individual level $\Rightarrow$ General income tax $Tax = T(wl + rk)$

With a finite shifting elasticity, differential MTRs for labor and capital income taxation induce an additional shifting distortion

The higher the shifting elasticity, the closer the tax rates on labor and capital income should be [Christiansen and Tuomala ITAX'08]

In practice, this seems to be a very important consideration when designing income tax systems [especially for top incomes] $\Rightarrow$ Strong reason for having $\tau_L = \tau_K$ at the top

TAXATION OF INHERITANCES: WELFARE EFFECTS

Definitions: donor is the person giving, donee is the person receiving

Inheritances and inter-vivos transfers raise difficult issues:

(1) Inequality in inheritances contributes to economic inequality: seems fair to redistribute from those who received inheritances to those who did not

(2) However, it seems unfair to double tax the donor who worked hard to pass on wealth to children

$\Rightarrow$ Double welfare effect: inheritance tax hurts donor (if donor altruistic to donee) and donee (which receives less)

TAXATION OF INHERITANCES: BEHAVIORAL RESPONSES

Potential behavioral response effects of inheritance tax:

(1) reduces wealth accumulation of altruistic donors (and hence tax base)

(2) reduces labor supply of altruistic donors (less motivated to work if cannot pass wealth to kids)

(3) induces donees to work more through income effects (Carnegie effect, Holtz-Eakin,Joulfaian,Rosen QJE'93)

Critical to understand why there are inheritances to decide on optimal inheritance tax policy. 4 main models of bequests: (a) accidental, (b) warm glow, (c) manipulative bequest motive, (d) dynastic

ACCIDENTAL BEQUESTS

People die with a stock of wealth they intended to spend on themselves: Such bequests arise because of imperfect annuity markets

Annuity is an insurance contract converting lumpsum amount into a stream of payments till end of life [insurance against risk of living too long]

Annuity markets are imperfect because of adverse selection [Finkelstein-Poterba EJ'02, JPE'04] or behavioral reasons [inertia, lack of planning]

Public retirement programs [and old defined benefit private pensions] are in general annuities

Newer defined contribution private pensions [401(k)s in the US] are in general not annuitized

ACCIDENTAL BEQUESTS

Bequest taxation has no distortionary effect on behavior of donor and can only increase labor supply of donees (through income effects) $\Rightarrow$ strong case for taxing bequests heavily

Kopczuk JPE '03 makes the point that estate tax plays the role of a second best annuity:

Estate tax paid by those who die early, and not by those who die late $\Rightarrow$ Implicit insurance against risk of living too long

Same tax policy conclusion arises if donors have wealth in their utility function [social status or power, Carroll '00]

WARM GLOW OR ALTRUISTIC BEQUESTS

$u(c)-h(l)+\delta v(b)$ where c is own consumption, l is labor supply, and b is net-of-tax bequests left to next generation and $v(b)$ is warm glow utility of bequests

Budget with no estate tax: $c + b/(1 + r) = wl - T_L(wl)$

Budget with estate tax at rate τ_E: $c + b/[(1 + r)(1 - \tau_E)] = wl - T_L(wl)$

Suppose first that b is not bequeathed but used for "after-life" consumption [e.g., funerary monument of no value to next generation]

$\Rightarrow$ Atkinson-Stiglitz implies that b should not be taxed [$\tau_E = 0$] and that nonlinear tax on wl is enough for redistribution

WARM GLOW OR ALTRUISTIC BEQUESTS

Suppose now that b is given to a heir derives utility $v^{heir}(b)$ $\Rightarrow$ b creates a positive externality and should be subsidized $\Rightarrow$ $\tau_E < 0$ is optimal

Kaplow '01 makes this point informally

Farhi-Werning QJE'10 develop formal model of non-linear Pigouvian subsidization of bequests with 2 generations and social Welfare: $SW = \int [u(c) - h(l) + \delta v(b) + v^{heir}(b)] f(w) dw$

The marginal external effect of bequests is dv^{heir}/db and hence should be smaller for large b

$\Rightarrow$ Optimal subsidy is smaller for large estates $\Rightarrow$ progressive estate subsidy

WARM GLOW BEQUESTS: ISSUES

(a) If past inheritances come from untaxed labor income, then it is desirable to tax inheritances [important when income tax starts]

