Lectures 9 and 10: Capital Taxation

Stefanie Stantcheva

Fall 2016
GOALS OF THIS LECTURE

(1) Understand basic concepts about capital income, savings, and wealth.

(2) Study recent new measurements and data about evolution of capital income and wealth.

(3) Theory of capital taxation: series of (quite different) models. Try to understand today, then re-read and see the different assumptions.

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

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.
MACRO FRAMEWORK

Constant return to scale aggregate production:

\[ Y = F(K, L) = rK + wL = \text{output} = \text{income} \]

\[ K = \text{capital stock (wealth)}, \ L = \text{labor input} \]

\[ r = \text{rate of return on capital}, \ w = \text{wage rate} \]

\[ rK = \text{capital income}, \ wL = \text{labor income} \]

\[ \alpha = \frac{rK}{Y} = \text{capital income share (constant \( \alpha \) when } F(K, L) = K^\alpha L^{1-\alpha} \text{ Cobb-Douglas), } \alpha \approx 30\% \]

\[ \beta = \frac{K}{Y} = \text{wealth to annual income ratio, } \beta \approx 4 - 6 \]

\[ r = \left( \frac{rK}{Y} \right) \cdot \left( \frac{Y}{K} \right) = \frac{\alpha}{\beta}, \ r = 5 - 6\% \]
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]
Figure 12: Capital shares in factor-price national income
1975-2010

Source: Piketty and Zucman (2014)
Private wealth / national income ratios, 1970-2010

Authors' computations using country national accounts. Private wealth = non-financial assets + financial assets - financial liabilities (household & non-profit sectors)

Source: Piketty and Zucman '13
Private wealth / national income ratios 1870-2010

Authors' computations using country national accounts. Private wealth = non-financial assets + financial assets - financial liabilities (household & non-profit sectors)

Source: Piketty and Zucman '13
The changing nature of national wealth, UK 1700-2010

National wealth = agricultural land + housing + other domestic capital goods + net foreign assets

Source: Piketty, Handbook chapter, 2014
National wealth = agricultural land + housing + other domestic capital goods + net foreign assets

Source: Piketty, Handbook chapter, 2014
The changing nature of national wealth, US 1770-2010 (incl. slaves)

National wealth = agricultural land + housing + other domestic capital goods + net foreign assets

Source: Piketty and Zucman '13

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

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

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”

$\Rightarrow$ Capital taxation reduces $r$ to $\bar{r} = r \cdot (1 - \tau_K) \Rightarrow$ 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 $\sim 4\times GDP \sim 60$ 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.12 For the pre-1962 period, no

---

**Figure II**

Aggregate US Household Wealth, 1913–2013

The figure depicts the level and composition of aggregate household wealth from 1913 to 2013 expressed as a percentage of national income. Housing (net of mortgages) includes owner- and tenant-occupied housing net of mortgage debt. Sole proprietorships and partnerships are business assets including sole proprietorships, farms including land and equipment, partnerships, and intellectual property products. Equities are corporate equities for both publicly traded and closely held corporations including S-corporations. Currency, deposits, and bonds are net fixed income claims including bonds, saving and current deposits, and currency, and are net of all non-mortgage debt. Pensions include individual retirement accounts, defined contribution pensions funds such as 401(k)s, funded defined benefits pensions, and life insurance reserves, but exclude unfunded defined benefit entitlements and Social Security. Pensions are typically invested in both fixed income claims and corporate equities. Source: Online Appendix Table A2.
The composition of capital income in the U.S., 1913-2013

Housing rents (net of mortgages)
Noncorporate business profits
Corporate profits
Profits & interest paid to pensions
Net interest

Source: Saez and Zucman (2014)
INDIVIDUAL 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 minus gifts given).

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

\[
W_t = \sum_{k=1}^{t} (E_k - C_k + I_k) \prod_{j=k+1}^{t} (1 + r_j)
\]
INDIVIDUAL WEALTH AND CAPITAL INCOME

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

1st term is **life-cycle** wealth, 2nd term is **inheritance** wealth

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 \( I_t \) [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

US public underestimates extent of wealth inequality and thinks the ideal wealth distribution should be a lot less unequal [Norton-Ariely ’11]
agreed that such redistribution should take the form of moving wealth from the top quintile to the bottom three quintiles. In short, although Americans tend to be relatively more favorable toward economic inequality than members of other countries (Osberg & Smeeding, 2006), Americans' consensus about the ideal distribution of wealth within the United States

Fig. 2. The actual United States wealth distribution plotted against the estimated and ideal distributions across all respondents. Because of their small percentage share of total wealth, both the “4th 20%” value (0.2%) and the “Bottom 20%” value (0.1%) are not visible in the “Actual” distribution.

Source: Norton and Ariely 2011
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]
A. Foundations top wealth vs. capitalized income shares

Panel A depicts top foundation wealth shares using balance sheet wealth (solid line) and foundations' capitalized incomes (dashed line). Since income from bonds and stocks is lumped together on foundation reporting forms 990-PF, we only capitalize dividends and interest on the one hand and rents on the other. Panel B depicts top household wealth shares using the reported wealth (solid line) and the capitalized incomes (dashed line) of SCF respondents. Wealth includes fixed income claims (savings, checking, money market, and call accounts, certificates of deposits, holdings of savings bonds, direct holding of tax-able bonds, and holdings of taxable bonds through mutual funds), corporate equi-ties (held directly and through mutual funds), business assets, rental real estate, and miscellaneous financial assets. Wealth excludes the net value of owner-occupied houses and pension wealth. Hence, the level and trend of wealth shares are not comparable with full wealth SCF estimates, discussed later. For the SCF of year $t$, wealth is measured in year $t$ but capital income is measured in year $t-1$.

Sources: Panel A: Publicly available Statistics of Income tax data, see Online Appendix Tables C11 and C13. Panel B: SCF AQ22 micro-data, see Online Appendix Table C1.
Testing the Capitalization Method Using SCF and Foundation Data

Panel A depicts top foundation wealth shares using balance sheet wealth (solid line) and foundations’ capitalized incomes (dashed line). Since income from bonds and stocks is lumped together on foundation reporting forms 990-PF, we only capitalize dividends and interest on the one hand and rents on the other. Panel B depicts top household wealth shares using the reported wealth (solid line) and the capitalized incomes (dashed line) of SCF respondents. Wealth includes fixed income claims (savings, checking, money market, and call accounts, certificates of deposits, holdings of savings bonds, direct holding of tax-exempt bonds, and holdings of taxable bonds through mutual funds), corporate equities (held directly and through mutual funds), business assets, rental real estate, and miscellaneous financial assets. Wealth excludes the net value of owner-occupied houses and pension wealth. Hence, the level and trend of wealth shares are not comparable with full wealth SCF estimates, discussed later. For the SCF of year $t$, wealth is measured in year $t$ but capital income is measured in year $t-1$.

