Econ. 704, Spring 2007, **François Gourio**

Handout 1: Deterministic Dynamic Programing

In this handout we look at dynamic problems under certainty, and consider two approaches to solving them: the sequence problem (SP) approach, which uses first-order condition (*aka* Euler equations) to characterize the solution as an optimal path, and the dynamic programming approach (DP), which views the solution through the lens of a value function and a policy function. Both formulations are useful, and depending on the question at hand, one of the approaches may be easier to adapt.

1 Statement of the two approaches and FOCs

We start with a statement of an optimization problem (e.g. utility maximization or social welfare maximization), which is defined by its primitives.

Primitives: state space X, return function $F : X \times X \to \mathbb{R}$, feasibility correspondence $\Gamma : X \rightrightarrows X$, discount factor $0 < \beta < 1$.

Sequence problem (SP):

$$\begin{aligned} v^*(x_0) &= \max_{\{x_{t+1}\}_{t=0}^{\infty}} \sum_{t \geq 0} \beta^t F(x_t, x_{t+1}) \\ s.t. &: x_{t+1} \in \Gamma(x_t), \text{ for all } t \geq 0 \\ & \quad x_0 \text{ given} \end{aligned}$$

v^* is the *value function.* An *optimal plan* is a sequence $\{x_t\}_{t=0}^{\infty}$ that solves this problem.

The first order condition (assuming differentiability of F and an interior solution) are obtained by the standard Lagrangean, yielding:

$$\text{for all } t \geq 0 : F_2(x_t, x_{t+1}) + \beta F_1(x_{t+1}, x_{t+2}) = 0, \qquad (EE)$$

where F_1 is the partial derivative of F with respect to its first argument, and F_2 the partial with respect to the second argument. This is the standard Euler equation. As you know, this equation is not sufficient for optimization - you need to add the transversality condition. More precisely, given x_0 and x_1 there is "in general" a unique solution to this second-order (nonlinear) difference equation (EE). However, while x_0 is an initial condition which is given in the statement of the problem, x_1 is not given and must be determined. To find it, we need to add the "transversality condition" (TVC):

$$\lim_{t \to \infty} \beta^t F_x(x_t, x_{t+1}) x_t = 0.$$

Result : if the function F is increasing in its first argument, is concave, and if the solution is interior, then EE and TVC are necessary and sufficient for the problem (SP).

A slightly different formulation of the sequence problem is the "state-controls formulation":

$$\begin{aligned} v^*(x_0) &= \max_{\{u_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t r(x_t, u_t), \\ s.t. &: x_{t+1} = h(x_t, u_t) \text{ for all } t \geq 0 \\ & \quad x_0 \text{ given.} \end{aligned}$$

$\{u_t\}$ = control variable(s) and $\{x_t\}$ = state variable(s), and the equation $x_{t+1} = h(x_t, u_t)$ is referred to as the *state law of motion*. You need to choose the path of controls $\{u_t\}$ to maximize your objective function, taking into account its impact on future state variables. The solution takes the form of a control rule $u_t = p(x_t)$. This formulation is equivalent to (SP) (see Problem 1). Optimal control was developed by engineers, so a concrete example would be x_t = position of a rocket at time t and u_t =engine power at time t.

Bellman equation (BE)

$$v(x) = \sup_{y\in\Gamma(x)} \{F(x,y) + \beta v(y)\}$$

The optimal y in the Bellman equation for a given x yields the *policy function* $y^* = g(x)$. To solve the Bellman equation is to find a *function* v which satisfies this equation, and once v is found we can compute the policy function g by computing, for each value of x, the right-hand side of this equation.

Assuming that the return function F and te function v are both differentiable, and that the solution is interior, we have the first-order condition:

$$F_2(x, g(x)) + \beta v'(g(x)) = 0,$$

and the envelope condition:

$$v'(x) = F_1(x, g(x)).$$

Example: Neoclassical Growth Model.

$$\max_{\{k_{t+1}\}_{t=0}^{\infty}} \sum_{t\geq 0} \beta^t U(c_t)$$

$$\begin{aligned} s.t. &: \quad c_t + k_{t+1} \leq (1-\delta)k_t + F(k_t) \\ k_{t+1} &\geq 0, \qquad k_0 \text{ given.} \end{aligned}$$

In this case the primitives are:

$$X = \mathbb{R}^+,$$

$$F(x,y) = U((1-\delta)x + F(x) - y),$$

$$\Gamma(x) = [0, +\infty[,$$

and β is β!

Exercise: check that the FOC of the sequence problem gives you the usual consumption Euler equation.

2 Theorems

Principle of Optimality

Under weak conditions, the value function v^*solution to (SP) satisfies the equation (BE).

(Proof : break down the problem between today $t = 0$ and from tomorrow on $t = 1$.)

This key result allows us to try and solve (BE) which is easier. Note that v is on both sides of (BE) though so it is not so simple!

Existence and Uniqueness of a solution to the Bellman equation

Assumptions (AA): X is convex; $\Gamma : X \rightrightarrows X$ is a continuous correspondence such that for all $x \in X$, $\Gamma(x)$ is a compact space; F is continuous and bounded; and $0 < \beta < 1$.

Then: there exists a unique continuous bounded function v satisfying (BE).

Outline of proof: let T be the *Bellman operator* mapping any function w into another function Tw defined for each $x \in X$ by:

$$(Tw)(x) = \sup_{y \in \Gamma(x)} \{F(x, y) + \beta w(y)\},$$

i.e. Tw is the right-hand side of the Bellman equation if you plug w into it. A solution to (BE) is a function v such that $v = Tv$, a fixed point of T. We use a math theorem which states that any contraction mapping on a complete space[1] has a fixed point. A contraction mapping is a function T such that there is some $0 < k < 1$, such that $\|Tv - Tw\| \leq k \|v - w\|$, for all functions v, w, where $\|f\| = \sup_{x \in X} |f(x)|$ is a measure of "distance" between two functions. To prove that T is indeed a contraction, we use the following important result which you should know and be able to use:

Blackwell's sufficient conditions for T to be a contraction:

If T satisfies (i) for any functions v, w: $v \leq w \Rightarrow Tv \leq Tw$ and (ii) there is a $0 < k < 1$ such that for any function v and constant a: $T(v + a) \leq Tv + ka$, Then T is a contraction.

To conclude our outline: check that in our case we can apply Blackwell's sufficient conditions.

Note: in many cases we will use this theorem without the assumption that F is bounded. It is possible to justify this by extending the theorem under some assumptions.[2]

Guess-and-verify

This result tells us that (under some conditions) there is a unique solution to (BE), and it is the solution to (SP). Hence if we find one solution to (BE), we're done. This means that if we make some guess and check that it satisfies (BE) then we're done. Often we have an "approximate guess", meaning we guess the solution up to some parameters, plug it in the right-hand side of (BE), compute the right-hand side (i.e. perform the maximization over $y \in \Gamma(x)$), and find the parameter that make the right-hand side equal to our (left-hand side) guess.

---

[1]A complete space is a vector space where all Cauchy sequences have a limit. (See SLP for definitions and details.) You do not need to know this, but it is useful to know that any closed subspace of a complete space is complete. Simply recall that (i) any finite-dimensional vector space is complete; (ii) the space of continuous and bounded functions on a space X is complete; (iii) the space of bounded functions on a space X is complete.

[2]The technical reference (not required at all!) is Alvarez and Stokey, Journal of Economic Theory 1998.