(b) Double counting issue: should social welfare double count bequests? [both for donor and donee]

Yes under utilitarian framework [Kaplow '01]

No: utilitarian framework with double counting generates predictions that conflict with horizontal equity:

- Govt should tax less those well loved by other people

- Govt should care more about kids with parents than orphans

MANIPULATIVE BEQUESTS

Parents use potential bequest to extract favors from children

Empirical Evidence: Bernheim-Shleifer-Summers JPE '85 show that number of visits of children to parents is correlated with bequeathable wealth but not annuitized wealth of parents

$\Rightarrow$ Bequest becomes one additional form of labor income for donee and one consumption good for donor

$\Rightarrow$ Inheritances should be counted and taxed as labor income for donees

SOCIAL-FAMILY PRESSURE BEQUESTS

Parents may not want to leave bequests but feel compelled to by pressure of heirs or society: bargaining between parents and children

With estate tax, parents do not feel like they need to give as much $\Rightarrow$ parents are made better-off by the estate tax $\Rightarrow$ Case for estate taxation stronger [Atkinson-Stiglitz does not apply and no double counting of bequests]

Empirical evidence:

Aura JpubE'05: reform of private pension annuities in the US in 1984 requiring both spouses signatures when worker decides to get a single annuity or couple annuity: reform $\uparrow$ sharply couple annuities choice

Equal division of estates [Wilhelm AER'96, McGarry]: estates are very often divided equally but gifts are not

DYNASTIC MODEL OR INFINITE HORIZON

Special case of warm glow: $V_t = u(c_t, l_t) + \delta V_{t+1}$ implies

$$V_0 = \sum_t \delta^t u(c_t, l_t)$$

st $\sum_t c_t/(1+r)^t \leq \sum_t w_t l_t/(1+r)^t$

Dynasty with Ricardian equivalence: consumption depends only on PDV of earnings of dynasty

Poor empirical fit:

1) Altonji-Hayashi-Kotlikoff AER'92, JPE'97 show that income shocks to parents have bigger effect on parents consumption than on kids consumption

2) Temporary tax cut debt financed [fiscal stimulus] should have no impact on consumption

INFINITE HORIZON MODEL: CHAMLEY-JUDD

Infinite horizon with no uncertainty. Govt can collect taxes using labor income tax or capital income taxes (but cannot confiscate initial wealth which would be optimal) that vary period by period

$$V_0 = \sum_t u(c_t, l_t)/(1+\delta)^t$$

st $\sum_t q_t c_t \leq \sum_t q_t w_t (1-\tau_L^t) l_t + A_0$ (λ)

$q_0 = 1$, $q_1 = 1/(1+r_1(1-\tau_K^1)$, ..., $q_t = 1/\prod_{s=1}^t (1+r_s(1-\tau_K^s))$

FOC in l_t and $c_t \Rightarrow w_t(1-\tau_L^t)u_c^t - u_l^t = 0$, $u_c^{t+1}/u_c^t = (1+\delta)/(1+r(1-\tau_K^t))$

With constant tax rate τ_K and constant r: Before tax price: $p_t = 1/(1+r)^t$ and after-tax price $q_t = 1/(1+r(1-\tau_K))^t \Rightarrow$

Price distortion q_t/p_t grows exponentially with time

CHAMLEY-JUDD: RESULTS

Chamley-Judd show that the capital income tax rate always tends to zero asymptotically: no capital tax in the long-run. This is due to 2 reasons, each of which is actually sufficient:

(1) Infinite supply elasticity of the capital stock k with respect to the net-of-tax rate of return $r(1 - \tau_K)$

(2) Government objective maximizes welfare of the dynasty seen from the first generation [V_0 objective]

CHAMLEY-JUDD, INFINITE ELASTICITY

Two classes: capitalists save as in infinite horizon model, workers do not save (consume wages w with no labor supply effects)

Can capital tax at rate τ be desirable for workers in steady-state?

$r = f'(k)$ and $w = f(k) - kf'(k)$, tax τ, net return is $(1-\tau)f'(k)$

Infinite horizon: modified Golden rule: discount rate $\delta = (1-\tau)f'(k)$ (if $>$, save more and k increases, if $<$, save less and k decreases).