Sources: Panel A: Publicly available Statistics of Income tax data, see Online Appendix Tables C11 and C13. Panel B: SCF AQ22 micro-data, see Online Appendix Table C1.
Panel A plots the wealth share of the top 10% in the United States from 1917 to 2012 using the capitalization method. We also report the official wealth share estimates of the top 10% from the SCF for the period 1989–2013 from Kennickell (2009b, 2011) and Bricker et al. (2014). Panel B plots the top 1% and next 9% wealth shares in the United States from 1913 to 2012. For our estimates, the unit is the family (single adult person aged 20 or more, with or without children dependents, or married couple with or without dependents). For the SCF, the unit is the household (a household can include several families) and wealth includes durables such as cars but excludes defined benefit funded pensions. Source: Online Appendix Table B1 and C4.
Panel A compares our top 0.1% wealth share estimates with top wealth share estimates from using estate tax returns (Kopczuk and Saez (2004) for 1917–2000, which we extended to 2001–2012) and the Survey of Consumer Finances (SCF). To improve comparability, starting from the SCF baseline estimates of Kennickell (2009b, 2011), we adjust the SCF series by: (1) defining fractiles relative to total families instead of households; (2) adjusting individual wealth components to match household balance sheet totals asset class by asset class; (3) adding back the Forbes 400 that are excluded by design from the SCF.

Panel B compares the top 0.1% capital income shares estimates from the SOI income tax data, the SCF, and decedents using the weights of Kopczuk and Saez (2004) (income is measured for the calendar year before death). In all three cases, we use the same definition of capital income (as the SCF reports income following the lines of the income tax return). Namely, capital income is the sum of taxable interest income, dividends, realized capital gains, profits from sole proprietorships, partnerships and S-corporations, rents, and royalties (schedule C and schedule E income). Source: Online Appendix Tables C2, C3, C4, C4b, and C8.
Panel A compares our top 0.1% wealth share estimates with top wealth share estimates from using estate tax returns (Kopczuk and Saez (2004) for 1917–2000, which we extended to 2001–2012) and the Survey of Consumer Finances (SCF). To improve comparability, starting from the SCF baseline estimates of Kennickell (2009b, 2011), we adjust the SCF series by: (1) defining fractiles relative to total families instead of households; (2) adjusting individual wealth components to match household balance sheet totals asset class by asset class; (3) adding back the Forbes 400 that are excluded by design from the SCF.

Panel B compares the top 0.1% capital income share estimates from the SOI income tax data, the SCF, and decedents using the weights of Kopczuk and Saez (2004) (income is measured for the calendar year before death). In all three cases, we use the same definition of capital income (as the SCF reports income following the lines of the income tax return). Namely, capital income is the sum of taxable interest income, dividends, realized capital gains, profits from sole proprietorships, partnerships and S-corporations, rents, and royalties (schedule C and schedule E income). Source: Online Appendix Tables C2, C3, C4, C4b, and C8.
Although the top 0.1% is a small group—it includes about 160,000 tax units with net assets above $20 million in 2012—carefully measuring its wealth is important. The public cares about the distribution of economic resources, and since wealth is highly concentrated (much more than labor income due to the dynamic processes that govern wealth accumulation), producing reliable estimates requires paying careful attention to the top. This is difficult to achieve with surveys, even the SCF (see Bricker et al. 2015 and Kennickell 2015 for recent careful evaluations), and motivates our attempt at using tax records covering all the richest families. The top 0.1% also matters from a macroeconomic perspective: it owns a sizable share of total wealth and accounts for a large fraction of its growth. From 1986 to 2012, for example, almost half of US wealth accumulation has been due to the top 0.1% alone.

A number of studies have used the income capitalization method in the past, notably King (1927), Stewart (1939), Wolff (1980), Greenwood (1983) in the United States, and Atkinson and Harrison (1978) in the United Kingdom. But these studies provide estimates for just a few years in isolation, do not use micro-

---

**Figure I**

Top 0.1% Wealth Share in the United States, 1913–2012
Third, some of the profits of partnerships and S-corporations include a labor income component, so that part of the rise of the top 0.1% share reflects a rise of top entrepreneurial rather than pure capital income. However, the concentration of pure capital income has also increased significantly. The share of dividends earned by the top 0.1% dividend-income earners rose from 35% in 1962 to 50% in 2012. The increase is even more spectacular for taxable interest, from 12% to 47%. In brief, the tax-return data are consistent with the view that capital inequality has risen enormously over the last decades. As we shall see, however, the concentration of wealth has increased less than that of taxable capital income.

**Figure III**

The Top 0.1% Taxable Capital Income Share in the United States, 1962–2012
Top Wealth Shares in the United States, 1913–2012

Panel A plots the wealth share of the top 10% in the United States from 1917 to 2012 using the capitalization method. We also report the official wealth share estimates of the top 10% from the SCF for the period 1989–2013 from Kennickell (2009b, 2011) and Bricker et al. (2014). Panel B plots the top 1% and next 9% wealth shares in the United States from 1913 to 2012. For our estimates, the unit is the family (single adult person aged 20 or more, with or without children dependents, or married couple with or without dependents). For the SCF, the unit is the household (a household can include several families) and wealth includes durables such as cars but excludes defined benefit funded pensions. Source: Online Appendix Table B1 and C4.
A. Composition of the bottom 90% wealth share

Panel A plots the wealth share of the bottom 90% and its composition from 1917 to 2012 lumping together the category of equities, fixed claim assets net of all non-mortgage debt. Panel B depicts the average real wealth of families in the bottom 90% (right y-axis) and families in the top 1% (left y-axis) from 1946 to 2012. Wealth is expressed in constant 2010 dollars, using the GDP deflator.

Source: Online Appendix Tables B3 and B5.
Wealth of the Bottom 90% of the Distribution
A. Share of wealth held by elderly households (aged 65+)

Panel A depicts the fraction of wealth held by elderly families for 3 groups: (1) the full population; (2) the bottom 90%; and (3) the top 0.1%. An elderly family is defined as a tax unit where either the primary filer or the secondary filer (for married tax units) is aged 65 or over. The series covers 1962 to 2012, years for which this information is available.