Example of guess and verify to work out in class:

$$\begin{aligned} v^*(k) &= \sup_{\{k_{t+1}\}_{t=0}^{\infty}} \sum_{t\geq 0} \beta^t \frac{c_t^{1-\gamma}}{1-\gamma} \\ s.t. &: \\ k_{t+1} &= A(k_t - c_t), \\ k_0 &> 0. \end{aligned}$$

Convergence by value iteration

Under the same assumptions (AA), pick an initial guess for v, say a function v_0. Next for $j = 0, 1, ...$, compute a new function recursively (i.e. by iteration):

$$v_{j+1}(x) = \sup_{y\in\Gamma(x)} \left\{F(x,y) + \beta v_j(x)\right\},$$

then our guesses v_j will converge as $j \to \infty$ to the value function v : $\lim_{j\to\infty} \|v_j - v\| = 0$, whatever the initial guess. This gives a practical way of solving numerically for a value function. Vladimir on Friday will go through an example of this method.

Proving that the value function is increasing

Assumptions: assume (AA) and moreover (i) if $x \leq x'$ then $\Gamma(x) \subset \Gamma(x')$; (ii) F is increasing in the first argument.

Result: the value function v is (weakly) increasing.

Idea of the proof: We show that the Bellman operator T maps weakly increasing functions into weakly increasing functions. This implies that T has a fixed point which is weakly increasing, since T is a contraction on the closed (thus complete) subspace of the weakly increasing functions. Concretely, we need to check that if v is increasing, then so is Tv. Check now that the assumptions I give here guarantee this.

Note: actually one can improve on this and show that v is strictly increasing.

Proving that the value function is concave

Assumptions: assume (AA) and moreover (i) F is concave, (ii) Γ is convex i.e. $y \in \Gamma(x)$, $y' \in \Gamma(x')$ implies that $\theta y + (1-\theta)y' \in \Gamma(\theta x + (1-\theta)x')$ for all $\theta \in [0,1]$.

Result: the value function v is concave in x.

Idea of the proof: apply the same method than for the "increasing" result.

Proving that the value function is differentiable, and the Benveniste-Scheinkman formula

Assumptions: F is continuously differentiable and [as in the concavity result] assume (AA) and moreover (i) F is concave, (ii) Γ is convex i.e. $y \in \Gamma(x)$, $y' \in \Gamma(x')$ implies that $\theta y + (1-\theta)y' \in \Gamma(\theta x + (1-\theta)x')$ for all $\theta \in [0,1]$.

Result: If $x \in int(X)$ and if $y^*(x) = g(x) \in int(X)$, then v is differentiable at x and

$$v'(x) = F_1(x, g(x)),$$

where F_1 means the derivative with respect to first argument.

Note: $int(X)$ means the "interior" points of the set X, i.e. the points which are not at the "boundary". The formula can be seen as a form of envelope theorem:

$$v(x) = \sup_{y\in\Gamma(x)} \left\{F(x,y) + \beta v(y)\right\} \to v'(x) = F_x(x, g(x)).$$

3 Dynamics and linear approximation

Dynamics of a first-order difference equation: the linear case

Consider a first-order linear difference equation: for all $t = 0, 1, 2, ...$:

$$x_{t+1} = bx_t + c,$$

where b and c are real numbers. Assume we have some starting value x_0. Iterating on this equation will give us the sequence $\{x_t\}$. Depending on the values of b and c, this sequence may or may not be monotonic, and it may or may not converge to some limiting value as $t \to \infty$. It is easy, however, to derive conditions under which there will be convergence. Define x^* as the steady-state of this 'system', i.e. a constant which satisfies the equation for all $t \geq 0$:

$$x^* = bx^* + c,$$

i.e. $x^* = \frac{c}{1-b}$, if $b \neq 1$. Then, we have

$$x_{t+1} - x^* = b\left(x_t - x^*\right).$$

By induction, it is clear that for any value of $t = 0, 1, 2, ...$:

$$x_t - x^* = b^t\left(x_0 - x^*\right).$$

From this equation, we can see that if $|b| < 1$, $x_t - x^* \to 0$ as $t \to \infty$, whatever is the initial value of x_0. In this case, we say that the system is globally convergent towards x^*. On the other hand, if $|b| > 1$, unless $x_0 = x^*$, there will never be convergence to x^* since the sequence $|x_t - x^*|$ is increasing.

Dynamics of a first-order difference equation: the nonlinear case

Now consider a nonlinear first-order difference equation: $x_{t+1} = g(x_t)$, but g need not be linear. First, a definition:

Steady-state

A steady-state is a *constant* solution: $x_t = x^*$ for all $t \geq 0$. Equivalently, $g(x^*) = x^*$.

Linearization of the policy function around the steady-state

The trick is to make the equation 'locally linear'. Since $x_{t+1} = g(x_t)$, and $g(x^*) = x^*$, in this case we have for x_t close to x^* the following Taylor first-order approximation:

$$\begin{aligned} x_{t+1} - x^* &= g(x_t) - g(x^*) \\ &\simeq g'(x^*) \times (x_t - x^*). \end{aligned}$$

If x_t starts close enough to x^*, this approximation will always be valid. Hence using the results for linear equations, if $|g'(x^*)| < 1$, the system is locally stable: if x_0 is close enough to x^*, then $g'(x^*)$ measures the speed of convergence to the steady-state: if $g'(x^*)$ is close to 1, the convergence will take more time than if it is close to zero. If $|g'(x^*)| > 1$, the system is locally unstable since $\{x_t\}$ will in general not converge to x^*.

[plot on the blackboard]

Application to our problem

Our problem's solution is characterized by the policy function $g : x_{t+1} = g(x_t)$

We can find $g'(x^*)$ by using the FOCs when everything is "smooth enough":

$$
\begin{aligned}
\forall x &\in X : F_2(x, g(x)) + \beta v'(g(x)) = 0 \\
&\rightarrow F_{21}(x, g(x)) + F_{22}(x, g(x))g'(x) + \beta v''(x)g'(x) = 0
\end{aligned}
$$

$$
\begin{aligned}
\forall x &\in X : v'(x) = F_1(x, g(x)) \\
&\rightarrow v''(x) = F_{11}(x, g(x)) + F_{12}(x, g(x))g'(x)
\end{aligned}
$$

$$F_{21} + F_{22}g' + \beta\left(F_{11} + F_{12}g'\right) \times g' = 0$$

$$F_{21} + \left(F_{22} + \beta F_{11}\right) g' + \beta F_{12}g'^2 = 0$$

For $x = x^*$, we know that $g(x) = x^*$ as well so we can evaluate the terms in F_{21}, F_{22}, etc. and then solve this quadratic equation.

Linearizing the Euler equation

An equivalent method is to use the Euler equation. The Euler equation can also be linearized around the steady-state:

$$\forall t \geq 0 : F_2\left(x_t, x_{t+1}\right) + \beta F_1\left(x_{t+1}, x_{t+2}\right) = 0.$$

$$F_{21}^* \times \left(x_t - x^*\right) + F_{22}^* \times \left(x_{t+1} - x^*\right) + \beta F_{11}^* \times \left(x_{t+1} - x^*\right) + \beta F_{12}^* \times \left(x_{t+2} - x^*\right) = 0,$$

where F_{ij}^* is the partial derivative evaluated at the steady-state (x^*, x^*). Simplify:

$$F_{21}^* \times \left(x_t - x^*\right) + \left(F_{22}^* + \beta F_{11}^*\right)\left(x_{t+1} - x^*\right) + \beta F_{21}^* \times \left(x_{t+2} - x^*\right) = 0.$$

This is now a second-order linear difference equation which we can solve. The solution is of the form $x_t = A\mu_1^t + B\mu_2^t$ where μ_1, μ_2 solve the quadratic equation in X

$$F_{11}^* + \left(F_{22}^* + \beta F_{11}^*\right) X + \beta F_{22}^* X^2 = 0,$$

and where A and B are constants to be determined. One constant is determined so that $x_0 = A\mu_1^0 + B\mu_2^0$, the other constant is determined so that the transversality condition (TVC) is satisfied.

