Capital Taxes and Intro to Uncertainty

Christine Braun

University of Warwick EC9A2

Overview

Last lecture: Neoclassical growth model with labor choice

- Agents choose consumption, savings, and labor
- Firms choose capital demand and labor demand
- Labor choice creates feedback loop

Today:

- Part 1: Example policy analysis, capital taxes
- Part 2: Introduction to uncertainty

Part 1: Capital Taxes

Why Study Capital Taxation?

Policy relevance:

- Corporate income taxes
- Capital gains taxes
- Property taxes on capital
- Investment tax credits (negative taxes)

Key questions:

- How do capital taxes affect long-run capital accumulation?
- What happens to labor supply when capital is taxed?
- How do the effects depend on preference parameters?
- What are the welfare costs of capital taxation?

Basic Framework - no taxes

Representative agent utility:

$$U = \sum_{t=0}^{\infty} \beta^t u(c_t, 1 - L_t)$$

Production function:

$$Y_t = F(K_t, L_t)$$

Capital accumulation:

$$K_{t+1} = (1 - \delta)K_t + I_t$$

Resource constraint:

$$Y_t = c_t + I_t$$

Firms problem:

$$\max_{K_t^d, L_t^d} \left\{ F(K_t^d, L_t^d) - r_t K_t^d - w_t L_t^d \right\}$$

Capital Tax Policy Setup

Government imposes tax rate $\tau_k \in [0, 1)$ on capital income.

- collects $T_t = \tau_k r_t K_T$ each period

Assume the government gives T_t back to agents in a lump sum transfer

- no resources wasted
- isolates the pure **incentive effects** of capital taxes

Alternatively, if the government "consumed" T_t each period

- **incentive effects**: less valuable to save
- **fiscal effects**: agents have less overall to spend on c_t and k_{t+1}

Capital Tax Policy

Agent receives after-tax return:

$$(1-\tau_K)r_t \text{ per unit of capital}$$

Firm still pays pre-tax return:

$$r_t = F_K(K_t, L_t)$$

Government budget balance:(lump-sum transfer to agent)

$$T_t = \tau_K r_t K_t$$

Modified budget constraint:

$$K_{t+1} = (1-\delta)K_t + (1-\tau_K)r_t K_t + w_t L_t - c_t + T_t$$

Substituting government budget, gives the **resource constraint**:

$$K_{t+1} = (1-\delta)K_t + r_t K_t + w_t L_t - c_t$$

Key insight: Resource constraint unchanged, model will tell us the incentive effects of capital taxes

Agent's Dynamic Problem

State variable: K_t (capital stock at beginning of period t)
Control variables: c_t (consumption), ℓ_t (labor supply)

Value Function:

$$V(k, K) = \max_{c,\ell,k'} \left\{ u(c, 1-\ell) + \beta V(k', K') \right\}$$

subject to:

$$\begin{aligned} k' &= (1-\delta)k + (1-\tau)rk + w\ell - c + T \\ \ell &\in [0,1], \quad k' \geq 0, \quad c \geq 0 \\ K' &= G(K) \end{aligned}$$

Note: Firm's problem unchanged.

Recursive Competitive Equilibrium with Capital Tax

A **Recursive Competitive Equilibrium** in which the government imposes capital taxes to collect T consists of:

1. Value function: $V(k, K)$
2. Policy functions: $g_c(k, K)$, $c_\ell(k, K)$, and $g_k(k, K)$
3. Factor price functions: $r(K, L)$, $w(K, L)$
4. Aggregate law of motion $G(K)$
5. Government policy: Tax rate $\tau_K(K)$

Such that:
(i) Agent optimization: $V(k, K)$, $g_c(k, K)$, $c_\ell(k, K)$, and $g_k(k, K)$ solve the household problem
(ii) Firm optimization: $r(K, L) = F_K(K, L)$, $w(K, L) = F_L(K, L)$ (market clearing)
(iii) Consistency: $K' = g_K(k, K) = G(K)$ and $L = g_\ell(k, K)$
(iv) Government budget balance: $T = \tau_K(K)r(K, L)K$