Panel B depicts the shares of total national income and total labor income accruing to top 0.1% wealth holders from 1960 to 2012. Source: Online Appendix Tables B4, B25, and B28.

FIGURE VIII

Elderly Wealth and Income Shares of Top 0.1% Wealth Holders
Panel A depicts the fraction of wealth held by elderly families for 3 groups: (1) the full population; (2) the bottom 90%; and (3) the top 0.1%. An elderly family is defined as a tax unit where either the primary filer or the secondary filer (for married tax units) is aged 65 or over. The series covers 1962 to 2012, years for which this information is available. Panel B depicts the shares of total national income and total labor income accruing to top 0.1% wealth holders from 1960 to 2012. Source: Online Appendix Tables B4, B25, and B28.
A. Saving rates by wealth class (decennial averages)

Panel A plots the synthetic saving rates (see definition in the text) for the top 1%, the top 10–1% (next 9%), and the bottom 90% averaged by decade from 1913 to 2012 (the first dot includes only 3 years 1917 to 1919, while the last dot includes only 3 years 2010 to 2012). The average private (household + corporate) saving rate was 11.4% over 1913–2013, but the rich save more as a fraction of their income, except in the 1930s when there was large dis-saving through corporations.

Panel B plots the share of wealth and income of the bottom 90% wealth holders. Income is defined so as to match (pre-tax) national income in the national accounts. If the bottom 90% saving rate had been equal to 3% every year from 1985 to 2012, then all else being equal (in particular keeping the top 10% saving constant) the bottom 90% wealth share would be 29.7% in 2012 instead of 22.8% in the data. If, in addition, the income share of the bottom 90% had remained equal to 70% (its 1970–1985 average value) then the wealth share of the bottom 90% would be 32.7% in 2012.

Source: Online Appendix Tables B1, B25, B33, and B33c.
The top 10% wealth holders own about 80% of total wealth in 1929, and 75% today.
The top decile (the top 10% highest wealth holders) owns 80-90% of total wealth in 1810-1910, and 60-65% today.

Source: Piketty and Zucman '14, handbook chapter
The top decile owns 80-90% of total wealth in 1810-1910, and 70% today.

Source: Piketty and Zucman '14, handbook chapter
The top 10% holds 80-90% of total wealth in 1810-1910, and 55-60% today.

Source: Piketty and Zucman '14, handbook chapter
CAPITAL TAXATION IN THE US


1) Corporate Income Tax (fed+state): 35% 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 $5.5M exemption (only .1% of deceased liable), tax rate is 40% above exemption (2013+)

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 VS. INHERITED WEALTH

Old view: Tobin and Modigliani: life cycle wealth accounts for the bulk of the wealth held 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 Modeling: what accounts for wealth accumulation and inequality? Is widely used life-cycle model with no bequests a good approximation?

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

Key problem is that the definition of life-cycle vs. inherited wealth is not conceptually clean (Modigliani does not capitalize inherited wealth while Kotlikoff-Summers do)
LIFE CYCLE VS. INHERITED WEALTH

Piketty-Postel-Vinay-Rosenthal EEH’14 (PPVR) propose better definition to resolve Modigliani vs. Kotlikoff-Summers controversy (see Piketty-Zucman Handbook chapter ’14)

Individual wealth accumulation:

\[ W_t = \sum_{k=1}^{t} (E_k - C_k) \cdot (1 + r)^{t-k} + \sum_{k=1}^{t} I_k \cdot (1 + r)^{t-k} \]

If \( W_t > \sum_{k=1}^{t} I_k \cdot (1 + r)^{t-k} \) then individual also saves out of labor income \( E_k \) and inherited wealth is \( \sum_{k=1}^{t} I_k \cdot (1 + r)^{t-k} \)

If \( W_t \leq \sum_{k=1}^{t} I_k \cdot (1 + r)^{t-k} \) then individual consumes part of inheritances (in addition to labor income) and inherited wealth is \( W_t \)

PPVR requires micro-data for implementation. If we assume uniform saving rate \( s \), there is a simplified formula for share of inherited wealth \( b_y / [b_y + (1 - \alpha) \cdot s] \) with \( b_y \) bequest flow/national income and \( \alpha \) capital share
LIFE CYCLE VS. INHERITED WEALTH

How do the shares of inheritance vs. life-cycle evolve over time? First measure is inheritance flow to national income

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

Inheritance share ↓ in 20th century but has ↑ recently in France (Piketty QJE’11, Piketty-Zucman ’14 handbook chapter)

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

In general \( r > n + g \) in which case inheritances play a large role in aggregate wealth and wealth concentration is going back (Western countries moving in that direction, Piketty ’14)
If we take a long-run perspective, then the twentieth-century U-shaped pattern looks even more spectacular. The inheritance flow was relatively stable around 20–25% of national income throughout the 1820–1910 period (with a slight upward trend), before being divided by a factor of about 5–6 between 1910 and the 1950s, and then multiplied by a factor of about 3–4 between the 1950s and the 2000s.

These are truly enormous historical variations, but they appear to be well founded empirically. In particular, we find similar patterns with our two fully independent estimates of the inheritance flow. The gap between our "economic flow" (computed from national wealth estimates, mortality tables, and observed age-wealth profiles) and "fiscal flow" series (computed from bequest and gift tax data) can be interpreted as a measure of tax evasion and other measurement errors. This gap appears to approximately constant over time and relatively small, so that our two series deliver fairly consistent long-run patterns.

If we use disposable income (national income minus taxes plus cash transfers) rather than national income as the denominator, then we find that the inheritance flow observed in the early twenty-first century is back to about 20%, that is, approximately the same level as that observed one century ago. This comes from the fact that disposable income was as high as 90–95% of national income.

**FIGURE I**
The inheritance flow follows a U-shaped curve in France as well as in the U.K. and Germany. It is possible that gifts are underestimated in the U.K. at the end of the period.

Source: Piketty and Zucman '14, handbook chapter
Inherited wealth represents 80-90% of total wealth in France in the 19th century; this share fell to 40%-50% during the 20th century, and is back to about 60-70% in the early 21st century.

Source: Piketty and Zucman ‘14, handbook chapter
The inheritance share in aggregate wealth accumulation follows a U-shaped curve in France and Germany (and to a more limited extent in the U.K. and Germany. It is possible that gifts are under-estimated in the U.K. at the end of the period.