Note that the product of the two roots if β. Hence if one root is greater in absolute value than $\frac{1}{\beta} > 1$, then the other root is (in absolute value) below 1. In this case the coefficient in front of the large root must be set to zero to ensure that (TVC) holds. Then the other coefficient is chosen to match x_0 and so there is a unique solution. This is known as the *saddle-path stable* case.

In the case where the larger root in absolute value is smaller than $\frac{1}{\beta}$, it may be that there is indeterminacy, i.e. there are many solutions. This doesn't happen much in practice.

Extension to the vector case (Not required)

In general, we can consider the first-order vector difference equation $x_{t+1} = ax_t$, where x_t is a $k \times 1$ vector and a is a $k \times k$ matrix. (This includes cases where there is more than one lag, by stacking equations.) Here's the general method to deal with such equations. You can often write the matrix a as $a = P^{-1}\Lambda P$ where Λ is a diagonal matrix and P is a nonsingular matrix.[3] This yields $Px_{t+1} = \Lambda Px_t$. Defining $y_t = Px_t$, i.e. a new vector which is a linear combination of x_t, we have $y_{t+1} = \Lambda y_t$.

Given that the solution is pinned down with two initial conditions

[3] If this not possible, there are other matrix decompositions you can use: the Jordan decomposition is used for proofs, and there other decompositions which are also useful for numerical applications (because they are more robust numerically than Jordan).

Handout to clarify the last lecture

We can study difference equations $x_{t+1} = g(x_t)$ or $x_{t+2} = h(x_{t+1}, x_t)$ without knowing where they come from. In this case we need the 'right' number of initial conditions to run them (1 in the first case, two in the second), and we can look at the three questions:

(1) is there a steady-state x^*?

(2) is it unique?

(3) is it stable, i.e. if we start with a point $x_0 \neq x^*$, do we converge to the steady-state or not? is it locally stable, i.e. if x_0 is close enough to x^*, will we converge towards it? (We know that the answer to the last question depends, for the case $x_{t+1} = g(x_t)$, on whether $|g'(x^*)|$ is greater or less than one.)

- When we study difference equations which come from maximization problems, things are somewhat different. In some cases we know already that the system must converge to some steady-state; in other cases we are not sure, but we usually want to keep the solution bounded. Now it depends which approach you use to solve the problem:

(a) if you consider the policy function, found from solving the Bellman equation, you have a difference equation $x_{t+1} = g(x_t)$ and one initial condition x_0; so you can compute numerically $\{x_t\}$ by iterating on this equation. When you compute the linear approximation around the steady-state $g'(x^*)$, you want to pick the root which is less than 1 in absolute value since this is the one that will give you convergence.

(b) if instead you use the Euler equation, you will have a second-order difference equation, with only one initial condition. As we know, the TVC condition will imply how you should pick the other initial condition. Mathematically, once you have linearized the Euler equation, you will have a second-order difference linear equation, for which there are usually two parameters to set using initial conditions. But one parameter will be set by the requirement that the solution is bounded (equivalently, converges to the steady-state). Again, there are some 'weird' cases when you do not have one root greater than one and another root smaller than one, but this is infrequent.

Second order linear difference equations

Consider the second-order difference equation: $aX_{t+2} + bX_{t+1} + cX_t = 0$, for all $t \geq 0.$, Clearly, if we want to be able to run this recursion, we need two initial conditions X_0 and X_1.

Result: the solutions of this second-order difference equation is $X_t = A \times \lambda_1^t + B \times \lambda_2^t$, where λ_1 and λ_2 are the roots of the quadratic equation $aX^2 + bX + c$, and A and B are two numbers which are picked to match the initial conditions X_0, X_1[4], i.e.

$$\begin{aligned} X_0 &= A \times \lambda_1^0 + B \times \lambda_2^0 = A + B, \\ X_1 &= A \times \lambda_1^1 + B \times \lambda_2^1 = A\lambda_1 + B\lambda_2. \end{aligned}$$

The proof is to check (as we did in class) that this formula for X_t works for any A and B. Given two initial conditions X_0 and X_1, we can clearly find the sequence X_t satisfying the difference equation. Hence, we have all the solutions since we can replicate any X_0, X_1 by picking A and B appropriately.

In the case where we want to keep only bounded solutions, and $|\lambda_1| > 1, |\lambda_2| < 1$,we will set $A = 0$ and pick B to match X_0.

[4] This works if the roots are different. If there is only one root, the solutions are of the form $(A + B \times t) \times \lambda^t$. If the roots are complex, λ_1 and λ_2 are complex conjugate numbers, but as long as X_0 and X_1 are real numbers, the solution will be made of real numbers.

Handout 2: Stochastic Processes and Stochastic Dynamic Programing

In this handout we start considering maximization problems when the decision make faces some uncertainty. Prominent examples of applications include a consumer who is uncertain about his future income or health, a firm which is uncertain about its future profitability, an unemployed who doesn't know if and when he will find a job, etc.

Solving these problems requires us to model the uncertainty that these people face, and then to incorporate this uncertainty into our optimizaton problem. So we start by considering Markov chains and ARMA (autoregressive-moving average) processes; these are simple and tractable processes which can approximate many time series in practice. (See the Sargent-Ljungqvist textbook for very nice chapter on time series.) Then we go over dynamic programming with uncertainty.

4 Stochastic Processes

A (discrete-time) stochastic process is a sequence (X_t) of random variables for $t = 0, 1, ..$

We will consider two classes stochastic processes, markov chains and ARMA processes.

Markov chains

A stochastic process is a (time-homogeneous) Markov *chain* if it takes values in a finite set S (called the state space) and if it satisfies the following Markov property:

$$\Pr\left(X_{t+1} \mid X_t\right) = \Pr\left(X_{t+1} \mid X_t, ..., X_0\right),$$

and moreover $\Pr\left(X_{t+1} \mid X_t\right)$ is independent of time. This means that the past only matters through the last period value of X.

This means that we can summarize a Markov chain by the following *conditional probabilities*:

$$\Pr\left(X_{t+1} = j \mid X_t = i\right), \text{ for } i, j = 1 \text{ to } S.$$

These are often represented by a $S \times S$ matrix P with typical element:

$$P(i,j) = \Pr\left(i \rightsquigarrow j\right) = \Pr\left(X_{t+1} = j \mid X_t = i\right),$$

the probability of going in one step from state i to state j. Note that this matrix satisfies $P(i,j) \geq 0$ for all i,j and moreover $\sum_{j=1}^{S} P(i,j) = 1$ for each $i = 1...S$. (Each row sums to one.)[5]

Of course to compute the evolution of X_t, we need an initial condition, i.e. X_0 is draw from a distribution μ_0 over $\{1...S\}$.

Graphical representation of a Markov chain: see blackboard.

Examples

(1) (detailed). $S = \{1, 2\}, \mu_0 = [1, 0]'$, $P = \begin{bmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{bmatrix}$. This matrix Q means that when you are in state 1, with probability 3/4 you stay in state 1, and with probability 1/4 you go to state 2. When you are in state 2, with probability 1/4 you go to state 1 and with probability

[5] Be careful as some authors sometimes consider the transpose of this matrix, i.e. $P(j,i)$.