Workers get $w + f'(k)\tau k = f(k) - (1-\tau)kf'(k) = f(k) - \delta k$, maximized when $f'(k) = \delta$, ie $\tau = 0$

Intuition: Supply of k is infinitely elastic: taxing an infinitely elastic good cannot be desirable [even in the steady state]

CHAMLEY-JUDD, V_0 objective

It is possible to build a model with endogenous discount rate $\delta(c)$ so that elasticity of k stock with respect to long-term return r is finite

Judd JpubE '85 shows that:

If workers have the same discount rate as capitalists (asymptotically) then long-run zero capital income tax result carries over

This is about inter-temporal distortions: a constant capital income tax rate τ_K produces a growing distortion overtime while Ramsey recommends to spread taxes across goods

ISSUES IN INFINITE HORIZON MODEL

1) Taxing initial wealth would be most efficient [as this would be lumpsum taxation]

2) Chamley-Judd tax is not time consistent: the government would like to renege and start taxing capital again

3) Zero-long run tax result is not robust to using progressive income taxation [Piketty, '01, Saez '02]

4) Dynastic model requires strong homogeneity assumptions (in discount rates) to generate reasonable steady states [unlikely to hold in practice]

Bottom line: Not very useful model for thinking about capital income taxation

PROGRESSIVE TAX IN ∞ HORIZON: PIKETTY '01

Dynastic utilities with inelastic labor supply

$$W = \sum_t u(c_t)/(1+\delta)^t$$

$r_t = f'(k_t)$, $w_t = f(k_t) - r_t k_t$

Distribution of wealth: a_t with density $f_t(a_t)$ so that $k_t = \int a g_t(a) da$

Golden rule capital stock k^*: $f'(k^*) = \delta$

With no taxes: In steady state: $f'(k_\infty) = \delta$ (i.e $k_\infty = k^*$), any $g_\infty(a)$ possible as long as $k^* = \int a g_\infty(a) da$

Proof: suppose $r_\infty = f'(k_\infty) > \delta \Rightarrow u'(c_{t+1})/u(c_t) = (1+\delta)/(1+r_t) < 1$ i.e., $c_{t+1} > c_t \Rightarrow$ Individuals want to shift consumption toward future $\Rightarrow$ Save more and accumulate capital indefinitely [not a steady state]

PROGRESSIVE TAX IN ∞ HORIZON: PIKETTY '01

Suppose a progressive capital income tax is introduced: $\tau_K = 0$ when $a \leq \bar{k}$ and $\tau_K = \tau > 0$ when $a \geq \bar{k}$

Assume $\bar{k} > k^*$

In the steady state:

1) Golden rule capital stock: $k_\infty = \int a g_\infty(a) da = k^*$

2) Truncated wealth distribution: $a \leq \bar{k}$ for all individuals

Proof: In steady state, all individuals must face same net-of-tax rate $r_\infty(1-\tau_K) \Rightarrow$ All individuals in same tax bracket $[0, \bar{k}]$ or $(\bar{k}, \infty)$. But $(\bar{k}, \infty)$ is impossible because $k_\infty = \int a g(a) da \geq \bar{k} > k^*$ and hence $f'(k_\infty) < \delta$

PROGRESSIVE TAX IN ∞ HORIZON: SAEZ '02

Piketty '01 shows that progressive capital income tax with exemption up to k^* equalizes wealth without affecting long-run capital stock

Seems desirable from steady-state perspective

Saez '02 shows that such progressive taxation is desirable from period 0 perspective if

$a \cdot \sigma < 1$ where a is Pareto parameter of initial wealth distribution and σ is inter-temporal elasticity of substitution $u(c_t) = [c_t^{1-1/\sigma} - 1]/[1 - 1/\sigma]$

Long-run wealth distribution will then be truncated

MAKING PROGRESS IN OPTIMAL CAPITAL INCOME THEORY

Ideal research plan:

1) Develop tax formulas that are based on sufficient statistics that can be estimated empirically [behavioral responses and distributive factors]

2) Formulas should be robust to heterogeneity in preferences [accidental, warm glow, dynastic, manipulative]

3) Predictions from theory should be somewhat aligned to actual practice [taxing only earnings and not at all capital does not fit with actual practice]

4) Progress may require to deviate from utilitarian criterion

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

NEW DYNAMIC PUBLIC FINANCE

Key ingredient is uncertainty in future ability w

2 period simple model:

(0) Everybody is identical in period 0: no work and consume c_0, period 0 utility is $u(c_0)$

(1) Ability w revealed in period 1, work l and earn $z = wl$, consume c_1, period 1 utility $u(c_1) - h(l)$

Total utility $u(c_0) + \beta[u(c_1) - h(l)]$

Rate of return r, gross return $R = 1 + r$

Discount rate $\beta < 1$

STANDARD EULER EQUATION

No govt intervention: $c_0 + c_1/R = wl/R$

Solve model backward (assume c_0 given):

Period 1: $c_1 = wl - Rc_0$, choose l to maximize $u(wl - Rc_0) - h(l)$

$\Rightarrow$ FOC $wu'(wl - Rc_0) = h'(l) \Rightarrow l^* = l(w, c_0)$

Period 0: Choose c_0 to maximize: $u(c_0) + \beta \int [u(wl^* - Rc_0) - h(l^*)] f(w) dw$

FOC for c_0 (using envelope condition for l^*)

$$u'(c_0) = \beta R \int u'(c_1) f(w) dw$$

This is called the Euler equation

MECHANISM DESIGN

Government would like to redistribute from high w to low w. Government does not observe w but can observe $c_0, c_1, z = wl$ and can set taxes as a function of c_0, c_1, z

Equivalently (using revelation principle), govt can offer menu $(c_0, c_1(w), z(w))_w$ and let individuals truthfully reveal their w

Govt program: choose menu $(c_0, c_1(w), z(w))_w$ to maximize: $SW = u(c_0) + \beta \int [u(c_1(w)) - h(z(w)/w)] f(w) dw$ st

1) Budget: $c_0 + \int c_1(w) f(w) dw / R \leq \int z(w) f(w) dw / R$

2) Incentive Compatibility (IC): individual w prefers $c_0, c_1(w), z(w)$ to any other $c_0, c_1(w'), z(w')$

INVERSE EULER EQUATION

Inverse Euler equation holds at the govt optimum:

$$\frac{1}{u'(c_0)} = \frac{1}{\beta R} \cdot \int \frac{1}{u'(c_1(w))} f(w) dw$$

Proof: small deviation in menus offered: $\Delta c_0 = -\varepsilon / u'(c_0)$ and $\Delta c_1(w) = \varepsilon / [\beta u'(c_1(w))]$

Does not affect individual utilities in any state:

$u(c_0 + \Delta c_0) + \beta u(c_1(w) + \Delta c_1(w)) = u(c_0) + \beta u(c_1(w)) + \Delta c_0 u'(c_0) + \Delta c_1(w) \beta u'(c_1(w)) = u(c_0) + \beta u(c_1(w))$

$\Rightarrow$ (IC) continues to hold and SW unchanged

$\Rightarrow$ If deviation creates a surplus (or deficit) in govt budget, then initial menu not optimal $\Rightarrow$ Deviation must be budget neutral $\Rightarrow -\varepsilon / u'(c_0) + \int \varepsilon f(w) dw / [\beta R u'(c_1(w))] = 0$

INTERTEMPORAL WEDGE

Jensen Inequality $\Rightarrow$ $K(\int x(w)dF(w)) < \int K(x(w))dF(w)$ for $K(.)$ convex

Apply this to $K(x) = 1/x$ and $x(w) = u'(c_1(w)) \Rightarrow$

$$\frac{1}{\int u'(c_1(w))f(w)dw} < \int \frac{f(w)dw}{u'(c_1(w))} = \frac{\beta R}{u'(c_0)}$$

$$\Rightarrow u'(c_0) < \beta R \int u'(c_1(w))f(w)dw$$

$\Rightarrow$ Optimal govt redistribution imposes a positive tax wedge on intertemporal choice

DECENTRALIZATION AND INTUITION

Decentralization: Optimum can be decentralized with a tax on capital income [which depends on current labor income] along with a nonlinear tax on wage income [Kocherlakota EMA'06]

Economic intuition: If high skill person works less (to imitate lower skill person), person would also like to reduce c_0 and hence save more, so tax on savings is a good way to discourage imitation

Result depends crucially on rationality in inter-temporal choices, not clear yet how applicable this is in practice

Golosov-Tsyvinski JPE'04 present decentralization results in the case of disability insurance (generalizing Diamond-Mirrlees JpubE '78): govt imposes an asset test on recipients