Source: Piketty and Zucman ‘14, handbook chapter
TAXES IN OLG LIFE-CYCLE MODEL

\[
\text{max } U = u(c_1, l_1) + \delta u(c_2, l_2)
\]

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 $\tau_L$:

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

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 $\tau_K$:

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

$\tau_K$ distorts only inter-temporal consumption choice

Budget with comprehensive income tax $\tau$:

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

$\tau$ distorts both labor-leisure and inter-temporal consumption choices

$\tau$ imposes “double” tax: (1) tax on earnings, (2) tax on savings
EFFECT OF $r$ ON SAVINGS

Assume that labor supply is fixed. Draw graph. 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 
\[ 1 + t_c = \frac{1}{(1 - \tau_L)} \] 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 \( t_c \): 
\[ c_1^c = (w_1 - s_c)/(1 + t_c), \quad c_2^c = (1 + r)s_c/(1 + t_c) \]

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

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 with labor tax $\tau_L$ and 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
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]
Complex problem with many sub-literatures: Banks and Diamond Mirrlees Review ’09 provide recent survey

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

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


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

RAMSEY TAX IN LIFE-CYCLE MODEL


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

Individual maximization problem:

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

\[ \text{st} \quad c_1 + c_2 / (1 + r(1 - \tau_K)) = w(1 - \tau_L) \]

where \( q = 1 / (1 + r(1 - \tau_K)) \) and \( p = 1 / (1 + r) \) are post-tax and pre-tax prices of \( c_2 \)
Optimal tax rates can be obtained by solving the standard Ramsey problem:

$$\max_{\tau_K, \tau_L} V(q, w(1 - \tau_L)) \quad \text{st} \quad wL \tau_L + (q - p)c_2 \geq g \quad (\lambda)$$

where $g$ is exogenous tax revenue requirement

Can apply the results from the 3 good Ramsey model

Derive FOC for $\tau_K$ and $\tau_L$

Can express them in terms of compensated elasticities
Combining the two FOC to get rid of $\lambda$, you get:

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

where $\sigma_{LL} = \left(\frac{w(1 - \tau L)}{l}\right)\frac{\partial l^c}{\partial (w(1 - \tau L))} > 0$ is the compensated elasticity of labor supply wrt net wage rate

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

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

$\sigma_{2L} = \left(\frac{w(1 - \tau L)}{c_2}\right)\frac{\partial c_2^c}{\partial (w(1 - \tau L))}$

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] ⇒ 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 \( \sigma_{LL} \ll |\sigma_{22}| \) then \( \tau_K \) should be small relative to \( \tau_L \)

**Key lesson:** What matters is the relative size of elasticities, not the number of distortions
Feldstein JPE’78 makes famous theoretical argument why $\sigma_{22}$ can be large even if $\varepsilon_{sq}^u = (q/s)\partial s/\partial q$ [uncompensated savings elasticity] is zero:

Budget $c_1 + qc_2 = w(1 - \tau_L)l + y$

Slutsky equation [y is endowment =0 in equilibrium]:

$$\partial c_2^\xi/\partial q = \partial c_2/\partial q + c_2\partial c_2/\partial y \Rightarrow$$

$$\sigma_{22} = \varepsilon_{2q}^u + q\partial c_2/\partial y$$

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

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

$c_1 + qc_2 = w(1 - \tau_L)l + y \Rightarrow$

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

$$\sigma_{22} \approx \varepsilon_{sq}^u - \partial c_1/\partial y \approx -\partial c_1/\partial y \leq -0.75 [as\ saving\ rate\ modest]$$
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]

In that case, optimal \( \tau_K, \tau_L \) given by same static Ramsey rule

Problems of dynamic efficiency (optimal \( K \) stock) and efficiency within a generation (\( \tau_L, \tau_K \)) are orthogonal
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 rates 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
1) No redistributive concerns: Can extend model to the multi-person case ⇒ Higher rate $\tau_K$ 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.
Stop here: need to study a very important result: Atkinson-Stiglitz theorem
Famous Atkinson-Stiglitz JpubE’ 76 shows that

$$\max_{t,T(.)} SWF = \max_{t=0,T(.)} SWF$$

(i.e., commodity taxes not useful) under two assumptions on utility functions $u^h(c_1, .., c_K, z)$

1) Weak separability between $(c_1, .., c_K)$ and $z$ in utility

2) Homogeneity across individuals in the sub-utility of consumption $v(c_1, .., c_K)$ [does not vary with $h$]

$$(1) \text{ and } (2): \quad u^h(c_1, .., c_K, z) = U^h(v(c_1, .., c_K), z)$$

Original proof was based on optimum conditions, new straightforward proof by Laroque EL ’05, and Kaplow JpubE ’06.
Let $V(y, p + t) = \max_c v(c_1, \ldots, c_K)$ st $(p + t) \cdot c \leq y$ be the indirect utility of consumption $c$ [common to all individuals].

Start with $(T(.), t)$. Let $c(t)$ be consumer choice.

Replace $(T(.), t)$ with $(\tilde{T}(.), t = 0)$ where $\tilde{T}(z)$ such that $V(z - T(z), p + t) = V(z - \tilde{T}(z), p) \Rightarrow$ Utility $U^h(V, z)$ and labor supply choices $z$ unchanged for all individuals.

Attaining $V(z - \tilde{T}(z), p)$ at price $p$ costs at least $z - \tilde{T}(z)$.

Consumer also attains $V(z - \tilde{T}(z), p) = V(z - T(z), p + t)$ when choosing $c(t) \Rightarrow z - \tilde{T}(z) \leq p \cdot c(t) = z - T(z) - t \cdot c(t)$

$\Rightarrow \tilde{T}(z) \geq T(z) + t \cdot c(t)$: the government collects more taxes with $(\tilde{T}(.), t = 0)$.
ATKINSON-STIGLITZ INTUITION

With separability and homogeneity, conditional on earnings $z$, consumption choices $c = (c_1, \ldots, c_K)$ do not provide any information on ability

$\Rightarrow$ Differentiated commodity taxes $t_1, \ldots, t_K$ create a tax distortion with no benefit $\Rightarrow$ Better to do all the redistribution with the individual income tax

Note: With weaker linear income taxation tool (Diamond-Mirrlees AER ’71, Diamond JpubE ’75), need $v(c_1, \ldots, c_K)$ homothetic (linear Engel curves, Deaton EMA ’81) to obtain no commodity tax result

[Unless Engel curves are linear, commodity taxation can be useful to “non-linearize” the tax system]
Generalization of Atkinson-Stiglitz to Heterogeneous Tastes – Saez (2002)

Can we generalize AS to case with heterogeneous consumption preferences?