3/4 you stay in state 2. Note that the $i-$th row gives the probability of going tomorrow to each state, conditional on being in state i today. Hence, the sum of each row must be 1. Inversely, the column j gives the conditional probabilities of going to j, conditional on being in states $1, 2.., S$. he distribution μ_0 indicates we start for sure in state 1.

(2) $S = \{1, 2\}, \mu_0 = [1, 0]', P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. This is a deterministic sequence. If t is odd, $X_t = 2$, and if t is even, $X_t = 1$.

(3) $S = \{1, 2\}, P = \begin{bmatrix} \alpha & 1-\alpha \\ \alpha & 1-\alpha \end{bmatrix}$ for some $\alpha \in [0, 1]$. This is an iid process: not only is $\Pr(X_{t+1} \mid X_t, ..., X_0) = \Pr(X_{t+1} \mid X_t)$ independent of $X_{t-1}, ..., X_0$, it is also independent of X_t : $\Pr(X_{t+1} \mid X_t) = \Pr(X_{t+1})$.

(4) $S = \{1, 2, 3\}, P = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/4 & 3/4 & 0 \\ 0 & 0.1 & 0.9 \end{bmatrix}$. Notice how after some time the system will leave state 3 and never return to it.

(5) $S = \{1, 2, 3\}, Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 2/3 & 1/3 \end{bmatrix}$. Notice how depending on the initial condition, the system may either get stuck in state 1, or oscillate stochastically b/w state 2 and state 3.

(6) Mehra and Prescott (1985, JME) model US consumption growth by the following two-state process: $S = \{5.4\%, -1.6\%\}, Q = \begin{bmatrix} \frac{1}{2}+p & \frac{1}{2}-p \\ \frac{1}{2}-p & \frac{1}{2}+p \end{bmatrix}$. State 1 is a "boom" with high consumption growth and state 2 is a "recession" with low consumption growth. If $p > 0$ there is some persistence in consumption growth, i.e. booms tend to be followed by booms and recessions by recessions. MP set $p = -0.28$.

(7) Example of having dependence over more than one period. Suppose the distribution of X_t depends on X_{t-1} and X_{t-2}. Then, by redefining the state as a vector (X_t, X_{t-1}), we have a markov chain.

Numerical simulation

Simulating a markov chain is easy given a random number generator. Given the state $x(t) \in S$, you need to use the conditional distribution $[Q(x(t), 1), ..., Q(x(t), S)]$ to draw $x(t+1)$. Draw a number $s(t)$ from a uniform distribution between 0 and 1, and if $s(t) < Q(x(t), 1)$, then $x(t+1) = 1$; if $Q(x(t), 1) \leq s(t) < Q(x(t), 1) + Q(x(t), 2)$, then $x(t+1) = 2$; and so on. Once you know $x(t+1)$ you can repeat the process with the new distribution $[Q(x(t+1), 1), ..., Q(x(t+1), S)]$, and so on to get a time series of the desired length.

Computing the probability of a history

As always,

$$\Pr(X_t, X_{t-1}, ..., X_0) = \Pr(X_t \mid X_{t-1}, ..., X_0) \Pr(X_{t-1} \mid X_{t-2}, ..., X_0) ... \Pr(X_1 \mid X_0)\mu(X_0).$$

Now given the Markov assumption, we have the simplified formula:

$$\begin{aligned}
\Pr\left(X_t, X_{t-1}, ..., X_0\right) &= \Pr(X_t \mid X_{t-1}) \Pr(X_{t-1} \mid X_{t-2}) ... \Pr(X_1 \mid X_0) \mu(X_0), \\
&= Q(X_{t-1}, X_t) Q(X_{t-2}, X_{t-1}) ... Q(X_0, X_1) \mu(X_0).
\end{aligned}$$

Probability of transitions in k periods (Chapman-Kolmogorov formula)

What is the probability of being in state j at time $t+2$ if you were in state i at time t? To find out, condition on where you were at time $t+1$:

$$\begin{aligned}
\Pr(X_{t+2} &= j|X_t = i) = \sum_{k=1}^{S} \Pr(X_{t+2} = j|X_{t+1} = k) \times \Pr(X_{t+1} = k|X_t = i) \\
\Pr(X_{t+2} &= j|X_t = i) = \sum_{k=1}^{S} Q(k, j) Q(i, k).
\end{aligned}$$

You recognize the formula for the product of two matrices; hence if I can $Q^{(2)}$ the matrix which gives the probability of transitions in two periods, we have $Q^{(2)} = Q \times Q = Q^2$. More generally, the $k-$period transition probabilities are given by Q^k.

Computing the expectation of a function of a Markov chain

Let F be a function defined over the state space S. The expectation of F tomorrow depends on today's state and can be computed:

$$\begin{aligned}
E\left(F(X_{t+1}) \mid X_t = s\right) &= \sum_{s'=1}^{S} P(s, s') F(s') \\
&= \left(PF\right)_s,
\end{aligned}$$

where PF is the multiplication of the matrix and the $S \times 1$ vector $\{F(s)\}_{s=1}^{S}$, and the $()_s$ denotes we look only at the $s-$th row.

Stacking these equations yields

$$E\left(F(X_{t+1}) \mid X_t = .\right) = PF,$$

so that the function "expectation of F tomorrow conditional on value today" is just the matrix multiplication of P and F. Similarly you can check that

$$E\left(F(X_{t+k}) \mid X_t = .\right) = P^k F.$$

Law of motion for probabilities

Let $\mu_{t,i} = \Pr\left\{X_t = i\right\}$ be the unconditional probabilities. Conditioning on the value of X_{t-1}, I have

$$\begin{aligned}
\mu_{t,j} &= \Pr\left(X_t = j\right) \\
&= \sum_{i=1}^{S} \Pr\left(X_t = j \mid X_{t-1} = i\right) \Pr\left(X_{t-1} = i\right) \\
&= \sum_{i=1}^{S} P_{i,j} \mu_{t-1,i}.
\end{aligned}$$

Hence, we can recursively compute the unconditional probabilities given the transition matrix and the initial distribution. Note that these equations can be stacked as:

$$\boldsymbol{\mu}_t = P' \boldsymbol{\mu}_{t-1},$$

where $\boldsymbol{\mu}_t$ is the column vector $\left(\mu_{t,i}\right)_{i=1}^{N}$; or $\boldsymbol{\mu}_t' = \boldsymbol{\mu}_{t-1}' P$. Iterating yields directly $\boldsymbol{\mu}_t' = \boldsymbol{\mu}_0' P^t$.

Limiting behavior, invariant (or ergodic) distribution

We can now ask, is there a limiting behavior to the Markov chain? In general, there is no "steady-state" in the sense of the state settling down, but we can ask how much time will be spent on average in each state s. For this we need to know if $\lim_{t\to\infty} \boldsymbol{\mu}_t$ exists. You can work out in some of the previous examples if this is the case.

If $\boldsymbol{\mu}^* = \lim_{t\to\infty} \boldsymbol{\mu}_t$ exists, then it must satisfy

$$\boldsymbol{\mu}^* = P' \boldsymbol{\mu}^*.$$

The solutions to this last equations are called the invariant distributions (or ergodic, or stationary, distributions of the Markov chain P). There may be none, there may be several, they may be degenerate. And given one such invariant distribution, there may or not be convergence to it in the long run. While a very general analysis can be done (as in SLP chapter 9), here I will just give two very useful results:

Result 1: if $\forall i, j = 1$ to $S, P_{i,j} > 0$, then the Markov chain given by P has a unique invariant distribution $\boldsymbol{\mu}^*$, and for any $\boldsymbol{\mu}_0$, we have $\lim_{t\to\infty} \boldsymbol{\mu}_t = \boldsymbol{\mu}^*$.