In our case: $\tau_K(K) = \tau_K$, flat tax on capital

Optimality Conditions

With capital tax, agent's FOCs become:

$$u_c(c, 1-\ell) = \beta u_c(c', 1-\ell')[(1-\tau_K) r(K', L') + 1 - \delta]$$
$$\frac{u_{1-\ell}(c, 1-\ell)}{u_c(c, 1-\ell')} = w(K, L)$$

Key changes:

- **Intertemporal condition**: After-tax return $(1-\tau_K) r(K', L')$ instead of $r(K', L')$
- **Intratemporal condition**: Unchanged (no tax on labor income)

Key point: Transfer T doesn't appear in the FOC! Only the after-tax return $(1-\tau_K) r(K', L')$ matters for incentives.

Factor prices still determined by firms: unchanged

$$r(K, L) = F_K(K, L), \quad w(K, L) = F_L(K, L)$$

Steady State Effects of Capital Tax

New steady state condition: (from Euler equation)

$$1 = \beta[(1 - \tau_k)F_K(K^*, L^*) + 1 - \delta]$$

Rearranging:

$$F_K(K^*, L^*) = \frac{1}{1 - \tau_K}\left[\frac{1}{\beta} - 1 + \delta\right]$$

Since $\tau_K > 0$:

$$\frac{1}{1 - \tau_K} > 1 \Rightarrow F_K(K^*, L^*) \text{ must be higher}$$

With diminishing returns:

$$F_K \text{ higher} \Rightarrow K^* \text{ must be lower (for given } L^*)$$

First result: Capital tax reduces steady-state capital stock.

General Equilibrium Effects on Labor

Labor supply still determined by:

$$\frac{u_{1-\ell}(c^*, 1 - L^*)}{u_c(c^*, 1 - L^*)} = w^* = F_L(K^*, L^*)$$

Chain of effects:

1. Capital tax $\rightarrow$ lower K^* (direct effect)
2. Lower K^* $\rightarrow$ lower $F_L(K^*, L^*)$ (factor complementarity)
3. Lower marginal product of labor $\rightarrow$ lower wage w^*
4. Lower wage $\rightarrow$ lower labor supply L^* (substitution effect)
5. Lower L^* $\rightarrow$ even lower K^* (feedback through F_K condition)

Second result: Capital tax also reduces steady-state labor supply through general equilibrium effects.
Third result: Feedback effects amplify the initial capital reduction.

Key Takeaways from Capital Taxes with Lump Sum Transfers

Key insight: Capital tax creates distortions in BOTH factor markets through general equilibrium

Policy Insight: Shows that even "fiscally neutral" tax changes can have large efficiency effects through incentive distortions.

Key economic insight: Even when tax revenue is returned to taxpayers, the wedge between private returns and social returns creates real distortions to capital accumulation and labor supply decisions.

Role of Patience Parameter (β)

Steady state condition:

$$F_K(K^*, L^*) = \frac{1}{1-\tau_K}\left[\frac{1}{\beta} - 1 + \delta\right]$$

Effect of higher patience (higher β):

- **Without tax**: Higher $\beta \to$ lower required $F_K \to$ higher K^*
- **With tax**: Same relationship, but starting from higher required F_K

Tax effect depends on β:

$$\frac{\partial}{\partial \tau_K}\left[\frac{1}{1-\tau_K}\left(\frac{1}{\beta} - 1 + \delta\right)\right] = \frac{1}{(1-\tau_K)^2}\left(\frac{1}{\beta} - 1 + \delta\right)$$

Result: More patient agents (higher β) are **less sensitive** to capital taxes because their baseline required return is lower.

Role of Labor-Leisure Preference (γ)

Labor supply condition: (with CRRA utility)

$$\frac{1-\gamma}{\gamma}\frac{c^*}{1-L^*} = F_L(K^*, L^*)$$

Effect of higher work preference (higher γ):

- Higher baseline labor supply L^*
- Higher $L^* \to$ higher $F_K(K^*, L^*) \to$ supports higher K^*
- Capital and labor are **complementary**

Tax interaction with γ:

- Higher $\gamma \to$ higher baseline (K^*, L^*)
- Capital tax reduces both, but effects are **amplified** by complementarity
- Higher $\gamma \to$ **larger** decreases in capital and labor for the same τ

Policy insight: Work ethic (γ) and capital accumulation are complementary, so capital taxes are especially harmful in societies with strong preferences for work.