Individuals indexed by $h$, utility $U(c, z)$ with $c = (c_1, ..., c_K)$.

Nonlinear income tax $T(z)$.

Pre-tax prices: $p$, post-tax price: $q = p + t$.

Budget constraint $q \cdot c \leq z - T(z)$.

Demands: $c^h(q, R, z)$, labor supply $z^h(q, T)$, indirect utility $v^h(q, R, z)$.

Suppose that $T(z)$ is optimally chosen at zero commodity taxation $p = q$ to max

$$W = \sum_h \alpha^h v^h(p, z^h - T(z^h), z^h) \quad \text{s.t.} \quad \sum_h T(z^h) \geq E$$

Marginal social welfare weight $g^h = \alpha^h v^h_R / \lambda$. 
Can commodity taxation improve welfare?

Imagine $dt_1$.

Mechanical revenue effect: $dM_1 = \sum_h c^h_1 dt_1 = C_1 dt_1$.

Welfare effect (envelope theorem): $dU_1 = -\sum_h g^h c^h_1 dt_1$.

Behavioral labor supply response: $dB_1 = -\sum_h T'(z^h) dz^h_{t_1}$ with $dz^h_{t_1} = dt_1 \frac{\partial z^h}{\partial q_1}$.

Why no behavioral response on revenue from changes in consumption?

If no commodity tax introduction can increase welfare, need $dW = dM_1 + dU_1 + dB_1 = 0$. 
Can commodity taxation improve welfare? (II)

Find a small income tax reform that “mimics” commodity tax change: $dT(z) = C_1(z)dt_1$.

This reform has zero first order welfare impact, why?

Mechanical Revenue effect:
$$dM_T = \sum_h dT(z^h) = \sum_h C_1(z^h)dt_1 = C_1 dt_1 = dM_1 \text{ (why?)}$$

Welfare effect: $dU_T = -\sum_h g^h C_1(z^h)dt_1$

Behavioral effect: $dB_t = -\sum_h T'(z^h)dz^h_T$.

Subtract one from other (using that $dW_T = 0$).
Can commodity taxation improve welfare? (III)

\[
\frac{dW}{dt_1} = - \sum_h g^h [c^h_1 - C_1(z^h)] + \sum_h T'(z^h) \left[ \frac{dz^h_T}{dT(z^h)} \cdot \frac{dT(z^h)}{dt_1} - \frac{dz^h_1}{dt_1} \right]
\]

- Pure welfare effect is zero if conditional on \( z, g^h \) and \( c^h \) are uncorrelated.

- What does this mean? Is this reasonable? (recall weights are "generalized" since \( \alpha^h \) depends on \( h \) directly).

- Young/old? medical expenses?

- Always satisfied in "standard AS" assumptions.
Behavioral Effects under Commodity and Income Taxation

Since $T'(z^h) \geq 0$ (remember), increasing $dt_1 > 0$ more efficient than equivalent income tax increase if labor supply increase from commodity tax change larger than that of income tax change.

When is this the case? Can show that:

$$E[dz^h_{t1}] = -dt_1 \left( E \left[ \frac{z^h_c}{1 + T''(z)z^h_c} \frac{dc^h_1}{dz} \right] + E \left[ \frac{z^h_R}{1 + T''(z)z^h_c} c^h_1 \right] \right)$$

$$E[dz^h_T] = -dt_1 \left( E \left[ \frac{z^h_c}{1 + T''(z)z^h_c} \frac{dC_1(z)}{dz} \right] + E \left[ \frac{z^h_R}{1 + T''(z)z^h_c} C_1(z) \right] \right)$$

Need $E[dz^h_{t1}] = E[dz^h_T]$ for no commodity taxation.
Assumptions needed for behavioral effects to be the same under Commodity and Income Taxation

Assumption 2: Conditional on $z$, behavioral responses $z^h_c$ and $z^h_R$ independent of consumption patterns $c^h_1$ and $\frac{dc^h_1}{dz}$.

Do you think this holds?

Assumption 3: For any income level, $E\left( \frac{dc^h_1}{dz} | z^h = z \right) = \frac{dC_1(z)}{dz}$.

This is the key assumption. What does it say? Why is this not mechanically true?

$$\frac{dC_1(z)}{dz} = \lim_{dz \to 0} \frac{E(c^h_1|z^h=z+dz)-E(c^h_1|z^h=z)}{dz}$$

is cross-sectional variation in consumption of good 1 when income changes.

What is $E\left( \frac{dc^h_1}{dz} | z^h = z \right)$?
Assumptions needed for behavioral effects to be the same under Commodity and Income Taxation

Imagine 2 groups:

Group A: $z^h = z$. Consume $C_1(z)$ on average.

Group B: $z^h = z - dz$. Consumes on average $dc_1 = \frac{dC_1(z)}{dz}$ less of good 1.

Group A': Individuals from group A who are forced to reduce their income to $z - dz$. Reduce their consumption relative to group A by $dc_1' = E \left( \frac{dc_1^h}{dz} \mid z^h = z \right) dz$.

Assumption 3 says: Group B = Group A' for consumption of good 1.

Always true in AS since consumption only depends on after tax income with separability $C_1(z) = c_1^h(z)$.

Why would Group A' not have same consumption of good 1 as Group B?
Thought experiment we just did was: force high earners to work less and earn only as much as low earners: if high earners consume more of good $k$ than low earners, taxing good 1 is desirable.

1) High earners are “different” (since if left to chose, chose to work more. If they have a relatively higher/lower taste for good 1 (independently of income), tax more/less good 1. [indirect tagging] Cigarettes? Fancy wine? How would you see this empirically?