Result 2: if there is an $N \geq 1$ such that $P_{i,j}^{(N)} > 0$, where $P^{(N)}$ means the Nth power of P, then the Markov chain given by P has a unique invariant distribution $\boldsymbol{\mu}^*$, and for any $\boldsymbol{\mu}_0$, we have $\lim_{t\to\infty} \boldsymbol{\mu}_t = \boldsymbol{\mu}^*$.

ARMA processes

Let ε_t be an i.i.d. variable with mean 0 and variance σ^2. An $ARMA(p,q)$ process is defined by a recursion of the type:

$$x_t - \rho_1 x_{t-1} - ... - \rho_p x_{t-p} = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - ... - \theta_q \varepsilon_{t-q},$$

where μ is a constant and $\left(\rho_1, ..., \rho_p\right)$ and $(\theta_1, ..., \theta_q)$ are real numbers. These processes can replicate any kind of autocorrelation.

These processes are stationary meaning roughly that they replicate themselves on average over time, and thus the moments $E\left(x_t\right), Var\left(x_t\right), Cov\left(x_{t+k}, x_t\right)$ do not depend on calendar time t.

For simple examples, you should know (1) how to compute the mean, (2) how to compute the variance, (3) how to compute the autocorrelation (or covariances), (4) how to compute conditional expectations, (5) how to compute an impulse response function: i.e. what happens when there is a unit shock at time t :given $\varepsilon_t = 1$ and $\varepsilon_j = 0$ for all $j > t$, compute the path of $\{x_t\}$.

To derive all these things the key property to use is that ε_t is i.i.d. so $Cov\left(\varepsilon_t, \varepsilon_k\right) = 0$ for $t \neq k$ and $E_t \varepsilon_{t+j} = 0$ for any $j \geq 1$.

Key example: AR(1) process. (Very important.)

$$x_{t+1} = \rho x_t + \varepsilon_{t+1}.$$

Taking expectations yields $E x_{t+1} = \rho E x_t + 0 \to E x_t = 0$ for all t.

The variance satisfies:

$$V x_{t+1} = \rho^2 V x_t + \sigma^2,$$

since ε_{t+1} is independent of x_t. Hence

$$Vx_t = \frac{\sigma^2}{1-\rho^2} \text{ for all } t.$$

The autocorrelation is:

$$\begin{aligned} Corr\,(x_{t+k}, x_t) &= Corr\,(\rho x_{t+k-1} + \varepsilon_{t+k}, x_t) \\ &= ... \\ &= Corr\left(\rho^k x_t + \rho^{k-1}\varepsilon_{t+1} + ... + \varepsilon_{t+k}, x_t\right) \\ &= \rho^k, \end{aligned}$$

which decays at rate ρ. (Note for $\rho < 0$ you have oscillations.)

The conditional expectation is

$$\begin{aligned} E_t x_{t+k} &= E_t\left(\rho^k x_t + \rho^{k-1}\varepsilon_{t+1} + ... + \varepsilon_{t+k}\right) \\ &= \rho^k x_t. \end{aligned}$$

Finally the impulse response is given by $x_t = 1$, $x_{t+1} = \rho$, $x_{t+2} = \rho^2$, etc. (Plot.)

Second example: MA(1) process. $x_t = \varepsilon_t - \theta\varepsilon_{t-1}$. Check that $Ex_t = 0$ for all t, that $Cov\,(x_t, x_{t-k}) = 0$ for all $k \geq 2$ and compute $Var\,(x_t) = \sigma^2\left(1+\theta^2\right)$.

We will encounter such processes in the class and over time you will become familiar with them.

See the examples in the SL book.

5 Stochastic Dynamic Programing

This section doesn't go into the technical details. These were discussed in the case of deterministic dynamic programing and they are similar.

We now have a shock z which follows a Markov chain with transition matrix P. Besides this shock, the primitives are now a feasibility correspondence: $y \in \Gamma(x,z)$; a reward function $F(x,y,z)$; a discount factor β between 0 and 1; and a state space X.

Sequence problem

The problem is to choose a state-contingent plan to maximize the expected discounted value: i.e. at time 0 we think of all the possible events that may occur and we list what we will do in each state, i.e. which action y_t we will take if faced with the history of events $\{z_1, ..., z_t\}$. A history is denoted $z^t \in S^t$ (the product space) and the plan is $y_t(z^t)$, defined for all t and all histories z^t. (We will talk again about histories in the next class.)

Formally, the sequence problem is:

$$\begin{aligned} v^*(x_0, z_0) &= \max_{\{y_t(z^t)\}_{t,z^t}} \sum_{t\geq 0}\sum_{z^t\in S^t} \beta^t \pi_t\left(z^t\right) F(x_t\left(z^t\right), y_t\left(z^t\right), z_t) \\ s.t. &: \; y_t\left(z^t\right) \in \Gamma(x_t\left(z^t\right), z_t) \text{ for all } t, z^t, \\ & \quad\; x_0 \text{ and } z_0 \text{ given.} \end{aligned}$$

Here $\pi_t\left(z^t\right)$ is the probability of history z^t occuring. Of course this probability can be derived given z_0 and the transition matrix P.

The notation is hard, but it just reflects the sheer complexity (dimensionality) of this problem.

Bellman equation

In contrast, the Bellman equation is simple:

$$v(x,z) = \max_{y \in \Gamma(x,z)} \left\{ F(x,y,z) + \beta \sum_{z'=1}^{S} P(z,z')v(y,z') \right\}$$

We can generalize this to the case where z is a continuous random variable:

$$v(x,z) = \max_{y \in \Gamma(x,z)} \left\{ F(x,y,z) + \beta E_{z'/z} v(y,z') \right\},$$

where $E_{z'/z}$ means we take the expectation over z', conditional on z.

(Note: here and above, I assume that you are able to choose your state for sure tomorrow. Another formulation would be that $y' = g(y,z')$.)

Analysis

The analysis is similar to the case of deterministic dynamic programing:

(1) under weak conditions, the value function of the sequence problem solves the Bellman equation ("Principle of Optimality");

(2) under some continuity+boundedness conditions on F, compactness and continuity of $\Gamma(x,z)$, etc., we can apply the contraction mapping theorem and get that there is a unique solution to the Bellman equation;

(3) this warrants the "guess-and-verify" method and gives us an algorithm to solve the Bellman equation numerically;

(4) we can prove if F is increasing and Γ expands as x increases that the value function is increasing in x;

(5) we can similarly prove that it is concave in x;

(6) a novelty: we can prove now that the value function is increasing in the shock z; this requires that F is increasing in the state z; that $\Gamma(x,z_2) \subset \Gamma(x,z_1)$ for $z_1 \geq z_2$; and that the conditional probabilities "shift up" in the sense that the distribution $\{Q(z_1,i)\}_{i=1}^{S}$ first-order stochastically dominates[6] $\{Q(z_2,i)\}_{i=1}^{S}$ for $z_1 \geq z_2$;

(7) finally we can also study the policy functions which now depend on both x and $z : y(x,z)$.