Role of Risk Aversion (θ)

Risk aversion affects transition dynamics, not steady state (last lecture)

Steady state independent of θ:

- Long-run (K^*, L^*, c^*) same regardless of θ
- Risk aversion is about consumption smoothing, not levels

But θ affects tax policy response:

- **High θ (risk averse):** Slow adjustment to new steady state after tax introduction
- **Low θ (risk tolerant):** Fast adjustment, more volatile transition

Policy insight:

- More risk-averse populations take longer to adjust to tax changes
- Transition costs may be spread over longer periods

Part 2: Intro To Uncertainty

Until Now

Up until now the models we have discusses have been **deterministic**. The model where outcomes are completely determined by initial conditions.

Key Characteristics:

- **Predictability:** Given the same initial conditions, a deterministic model will always produce exactly the same outcome.
- **No Random Elements:** There are no probability distributions, random variables, or stochastic shocks.
- **Mathematical Certainty:** The relationship between inputs and outputs follows exact mathematical rules, i.e. the production function is not changing

Why Models With Uncertainty Are More Realistic

Real Economic Decisions Are Made Under Uncertainty: Consumers don't know their future income, firms don't know future demand, and policymakers don't know how shocks will hit the economy. Deterministic models assume away this fundamental feature of economic life.

Empirical Reality: Real economic data shows volatility, cycles, and unpredictable fluctuations that deterministic models simply cannot generate or explain.

Unique Insights from Models With Uncertainty

Risk and Precautionary Behavior: Deterministic models can't explain why people save beyond lifecycle needs or why insurance exists.

Business Cycles and Fluctuations: Deterministic models predict smooth paths to steady states - they can't generate the cyclical behavior we observe.

Asset Pricing and Risk Premiums: Deterministic models would predict identical returns across all assets - clearly counterfactual.

Moving Forward

We will now move to a studying consumption and savings decisions in a world of uncertainty.

Aggregate Uncertainty (Systemic Risk): Uncertainty that affects everyone at the same time - shocks that hit the entire economy.

- **Key Feature:** These risks don't cancel out - they affect everyone simultaneously, so there's genuine uncertainty about aggregate outcomes.

Individual Uncertainty (Idiosyncratic Risk): Uncertainty that affects individual people or firms differently - your personal "bad luck" or "good luck.

- **Key Feature:** These risks cancel out when you look at the whole economy, so aggregate states can be predicted. If 100 people each have a 10% chance of losing their job, about 10 will be unemployed, but we know roughly how many.

What is a Stochastic Process?

A stochastic process is a collection of random variables indexed by time:

$$\{X_t\}_{t \in T} = \{X_t : t \in T\}$$

where T is the index set (usually time).

Examples in Macroeconomics:

- $\{Y_t\}$: GDP over time
- $\{A_t\}$: Total Factor Productivity
- $\{r_t\}$: Interest rates
- $\{\pi_t\}$: Inflation rates

Each realization is called a **sample path** or **trajectory**.

Stationarity

Strict Stationarity: The joint distribution of $(X_{t_1}, X_{t_2}, \ldots, X_{t_n})$ is the same as $(X_{t_1+h}, X_{t_2+h}, \ldots, X_{t_n+h})$ for any h.

Weak Stationarity (Covariance Stationarity):

1. $\mathbb{E}[X_t] = \mu$ (constant mean)
2. $\text{Var}(X_t) = \sigma^2$ (constant variance)
3. $\text{Cov}(X_t, X_{t-j}) = \gamma_j$ (depends only on lag j)

Why Care? Many macro models assume stationarity for:

- Balanced growth paths
- Long-run equilibrium analysis
- Forecasting and policy analysis

White Noise Process

$\{\varepsilon_t\}$ is white noise if:

1. $\mathbb{E}[\varepsilon_t] = 0$ for all t
2. $\text{Var}(\varepsilon_t) = \sigma^2$ for all t
3. $\text{Cov}(\varepsilon_t, \varepsilon_s) = 0$ for $t \neq s$

Strong White Noise: Additionally requires ε_t to be independent across time.