2) High earners now have more leisure. If Good 1 positively related to leisure (consumption of 1 increases when leisure increases keeping after-tax income constant), tax it! [tax on holiday trips, subsidy on work related expenses such as child care]

In general Atkison-Stiglitz assumption is a good starting place for most goods ⇒ Zero-rating on some goods under VAT for redistribution is inefficient and administratively burdensome [Mirrlees review]
Life-Cycle model: Atkinson-Stiglitz JpubE ’76

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).
Life-cycle model: linear labor income tax

Suppose the government can only use linear earnings tax: $wl \cdot (1 - \tau_L) + E$

If sub-utility $v(c_1, c_2)$ is also homothetic of degree one [i.e., $v(\lambda c_1, \lambda c_2) = \lambda v(c_1, c_2)$ for all $\lambda$] then $\tau_K = 0$ is again optimal [linear tax counter-part of Atkinson-Stiglitz, see Deaton, 1979]

In the general case $V^h(c_1, c_2, l)$, optimal $\tau_K$ is not always zero
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 donors who worked hard to pass on wealth to children

⇒ Double welfare effect: inheritance tax hurts donor (if donor altruistic to donee) and donee (which receives less) [Kaplow, ’01]
MODELS OF BEQUESTS AND GIFTS

Individuals receive inheritances and inter-vivos gifts. Those arise because of:

1) Accidental bequests: uncertain life-time with imperfect annuity markets \(\Rightarrow\) people die with positive wealth

2) Altruistic bequests/gifts: people care about utility of children [Barro-Becker dynastic model is most famous example]

3) Warm glow bequests/gifts: people enjoy making transfers to children (the bequest itself enters utility).

4) Manipulative / social norms bequests: bequests used to extract services from heirs or pressure from heirs to leave bequests [equal estate division Wilhelm AER’96, spousal pensions Aura JpubE’05]
Estate Taxation in the United States

Estate federal tax imposes a tax on estates above $5.5M exemption (only about .1% of deceased liable), tax rate is 40% above exemption (2013+)

Charitable and spousal giving are fully exempt from the tax

E.g.: if Bill Gates / Warren Buffet give all their wealth to charity, they won’t pay estate tax

Support for estate tax is pretty weak (“death tax”) but public does not know that estate tax affects only richest

Support for estate tax increase shots up from 17% to 53% when survey respondents are informed that only richest pay it (Kuziemko-Norton-Saez-Stantcheva ’13 do an online Mturk survey experiment)
Besides the income tax, the government can also level the playing field with the federal estate tax.

The Federal Estate Tax (also known as the Death Tax) applies when a deceased person leaves more than $5 million in wealth to his or her heirs. Wealth left to a spouse or charitable organizations is exempt from estate tax.

Only 1 person out of 1000 is wealthy enough to face the estate tax.

Average Americans do not have anything close to $5 million in wealth, so the estate tax does not affect them and they can pass on their property to their children tax-free.

Eliminating the estate tax would allow the very richest families to pass down all of their wealth to their children tax-free. Hence, children of rich people would also start off very rich themselves.

Increasing the estate tax is a way to level the playing field between the children of wealthy parents and children of middle-class parents.
Taxation of Inheritances: Behavioral Responses

Potential behavioral response effects of inheritance tax:

(1) reduces wealth accumulation of altruistic donors (and hence tax base) [no very good empirical evidence, Slemrod-Kopczuk 01]

(2) reduces labor supply of altruistic donors (less motivated to work if cannot pass wealth to kids) [no good evidence]

(3) induces donees to work more through income effects (Carnegie effect, decent evidence from 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) bequests in the utility, (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 defined benefit private pensions] are in general mandatory 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) ⇒ 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 ⇒ Implicit insurance against risk of living too long

Kopczuk–Lupton REStud’07 shows that only 1/2 of people accumulate wealth for bequest motives.
Bequests in the Utility: Warm Glow Or Altruistic

\[ u(c) - h(l) + \delta v(b) \] 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 bequest tax at rate \( \tau_B \): \( c + b/[(1 + r)(1 - \tau_B)] = wl - T_L(wl) \)

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 \( [\tau_B = 0] \) and that nonlinear tax on \( wl \) is enough for redistribution.
Bequests in the Utility: Warm Glow Or Altruistic

Suppose now that $b$ is given to a heir who derives utility $v^{\text{heir}}(b) \Rightarrow b$ creates a positive externality (to donee) and hence should be subsidized $\Rightarrow \tau_B < 0$ 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:

$$SWF = \int [u(c) - h(l) + \delta v(b) + v^{\text{heir}}(b)] f(w) dw$$

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

$\Rightarrow$ Optimal subsidy rate is smaller for large estates $\Rightarrow$ progressive estate subsidy
Atkinson-Stiglitz applies when sole source of lifetime income is labor:

\[ c + \frac{b(\text{left})}{1 + r} = w l - T(w l) \quad (w = \text{productivity}, \ l = \text{labor supply}) \]

In GE, bequests provide an additional source of life-income:

\[ c + \frac{b(\text{left})}{1 + r} = w l - T(w l) + b(\text{received}) \]

⇒ conditional on \( w l \), high \( b(\text{left}) \) is a signal of high \( b(\text{received}) \) ⇒ \( b(\text{left}) \) should be taxed even with optimal \( T(w l) \)

⇒ Two-dim. inequality requires two-dim. tax policy tool

Extreme example: no heterogeneity in productivity \( w \) but pure heterogeneity in bequests motives ⇒ bequest taxation is desirable for redistribution
Measure one of individuals, who are both bequests receivers and bequest leavers (in ergodic general equilibrium)

Linear tax $\tau_B$ on bequests funds lumpsum grant $E$

Life-time budget constraint: $c_i + b_i = R(1 - \tau_B)b_i^r + y_{Li} + E$

with $c_i$ consumption, $b_i$ bequests left, $y_{Li}$ inelastic labor income, $b_i^r$ pre-tax bequests received, $R = 1 + r$ generational rate of return on bequests

Individual $i$ has utility $V_i(c, b)$ with $b = R(1 - \tau_B)b$ net-of-tax bequests left and solves

$$\max_{b_i} V_i(y_{Li} + E + R(1 - \tau_B)b_i^r - b_i, Rb_i(1 - \tau_B)) \Rightarrow V_{c_i} = R(1 - \tau_B)V_{b_i}^i$$
Government budget constraint is $E = \tau_B b$ with $b$ aggregate (=average) bequests. Govt solves:

$$\max_{\tau_B} \int \omega_i V^i(y_{Li} + \tau_B b + R(1 - \tau_B)b^r_i - b_i, Rb_i(1 - \tau_B))$$

with $\omega_i \geq 0$ Pareto weights

Meritocratic Rawlsian criterion: maximize welfare of those receiving no inheritances with uniform social marginal welfare weight $\omega_i V^i_c$ among zero-receivers