Steady-states and invariant distributions

In general a stochastic system, which is constantly disturbed by random shocks, will never settle down to a steady-state. Hence to study the long-run behavior of such a dynamic system we look for invariant distribution. The key remark is that the vector (x,z) is jointly Markov because (x_{t+1}, z_{t+1}) is given by (a) the optimal policy rule $x_{t+1} = h(x_t, z_t)$ and the transition matrix $P(z_t, z_{t+1})$. Hence we can analyze the vector (x,z) just like any Markov chain (or markov process more generally). In particular if we can apply our results 1 or 2 we know it will have an invariant distribution.

[6] A distribution F first-order stochastically dominates another distribution G if one of the following equivalent statements is true: (i) for any increasing function u, $\int u(x)dF(x) \geq \int u(x)dG(x)$; (ii) for all x, $F(x) \leq G(x)$; (iii) a variable drawn from F is equal to a variable drawn from G plus some positive random variable: $X = Y + Z$ with $Z \geq 0$.

Handout 3: General Equilibrium, Complete Markets, Representative Agent

In this lecture I apply the standard General Equilibrium model of Arrow and Debreu to economies with time and uncertainty. This will allow us to justify under some conditions the use of a representative agent.

Review of General Equilibrium

In GE theory, an economy is defined by:

(1) a commodity space, i.e. the set of goods that are available in the economy; mathematically a subset X of some vector space.

(2) a number of consumers $i = 1...I$, each of which has some preferences U_i over consumption bundles in X; each consumer also has an endowment $e_i \in X$ and shares in the technology j, θ_{ij}.

(3) a number of firms $j = 1...J$, each of which has a technology defined by a production set $Y_j \subset X$.

The consumer i's problem is:

$$\max_{x \in X} U_i(x)$$
$$s.t. \quad : \quad p.x \leq p.e_i + \sum_{j=1}^{J} \theta_{ij} \pi_j(p),$$

yielding a net demand $x_i(p)$. (Note that $x_i(p)$ is a vector, the dimension of which is the dimension of the commodity space.)

The firm j's problem is:

$$\pi_j(p) = \max_{y \in Y_j} p.y,$$

yielding a net supply $y_j(p)$.

A *competitive equilibrium* is a price vector p such that:

$$\sum_{i=1}^{I} x_i(p) = \sum_{i=1}^{I} e_i + \sum_{j=1}^{J} y_j(p).$$

i.e.: (i) each consumer solves his program, (ii) each firm solves his program, (iii) all markets clear.

Implicit in this definition is the fact that there are complete markets, i.e. one market for each good, and each consumer can buy or sell whatever amounts he wants in each market, subject only to his resource constraint.

A *feasible allocation* is a list of consumption vectors for each consumer, and of firm plans for each firm, that satisfies the aggregate resource constraint; mathematically $\{x_i\}, \{y_j\}$ such that

$$\sum_{i=1}^{I} x_i = \sum_{i=1}^{I} e_i + \sum_{j=1}^{J} y_j.$$

A *Pareto-optimal* allocation is a feasible allocation such that there is no feasible allocation that makes all consumers as well off, and at least one strictly better off.

An allocation $\{x_i\},\{y_j\}$ is Pareto-optimal allocation iff there exists a set of Pareto weights $\{\lambda_i\}$ with $\lambda_i \geq 0$ for all i and $\lambda_i > 0$ for some i, such that $\{x_i\},\{y_j\}$ solves

$$\max_{\{x_i\},\{y_j\}} \sum_{i=1}^{I} \lambda_i u(x_i) \tag{5.1}$$
$$s.t. \quad : \quad \sum_{i=1}^{I} x_i \leq \sum_{i=1}^{I} e_i + \sum_{j=1}^{J} y_j.$$

I now state the three key results. All rely on complete markets, i.e. there is a market for each good.

(1) First Welfare Theorem: if utility functions are strictly monotonic, any competitive equilibrium is a Pareto optimum.

(2) Second Welfare Theorem: under convexity assumptions[7], any Pareto optimum is a competitive equilibrium for some initial redistribution of wealth.

(3) Existence of equilibrium: under convexity assumptions, there exists at least one competitive equilibrium.

Application to Economies with Time and Uncertainty

To apply this theory to dynamic models, we only need to reinterpret the commodity space. Now a good is defined not only by its physical characteristics, but also by the date and state of nature in which it is available.

Example 1: a two-period consumption/savings problem.

$$\max \{u(c_1) + \beta u(c_2)\}$$
$$s.t. \quad : \quad c_1 + s = y_1$$
$$c_2 \quad = \quad y_2 + (1+r) \times s$$

Define the commodity space, and reframe this as a consumer problem from GE.

Example 2: an infinitely-lived consumer without uncertainty:

$$\max \sum_{t=0}^{\infty} \beta^t u(c_t),$$
$$s.t. \quad : \quad \sum_{t=0}^{\infty} p_t c_t \leq \sum_{t=0}^{\infty} p_t y_t + w_0.$$

Example 3: with uncertainty, the commodity space would include all state-contingent paths: $c_t(s^t)$, where s^t is the history $(s_0, ..., s_t)$.

Application of GE theory to dynamic stochastic models: Complete Markets and Consumption Allocation

Histories

We denote by $s \in S$ the "state of the economy" which keeps track of all the relevant information, e.g. income shocks, news about future income, etc.

Example 1: one agent, his income can be either high or low, then $S = \{high, low\}$ and $s = \{high\}$ or low.

[7] i.e. if utility functions are quasi-concave and production sets are convex.

Example 2: two agents A and B, each of which has an income which can be either high or low, then $S = \{hh, hl, lh, ll\}$.

s_0 is the initial known condition, and s_t is the realized state at time t. We denote a history by $s^t = (s_1, ..., s_t) \in S^t$. The probability of history s^t occuring is denoted $\pi_t(s^t)$. Of course $\sum_{s^t \in S^t} \pi_t(s^t) = 1$ for all s^t.

Example 3: Assume s_t is Markov with transition $P_{s,s'}$. Then $\pi_t(s^t) = P_{s_0 s_1} \times P_{s_1 s_2} \times ... \times P_{s_{t-1} s_t}$.

A basic GE economy to study consumption insurance

(a) commodity space: the set of all possible stochastic processes: $\{c_t(s^t)\}$ all t, s^t.

Example: Simple 3 dates, 2 states example.

(b) preferences: we will restrict preferences considerably, instead of defining them as functions of the entire stochastic processes, i.e. $U(\{c_t(s^t\})$, I will assume that preferences satisfy (1) expected utility [i.e. separability across states of natures], and (2) time-separability:

$$\begin{aligned} U &= E\left[\sum_{t \geq 0} \beta^t u^i(c_t\left(s^t\right) \mid s_0\right], \\ &= \sum_{t \geq 0} \sum_{s^t \in S^t} \beta^t \pi_t(s^t \mid s_0) u\left(c_t\left(s^t\right)\right). \end{aligned} \tag{5.2}$$

There are I agents, indexed by i, with utilities $u^i(.)$, endowments $e_{i,t}(s^t)$. Let $E_t(s^t) = \sum_{i=1}^I e_{i,t}(s^t)$ be the aggregate endowment at time t after history s^t.

(c) technology: this is an endowment economy. i.e. the only technology is free disposal.

Representative consumer

Assume markets are complete. Then any CE is PO. Any PO can be obtained from (5.1). Defining

$$\begin{aligned} V(E) &= \max_{\{x_i\}} \sum_{i=1}^{I} \lambda_i u(x_i) \\ s.t. &: \sum_{i=1}^{I} x_i \leq E, \end{aligned}$$

we see that the economy is equivalent to one where there is only one consumer, with utility function V. This justifies using a representative consumer (RC).