Notation: $\varepsilon_t \sim WN(0, \sigma^2)$

This is the building block for most macro stochastic processes.

Random Walk

Definition:

$$X_t = X_{t-1} + \varepsilon_t$$

where $\varepsilon_t \sim WN(0, \sigma^2)$.

Properties:

- $\mathbb{E}[X_t] = X_0$ (if X_0 is deterministic)
- $\text{Var}(X_t) = t\sigma^2$ (variance grows with time!)
- Non-stationary process
- $\mathbb{E}[X_t | X_{t-1}] = X_{t-1}$ (best forecast is current value)

Macro Examples:

- Log GDP (with drift)
- Asset prices (efficient markets)
- Some models of permanent income

Random Walk with Drift

$$X_t = \mu + X_{t-1} + \varepsilon_t$$

Solution by recursive substitution:

$$X_t = X_0 + \mu t + \sum_{i=1}^{t} \varepsilon_i$$

Properties:

- $\mathbb{E}[X_t] = X_0 + \mu t$ (trending mean)
- $\text{Var}(X_t) = t\sigma^2$
- Still non-stationary

This is often used to model trending macroeconomic variables like GDP.

Autoregressive Process AR(1)

$$X_t = \phi X_{t-1} + \varepsilon_t$$

where $\varepsilon_t \sim WN(0, \sigma^2)$.

Stationarity Condition: $|\phi| < 1$

Properties (when stationary):

- $\mathbb{E}[X_t] = 0$
- $\text{Var}(X_t) = \frac{\sigma^2}{1-\phi^2}$
- $\text{Cov}(X_t, X_{t-j}) = \phi^j \frac{\sigma^2}{1-\phi^2}$

Autocorrelations decay exponentially: $\rho_j = \phi^j$

AR(1) with Intercept

$$X_t = c + \phi X_{t-1} + \varepsilon_t$$

Long-run mean: $\mu = \frac{c}{1-\phi}$

Alternative representation:

$$X_t - \mu = \phi(X_{t-1} - \mu) + \varepsilon_t$$

Properties:

- $\mathbb{E}[X_t] = \mu = \frac{c}{1-\phi}$
- Mean reversion when $0 < \phi < 1$
- Process oscillates when $-1 < \phi < 0$

Why AR(1) Processes Are Fundamental in Macro

Core Economic Intuition:

- Most economic variables exhibit **persistence with mean reversion**
- Shocks have lasting effects but eventually fade away
- Captures the middle ground between white noise and random walks

AR(1) processes are commonplace in macro because they capture:

1. **Realistic Dynamics**: Most economic variables show persistence with mean reversion
2. **Parsimony**: One parameter captures essential time-series behavior
3. **Tractability**: Easy to analyze theoretically and computationally
4. **Flexibility**: Can model various degrees of persistence

Key Applications:

- Technology and productivity shocks
- Monetary and fiscal policy processes
- External shocks in open economy models
- Preference and demand disturbances
- Financial market conditions

Application: Technology Shocks

Total Factor Productivity (TFP): $Y_t = A_t F(K_t, L_t)$

$$\log A_t = \rho_A \log A_{t-1} + \varepsilon_t^A$$

Why This Specification?

- **Persistence**: Technological improvements don't disappear overnight ($\rho_A \approx 0.95$)
- **Mean Reversion**: Technology can't grow without bound relative to trend
- **Propagation**: Creates realistic business cycle fluctuations

Economic Mechanism:

1. Positive productivity shock increases output
2. Higher productivity persists, encouraging investment
3. Gradual return to trend creates boom-bust cycles

Random Walk: Would imply permanent productivity changes - unrealistic

Markov Property: Intuitive Definition

Simple Idea: The future depends only on the present, not on how we got here.