(e.g., people not responsible for $b^r_i$ but responsible for $y_{Li}$) ⇒

**Optimal inheritance tax rate:** $\tau_B = \frac{1 - \bar{b}}{1 + e_B}$

with $e_B = \frac{1 - \tau_B}{b} \frac{db}{d(1 - \tau_B)}$ elasticity of aggregate bequests and $\bar{b} = \frac{E[b_i|b^r_i=0]}{b}$ relative bequest left by zero-receivers
\[
SWF = \int_i \omega_i V^i_i (y_{Li} + \tau_B b_i - b_i, Rb_i (1 - \tau_B))
\]

[NB: removed term \( R(1 - \tau_B) b_i^r \) because \( \omega_i = 0 \) when \( b_i^r = 0 \)]

\[
0 = \frac{dSWF}{d\tau_B} = \int_i \omega_i \cdot \left( V^i_c \left[ b - \tau_B \frac{db}{d(1 - \tau_B)} \right] - Rb_i V^i_b \right) \Rightarrow
\]

\[
0 = \int_i \omega_i \cdot \left( V^i_c \cdot b \left[ 1 - \frac{\tau_B}{1 - \tau_B} e_B \right] - \frac{b_i}{1 - \tau_B} V^i_c \right) \Rightarrow
\]

\[
0 = b \left[ 1 - \frac{\tau_B}{1 - \tau_B} e_B \right] - \frac{1}{1 - \tau_B} \cdot \frac{\int_i \omega_i V^i_c \cdot b_i}{\int_i \omega_i V^i_c} \Rightarrow
\]

as \( \omega_i V^i_c \equiv 0 \) for \( b_i^r > 0 \) and \( \omega_i V^i_c \equiv 1 \) for \( b_i^r = 0 \) \( \Rightarrow \)

\[
0 = 1 - \tau_B - \tau_B \cdot e_B - \frac{E[b_i | b_i^r = 0]}{b} \Rightarrow \tau_B = \frac{1 - \bar{b}}{1 + e_B}
\]
Optimal inheritance tax rate: \( \tau_B = \frac{1 - \bar{b}}{1 + e_B} \)

1) Optimal \( \tau_B < \frac{1}{1 + e_B} \) revenue maximizing rate because zero-receivers care about bequests they leave

2) \( \tau_B = 0 \) if \( \bar{b} = 1 \) (i.e, zero-receivers leave as much bequest as average)

3) If bequests are quantitatively important, highly concentrated, and low wealth mobility then \( \bar{b} << 1 \)

4) Empirically \( e_B \) small (Kopczuk-Slemrod '01) but poorly known, \( \bar{b} = 2/3 \) in US (SCF data) but poorly measured

5) Formula can be extended to other social criteria, elastic labor supply, wealth loving preferences, altruistic preferences [see Piketty-Saez ECMA'13]
Optimal Capital Stock and Modified Golden Rule

**Modified Golden Rule:** \( r = \delta + \gamma \cdot g \)

with \( r = F_K(K, L) = f'(k) \) rate of return, \( \delta \) discount rate, \( \gamma \) curvature of \( u'(c) = c^{-\gamma} \), \( g \) growth rate (per capita).

Proof: $1$ extra in period \( t \) gives social welfare \( u'(c_t) \)

\[
\frac{(1+r)u'(c_{t+1})}{1+\delta} = \frac{(1+r)u'(c_t)}{1+\delta} \frac{u'(c_{t+1})}{u'(c_t)} = \frac{1+r}{(1+\delta)(1+g)^\gamma} u'(c_t) \Rightarrow
\]

\[ 1 + r = (1 + \delta)(1 + g)^\gamma \]

This is equivalent to \( r = \delta + \gamma \cdot g \) when the period is small. QED.

Normatively \( \delta = 0 \) seems justified. Small capital stock and \( r > g \) desirable if \( \gamma \) is high [controversy Stern vs. Nordhaus]
Optimal Capital Stock and Modified Golden Rule

Modified Golden Rule: \( r = \delta + \gamma \cdot g \)

Bequest and capital taxes affect capital stock

However, if govt can use debt, it can control capital stock

If debt used to set optimal capital stock at the Modified Golden Rule) then effects of taxes on \( K \) stock can be ignored \( \Rightarrow \) Optimal \( K \) stock and optimal redistribution are orthogonal

If \( K \) stock is not at Modified Golden Rule, then optimal \( K \) tax formulas include a corrective term

[see King 1980 and Atkinson-Sandmo 1980 in OLG life-cycle model, Piketty-Saez ECMA’13 for models with bequests]

In practice: no reason for MGR to hold, govts do not actively target \( K \) stock
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

⇒ Bequest becomes one additional form of labor income for donee and one consumption good for donor

⇒ 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 ⇒ parents are made better-off by the estate tax ⇒ 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 retiring worker decides to get a single annuity or couple annuity: reform ↑ sharply couple annuities choice

Equal division of estates [Wilhelm AER’96, Light-McGarry ’04]: estates are very often divided equally but gifts are not
DYNASTIC MODEL OR INFINITE HORIZON

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

\[
V_0 = \sum_{t \geq 0} \frac{u(c_t, l_t)}{(1 + \delta)^t}
\]

subject to \( \sum_{t \geq 0} \frac{c_t}{(1 + r)^t} \leq \sum_{t \geq 0} \frac{w_t l_t}{(1 + r)^t} \)

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 (and conversely)

2) Temporary tax cut debt financed [fiscal stimulus] should have no impact on consumption but actually do
INFINITE HORIZON MODEL: CHAMLEY-JUDD

Govt can collect taxes using linear labor income tax or capital income taxes that vary period by period $\tau_L^t, \tau_K^t$.

Goal of the government is to maximize utility of the dynasty

$$V_0 = \sum_{t} u(c_t, l_t)/(1 + \delta)^t \text{ st } \sum_{t} q_t c_t \leq \sum_{t} q_t w_t (1 - \tau_L^t) l_t + A_0 \quad (\lambda)$$

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

With constant tax rate $\tau_K$ 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:

Two equivalent ways to understand this result:

(1) A constant tax on capital income creates an exponentially growing distortion which is inefficient

(2) The PDV of the capital income tax base is infinitely elastic with respect to an increase in $\tau_K$ in the distant future [Piketty-Saez '13]

Intuition: $u_c(c_{t+1})/u_c(c_t) = (1 + \delta)/(1 + r(1 - \tau_K)) \Rightarrow$ savings decisions infinitely elastic to $r(1 - \tau_K) - \delta$

If $r(1 - \tau_K) > \delta$, accumulate forever. If $r(1 - \tau_K) < \delta$, get in debt as much as possible.
ISSUES IN INFINITE HORIZON MODEL

1) Taxing initial wealth is most efficient [as this is lumpsum taxation] ⇒ solutions typically bang-bang: tax capital as much as possible early, then zero

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 JpubE’13]

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

5) Introducing stochastic shocks in labor/preferences in dynastic model leads to finite elasticities (and reasonable optimal tax rates) [Piketty-Saez ECMA’13]
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

⇒ 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.