Note: in general the utility function of the RC depends on the weights λ_i, which correspond to the initial wealth distribution. Only if all utility functions are homothetic (i.e., all income elasticities are unity) is the distribution of wealth always irrelevant to find the aggregate allocation. (Of course the wealth distribution is always relevant to compute the individual allocation, but we may not care so much about it.) We will see an example below of representative consumer.

RESULT: if $\{x_i, y_j, p\}$ is an equilibrium for the economy $\{e_i, u_i, Y_j\}$, then $\{\sum x_i, y_j, p\}$ is an eq for the RC economy.

Pareto Optima: Social Planning Problem

I solve for a competitive equilibrium with complete markets, so I can use a planner problem.

$$\max_{\{c_{i,t}(s^t)\}_{i,t,s^t}} \sum_{i=1}^{I} \lambda_i \sum_{t \geq 0} \sum_{s^t \in S^t} \beta^t \pi_t(s^t \mid s_0) u^i\left(c_{i,t}\left(s^t\right)\right)$$

$$s.t. : \forall t, s^t : \sum_{i=1}^{I} c_{i,t}(s^t) = \sum_{i=1}^{I} e_{i,t}(s^t) = E_t(s^t).$$

The FOC w.r.t. $c_{i,t}(s^t)$ is (denoting $\mu_t(s^t)$ the Lagrange multiplier on the resource constraint):

$$\lambda_i \beta^t \pi_t(s^t) u_i'(c_{i,t}(s^t)) = \mu_t(s^t),$$

thus for any agent $i = 2...I$

$$\frac{\lambda_i}{\lambda_1} \frac{u_i'(c_{it}(s^t))}{u_1'(c_{1t}(s^t))} = 1$$

$$\rightarrow c_{it}(s^t) = u_i'^{-1}\left(\frac{\lambda_1}{\lambda_i} u_1'(c_{1t}(s^t))\right)$$

Plugging in the RC yields $c_{1t}(s^t)$. Clearly $c_{1t}(s^t)$ depends only on $E_t(s^t)$, not on the distribution of this endowment across agents $\{e_{i,t}\}$. This is the key result, stated in:

Perfect Risk Sharing and History-Independence

Under complete markets and expected discounted utility, there is perfect risk-sharing: the consumption of each agent depends only on the aggregate consumption, and not on his individual income. Moreover allocations are independent of history, i.e. of the sequence of realization of shocks s^t, and depends only on the current realization of the endowment $E_t(s^t)$.

Decentralization: Competitive Equilibrium

We can decentralize this planner problem. We need one price for each good, i.e. each time-state combination. Let $q_t(s^t)$ =price at time 0 of unit of consumption at time t in state s^t. A competitive equilibrium is an allocation such that

(1) Taking as given the price vector $\{q_t(s^t)\}$, each agent $i = 1...I$ maximizes his utility function (5.2) subject to the (time-zero) budget constraint:

$$\sum_{t\geq 0} \sum_{s^t \in S^t} q_t(s^t) c_{i,t}\left(s^t\right) \leq \sum_{t\geq 0} \sum_{s^t \in S^t} q_t(s^t) e_{i,t}\left(s^t\right).$$

(2) Markets clear: $\forall t \geq 0, \forall s^t \in S^t$:

$$\sum_{i=1}^{I} c_{i,t}\left(s^t\right) = \sum_{i=1}^{I} e_{i,t}\left(s^t\right) = E_t\left(s^t\right).$$

In class: check that this gives the same allocation.

Representative Consumer: a special case

Assume that all agents have the utility $u(c) = c^{1-\gamma}/(1-\gamma)$. Agent i has endowment $\{e_{i,t}(s^t)\}$. Let the equilibrium allocation be $\{c_{i,t}(s^t)\}$ and the prices $\{q_t(s^t)\}$. Then $\{q_t(s^t)\}$ are equilibrium prices for the representative agent economy, with preferences $u(c) = c^{1-\gamma}/(1-\gamma)$ and endowment $\{\sum_i e_{it}(s^t)\}$.

Proof: to be done in class.

Empirical Tests of Full Consumption Insurance

Cochrane (1991) and Mace (1991) were the first to test this implication with US data. They use the PSID to test for perfect consumption insurance by running regressions of the type

$$\Delta \log C_{i,t} = \alpha + \beta \Delta \log C_t + \gamma Z_{i,t} + \varepsilon_{i,t},$$

and testing for $\gamma = 0$. Z could be individual income growth, an unemployment indicator, a sickness indicator... This literature concludes that individual histories do matter, i.e. $\gamma \neq 0$, so there does not appear to be full insurance in the US.

Townsend (1994) test similarly for perfect insurance in Indian villages. Does an individual's consumption depends on his income, once you control for the village's total consumption. He finds that the complete model is rejected too, but by a narrow margin. Using data from Thailand the rejection is "larger".

Possible Extensions

If utility depends in a separable way of leisure (i.e. $u(c)+v(l)$), or something else (e.g. illness), the consumption predictions are unaffected. On the other hand, because the complete markets model makes marginal utilities of consumptions across agents all proportional across time and states, if *marginal utility of consumption* depends on something else besides consumption, then the complete markets model does not predict that consumption moves in lockstep between agents.

Testing Complete Markets using Asset Prices

Under complete markets, we have a RC whose utility depends only on aggregate consumption. As we'll see in the next class, this imply we can use his utility function to price assets. It has proved very hard to find a utility function that rationalizes facts about asset prices using aggregate consumption only. This is an "indirect" rejection of complete markets.

A word on Standard preferences

Most of macroeconomics is done with preferences that satisfy:

- expected utility;
- constant geometric discounting;
- time-separable;

i.e.

$$U = E\sum_{t\geq 0}\beta^t u(c_t),$$

with u increasing, concave, and satisfying generally the Inada conditions:

$$\lim_{c\to 0} u'(c) = +\infty, \quad \lim_{c\to +\infty} u'(c) = 0.$$

In particular we will most often use the CRRA utility function:

$$\begin{aligned} u(c) &= \frac{c^{1-\gamma}}{1-\gamma}, \text{ for any } \gamma > 0, \gamma \neq 1 \\ u(c) &= \log c, \text{ the limiting case } \gamma = 1). \end{aligned}$$

In this case γ is the coefficient of relative risk aversion, and $1/\gamma$ is the intertemporal elasticity of substitution (IES). Applied work typically uses $\gamma = 1$ or 2.

In some case we will use CARA preferences: $u(c) = -e^{-\gamma c}$, but this is only for convenience.

Extensions: Recent research uses different utility functions, in particular:

- not of the expected utility type, e.g. Epstein and Zin (1989); more generally preferences without the independence axiom; or preferences reflecting an aversion to ambiguity.
- not constant geometric discounting, e.g. Laibson (1997); this however leads to issues with time-consistency;
- not time-separable, e.g. habits (Abel 1990, Campbell and Cochrane 1999).

Handout 5: Asset Pricing

In this handout we take a look at the determination of asset prices. Essentially, we extend the formula that prices equals the present discounted value of dividends. First, I cover a general theoretical (static) setup where some info on prices can be deduced from no-arbitrage restrictions. Second, I look in detail at some asset pricing implications of the Euler equation.