Everyday Example - Weather:

- **Markov**: Tomorrow's weather depends only on today's weather
- **Non-Markov**: Tomorrow's weather depends on today's weather AND last week's patterns AND seasonal history

Economic Example - Unemployment:

- **Markov**: Tomorrow's employment status depends only on today's status
- **Non-Markov**: Tomorrow's status depends on current status AND duration of unemployment AND previous job history

Key Insight: Current state contains all relevant information for predicting the future.

Markov Property: Mathematical Definition

A stochastic process $\{X_t\}$ has the Markov property if:

$$\mathbb{P}(X_{t+1} \leq x | X_t, X_{t-1}, \ldots, X_0) = \mathbb{P}(X_{t+1} \leq x | X_t)$$

Intuition: The future depends only on the current state, not the entire history.

Conditional Independence:

$$\mathbb{E}[X_{t+1} | X_t, X_{t-1}, \ldots] = \mathbb{E}[X_{t+1} | X_t]$$

Why Important in Macro?

- State variables in dynamic programming
- Rational expectations models
- Computational tractability

Examples: Markov vs Non-Markov

Markov Process - Random Walk:

$$X_t = X_{t-1} + \varepsilon_t$$

$\mathbb{E}[X_{t+1} \mid X_t, X_{t-1}, \ldots] = X_t$ (only current value matters)

Markov Process - AR(1):

$$X_t = \rho X_{t-1} + \varepsilon_t$$

$\mathbb{E}[X_{t+1} \mid X_t, X_{t-1}, \ldots] = \rho X_t$ (only current value matters)

Non-Markov Process - AR(2):

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t$$

$\mathbb{E}[X_{t+1} \mid X_t, X_{t-1}, \ldots] = \phi_1 X_t + \phi_2 X_{t-1}$ (need two lags!)

Trick for AR(2): Define state vector $\mathrm{S_t} = \begin{pmatrix} X_t \\ X_{t-1} \end{pmatrix}$

Then the system becomes Markov in the expanded state space.

Why we want Markov Property: Computational Tractability

The Curse of Dimensionality:

- Without Markov property: must track entire history
- State space grows exponentially with time
- Computational nightmare for dynamic models

With Markov Property:

- State space remains fixed dimension
- Only need current state for optimal decisions
- Dynamic programming becomes feasible

Example - Savings Decision:

- **Without Markov**: Optimal investment depends on $(K_t, A_t, A_{t-1}, A_{t-2}, \ldots)$
- **With Markov**: Optimal investment depends only on (K_t, A_t)

Discrete State Space Markov Chain

Definition: A stochastic process that can only take on a finite (or countable) number of distinct values, where transitions between states follow the Markov property.

Key Components:

1. State space: $S = \{s_1, s_2, \ldots, s_N\}$ (finite set of possible values)
2. Transition Matrix: $P_{ij} = \mathbb{P}(X_{t+1} = s_j | X_t = s_i)$. Matrix form:

$$\mathrm{P} = \begin{pmatrix} P_{11} & P_{12} & \cdots & P_{1N} \\ P_{21} & P_{22} & \cdots & P_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ P_{N1} & P_{N2} & \cdots & P_{NN} \end{pmatrix}$$

   Each row sums to 1: $\sum_{j=1}^{N} P_{ij} = 1$
3. Markov Property: $\mathbb{P}(X_{t+1} | X_t, X_{t-1}, \ldots) = \mathbb{P}(X_{t+1} | X_t)$

Definition of Stationarity for Discrete Markov Chains

A discrete Markov chain is stationary if there exists a probability distribution $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_N)$ such that:

$$\boldsymbol{\pi} = \boldsymbol{\pi}\mathrm{P}$$

Or nn component form: $\pi_j = \sum_{i=1}^{N} \pi_i P_{ij}$ for all j.

Interpretation: If the process starts with distribution $\boldsymbol{\pi}$, it maintains this distribution forever.

Existence: Under regularity conditions (irreducibility, aperiodicity), a unique stationary distribution exists.

Macro Application: Long-run distribution of productivity shocks, employment states, etc.