⇒ Transitional dynamics instantaneous ⇒ 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 ⇒ Same optimal K tax formulas apply, with *appropriately defined elasticity of the tax base.*

Qualitatively: Lessons and intuitions from simpler model still valid.

Quantitatively: Sluggish adjustments reflected in elasticity.

The faster K adjustments, the closer to simpler model.

3) Numerically explore optimal taxation using U.S. IRS data.
A Simpler Model of Capital Taxation

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

Heterogeneous discount rate \( \delta_i \) (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:

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

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

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

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

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 \( k_i(t) \): \[ a_i'(k_i(t)) + \lambda_i(t) \cdot r \cdot (1 - T_K') = -\frac{d\lambda_i(t)}{dt} + \delta_i \cdot \lambda_i(t) \]

\[ \Rightarrow a_i'(k_i(t)) = \delta_i - \bar{r} \quad \text{where} \quad \bar{r} = r \cdot (1 - T_K') \]
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 \( k_{\text{init}}^i \) to \( k_i(t) \) (consumption quantum Dirac jump):

\[
U_i = \underbrace{rk_i(t) - T_K(rk_i(t)) + z_i(t)}_{c_i(t)} + a_i(k_i(t)) + \delta_i \cdot (k_{\text{init}}^i - k_i(t))
\]

Dynamic model equivalent to a static model:

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

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

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

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: $g_i = \text{value of } \$1 \text{ extra given to } i \ (\int_i g_i = 1)$.

(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 $\tau_K$

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

Revenues $\tau_K k^m(\bar{r})$ 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 $dSWF / d\tau_K = 0$. 
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:

- $\bar{g}_K = 1$ (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: $\tau_K = 1/(1 + a_{K}^{\text{top}} \cdot e_{K}^{\text{top}})$ with $a_{K}^{\text{top}}$ 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) \( \bar{G}_K(rk) \equiv \int_{\{i: rki \geq rk\}} \frac{g_i \, di}{P(rki \geq rk) \int_j g_i \, dj} \) 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)

⇒ 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 $k_i$ → $\tau_K = 0$. 
(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 ($k_i^{\text{init}}$) – 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.

\( \tilde{G}_K(rk) \) 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 $h_i(z)$.

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

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

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

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

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

\[ v_i(z_i) = 1 - T'_L(z_i), \quad v_{ik}(k_i, z_i) = \delta_i - \bar{r}, \quad c_i = \bar{r}k_i + z_i - T_L(z_i) \]

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 \( \bar{g}_K = \int_i k_i g_i \), \( e_Z, (1 - \tau_K) = \frac{dz^m}{d(1 - \tau_K)} \frac{(1 - \tau_K)}{z^m} \)
Comprehensive nonlinear income taxation \( T(rk + z) \)

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

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

Standard Mirrlees’ formula applies to comprehensive income tax problem

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

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

\( \alpha_Y(y) \) 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:

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

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 ≠ elasticities (housing vs. financial assets)

Different social judgments or distributional characteristics $g_K^j$.

Formulas hold asset by asset, determined by: $g_K^j$, $e_K^j$, and cross-elasticities $e_{Ks, (1-\tau_K^j)}$.

$$\tau_K^j = \frac{1 - g_K^j}{1 - g_K^j + e_K^j}$$

$$g_K^j = \frac{\int_i g_i \cdot k_i^j}{\int_i k_i^j}, \quad e_K^j = \frac{\bar{r}_j}{k_{m,j}} \cdot \frac{dk_{m,j}}{d\bar{r}_j} > 0, \quad e_{Ks, (1-\tau_K^j)} = \frac{\bar{r}_j}{k_{m,s}} \cdot \frac{dk_{m,s}}{d\bar{r}_j}$$
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 $e_{K^s,(1-\tau^j_K)}$.

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

$$
\bar{g}_K^j = \frac{\int_i g_i \cdot k_i^j}{\int_i k_i^j}, \quad e_K^j = \frac{\bar{r}_j}{k^{m,j}} \cdot \frac{dk^{m,j}}{d\bar{r}_j} > 0, \quad e_{K^s,(1-\tau^K_j)} = \frac{\bar{r}_j}{k^{m,s}} \cdot \frac{dk^{m,s}}{d\bar{r}_j}
$$
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.”
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 \((z_i, rK_i)\) 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 $T'_L(z)$

![Graph showing the marginal tax rates for different values of $e_L$. The $x$-axis represents labor income in $\$100,000$ intervals, and the $y$-axis represents marginal tax rates. There are three curves for $e_L = 1$, $e_L = 0.5$, and $e_L = 0.25$. Each curve shows how the marginal tax rate changes with labor income.]
Optimal Capital Income Tax Rate $T'_K(rk)$
Optimal Tax Rate on Comprehensive Income $T'_Y(y)$

<table>
<thead>
<tr>
<th>Total Income</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$200,000</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>$400,000</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>$600,000</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>$800,000</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>$1,000,000</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
</tbody>
</table>

$e_Y = 1$

$e_Y = 0.5$

$e_Y = 0.25$
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 \( z \)

\[ \Rightarrow \text{consumption smoothing} \Rightarrow \text{sluggish transitional dynamics (a sum of anticipatory and build-up effects).} \]

Convergence to steady state no longer instantaneous:

\[ \frac{u_i}{u_{ic}} = \delta_i - \bar{r}, \quad u_{ic} \cdot (1 - T'_L) = -u_{iz} \quad \text{and} \quad c = rk + z - T(rk, z). \]

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 $\tau_K$ 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:

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

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:

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

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

Optimal linear capital income tax in steady state:

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

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:

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

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:
\[ SWF_d = \int \omega_i \cdot V_i(\{c_i(t), k_i(t), z_i(t)\}_{t \geq 0}) \]

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 $e_K < \infty$.

$e^\text{ante}_K$ anticipation elasticity:

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

Reasonable?

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

$e^\text{post}_K$ adjustment to reform: sluggish in all models, except with no transitional dynamics (linear utility).
REFERENCES


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