Alternative Markov Models for TFP

Instead of continuous AR(1): $\log A_t = 0.95 \log A_{t-1} + \varepsilon_t$

Discrete Markov Chain:

- States: $\{A_L, A_M, A_H\} = \{0.95, 1.00, 1.05\}$ (Low, Medium, High TFP)
- Transition matrix:

$$P = \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.2 & 0.6 & 0.2 \\ 0.1 & 0.3 & 0.6 \end{pmatrix}$$

This means: if TFP is currently Low (A_L), there's a 60% chance it stays Low, 30% chance it goes to Medium, 10% chance it jumps to High.

When to Use Each Approach

Use AR(1) When:

- Data shows linear relationship with normal errors
- Want analytical tractability and closed-form solutions
- Parsimony is important (few parameters)
- Computational speed is crucial

Use General Markov When:

- Evidence of nonlinear dynamics or regime changes
- Asymmetric responses to positive vs negative shocks
- Need to model rare disasters or extreme events

Example 1: Regime-Based Economic Modeling

Economic Intuition: Sometimes TFP exhibits distinct "regimes" rather than smooth transitions.

Historical Examples:

- **1970s Oil Shocks**: Sharp discrete drop in energy productivity
- **IT Revolution (1990s)**: Sudden jump to higher productivity regime
- **2008 Financial Crisis**: Abrupt fall in financial sector productivity
- **COVID-19**: Discrete shift in remote work productivity

Policy Regimes:

$$\mathrm{P} = \begin{pmatrix} \text{Low Reg} & \text{High Reg} \\ 0.9 & 0.1 \\ 0.05 & 0.95 \end{pmatrix}$$

- Low Regulation: $A_L = 0.95$ (persistent, $P_{LL} = 0.9$)
- High Regulation: $A_H = 1.05$ (very persistent, $P_{HH} = 0.95$)

AR(1) Problem: Would smooth these transitions unrealistically.

Example 2: Asymmetric Dynamics

AR(1) Limitation: Symmetric responses to positive and negative shocks.

Discrete Chain Advantage: Can model asymmetric transitions.

Example - "Easy to Fall, Hard to Rise":

$$P = \begin{pmatrix} & L & M & H \\ L & 0.8 & 0.2 & 0.0 \\ M & 0.3 & 0.5 & 0.2 \\ H & 0.1 & 0.7 & 0.2 \end{pmatrix}$$

Economic Interpretation:

- **From Low**: Hard to jump directly to High (0% chance)
- **From High**: Easy to fall to Medium (70% chance)
- **From Medium**: About equally likely to rise or fall

Applications: Financial crises, technology adoption, reputation effects.

Example 3: Rare Disasters and Extreme Events

Two-State Example:

- **Normal Times**: $A = 1.0$, Probability $= 98.3\%$
- **Disaster**: $A = 0.64$ (36% drop), Probability $= 1.7\%$

$$P = \begin{array}{cc} \text{Normal} & \text{Disaster} \\ \multicolumn{2}{c}{\begin{pmatrix} 0.983 & 0.017 \\ 0.5 & 0.5 \end{pmatrix}} \end{array}$$

Key Features:

- Disasters are rare but severe
- Recovery takes time (50% chance to exit disaster each period)
- Fat tails in productivity distribution

AR(1) Problem: Normal distribution cannot generate extreme, rare events observed in data (Great Depression, wars, pandemics).

Example 4: Structural Break Modeling

U.S. Productivity Growth Regimes:

- 1947-1973: High growth (2.8% annually)
- 1974-1995: Slow growth (1.4% annually)
- 1996-2004: High growth again (2.5% annually)
- 2005-2015: Slow growth (1.2% annually)

Three-Regime Model:

$$P = \begin{pmatrix} & \text{Low} & \text{Med} & \text{High} \\ \text{Low} & 0.95 & 0.05 & 0.00 \\ \text{Med} & 0.10 & 0.85 & 0.05 \\ \text{High} & 0.00 & 0.10 & 0.90 \end{pmatrix}$$

Economic Interpretation:

- Regimes are very persistent (90-95% probability of staying)
- Transitions occur gradually through medium regime
- Captures decades-long productivity patterns