Static Setup with Uncertainty

Consider a one-period world with uncertainty: people make decisions before the uncertainty is realized, then once the uncertainy is realized, consumption takes place. The uncertainty is summarized by a state of nature $s = 1, ...S$. Let π_s be the probability that state s occurs; so $\pi_s \geq 0$ for all s and $\sum_{s=1}^{S} \pi_s = 1$. Consider a consumer i who has an endowment e_{is} in state s. The preferences map consumption in all the states into ex-ante utility: $U_i(c_{i1}, ..., c_{iS})$. (i.e. there is no assumption of expected utility). We now consider two possible market structures:

(1) *complete markets with Arrow-Debreu securities:* there are S securities, security s pays one unit of consumption good if state s occurs, and zero otherwise. This is the setup we discussed in the previous lecture. People can sell their endowment in state-contingent markets, then buy their consumption bundle in state-contingent markets. We know that the resulting equilibrium is Pareto-optimal, and it involves full risk-sharing: the consumption of any agent depends only on aggregate consumption.

When we have complete markets, we can price any security from the prices of the AD assets. For instance, the price of a security which yields one unit for sure is the sum of the prices of all the AD assets: these are two equivalent ways of representing the same payoffs, so they have the same price. Another example would be to price an asset which pays off if aggregate consumption is above a cutoff value. See SL p. 220 for more examples.

(2) *a general security market*: there are L securities; the consumer $i = 1...I$ can buy or sell securities $l = 1...L$ at price q_l before the uncertainty is realized; security l pays out a dividend d_{ls}. His budget constraint is, if θ_{il} is the # of securities l that he buys:

$$\sum_{l=1}^{L} q_l \theta_{il} \leq 0,$$

and his consumption is thereafter $c_{is} = e_{is} + \sum_{l=1}^{L} \theta_{il} d_{ls}$.

We can define a competitive equilibrium in the usual way: a price vector $q \in R^L$ such that:

(i) given q, each consumer maximizes his utility function, i.e. solve:

$$\max_{\{\theta_{il}\}} U_i(c_{i1}, ..., c_{iS})$$

$$s.t. \quad : \quad \forall s : c_{is} = e_{is} + \sum_{l=1}^{L} \theta_{il} d_{ls},$$

$$\text{and} \quad : \quad \sum_{l=1}^{L} q_l \theta_{il} \leq 0.$$

(ii) markets clear, i.e. the total supply of each security equals the number that exist in the economy. Assuming that these assets are in zero net supply:

$$\sum_{i=1}^{I} \theta_{il} = 0 \text{ for any } l = 1, ...L.$$

The intepretation of this important zero net supply market equilibrium is that these assets are issued by some people, and bought by other people, but there is no underlying supply of assets. (Think of people trading bonds.) In some cases there might be an outside supply of assets, e.g. there is a capital stock, or there are equities in positive supply to households.

Let $D = \{d_{ls}\}$ be the $L \times S$ matrix. Can write this problem as pick $\theta \in R^L$ s.t. $c = e + D'\theta$, s.t. $q'\theta \leq 0$.

Result: if $Span(D) = R^S$, then the allocation is the same as with complete markets.

[$Span(D) = R^S$ means that for any $x \in R^S$, there is a $y \in R^L$ st $x = D'\theta$.]

Note: this requires $L \geq S$.

Example: to add in class.

No-Arbitrage Pricing

Definition: the system of securities characterized by $\{d_{ls}\}$ and $\{q_l\}$ is arbitrage-free (or: there is no arbitrage opportunity, NOAO) if there is no vector of portfolio choices $\{\theta_l\}$ such that both

$$(1)\ \sum_{l=1}^{L} q_l\theta_l \leq 0 \text{ (i.e. } \theta \text{ is free, or has a negative price),}$$

and $\forall s = 1...S$, $\sum_{l=1}^{L} d_{ls}\theta_l \geq 0$, and > 0 for some state s (i.e. θ gives positive payoffs in all states).

Clearly, if the prices $\{q_l\}$ result from a competitive equilibrium as in (2) above, it must be that there is NOAO (otherwise consumers would demand an infinite amount of the combination which is free and gives a positive payoff.) But NOAO is more general and as a result requires very few assumptions.

Consequence of no arbitrage: assume you have two assets x and y and that in all possible states, x will pay you more dividends than y. Then the price of x should be (weakly) greater than the price of y. This basic principle imposes restrictions across prices of securities.

In some cases, NOAO allows you to pin down exactly the price of some other securities. Even w/o complete markets, assets that are *redundant*, i.e. which can be replicated by existing assets, can be priced if we assume no arbitrage.

Example: static economy with 7 possible states, 3 assets are traded, payoff matrix listing in each row the payoff of each asset in each of the states:

$$D = \begin{pmatrix} 1 & 2 & 1 & 0 & 0 & 0 & 0 \\ 1 & 3 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

the third asset = the second asset minus the first one. Thus its price is the difference of the prices of the first two assets.

An important practical application of NOAO is option pricing: Black and Scholes noted that the payoff of an option is (roughly, for small time intervals) a linear combination of the payoff of a stock and a bond, hence its value can be deduced from these two prices.

Euler equations

Consider a savings/portfolio choice model: the agent maximizes utility $E\sum_{t\geq 0}\beta^t u(c_t)$ by choosing consumption, and the number of share $\theta_{i,t}$ of each asset $i = 1...N$ to hold, subject to the sequence of budget constraints (explain in class):

$$y_t + \sum_{i=1}^{N} (P_{it} + D_{it})\,\theta_{it-1} = c_t + \sum_{i=1}^{N} P_{it}\theta_{it}.$$

The FOC w.r.t $\theta_{i,t}$ yields:

$$\begin{aligned} u'(c_t)P_{i,t} &= E_t\left(\beta u'(c_{t+1})\left(P_{i,t+1} + D_{i,t+1}\right)\right) \\ u'(c_t) &= E_t\left(\beta u'(c_{t+1})R^i_{t\to t+1}\right). \end{aligned}$$

where $R^i_{t\to t+1} = \frac{P_{i,t+1}+D_{i,t+1}}{P_{i,t}}$ is the gross return in any asset i available to the investor at time t (this assumes an interior solution, i.e. no binding short-sales constraint, etc.).

In finance this equation is often written

$$E_t\left(m_{t,t+1}R^i_{t\to t+1}\right) = 1,$$

for any asset i, where m is the *stochastic discount factor* which in our case is simply the IMRS (Intertemporal Marginal Rate of Substitution) of our representative consumer:

$$m_{t,t+1} = \frac{\beta u'(c_{t+1})}{u'(c_t)}.$$

The rest of this handout (and indeed a good deal of finance) plays with this Euler equation, which is a *restriction on the joint behavior of asset returns and consumption.*[8]

While this EE can be applied to individual data on consumption and asset returns, we will apply it mostly to aggregate data, i.e. we assume a representative agent. As you know from the previous class, this can be justified by complete markets. At least this gives us some intuition for the key forces that matter in the aggregate.

Technical notes: (1) Can also derive the equations from a BE:

$$\begin{aligned} V(W) &= \max_{c,\theta_i}\left\{u(c) + \beta EV(W')\right\}, \\ s.t. &: W' = \sum_i \theta_i R'_i \times (W - c). \end{aligned}$$

(2) and you can also derive the Euler equation from a full-fledged state-contingent sequence problem: choose $c_t(s^t), \theta_{it}(s^t)$, and so on.

Deterministic Euler Equation

First assume there is no uncertainty, then the Euler equation reads

$$\frac{\beta u'(c_{t+1})}{u'(c_t)} = \frac{1}{R_{t\to t+1}},$$

and assuming $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$ I get (taking logs)

$$\Delta \log c_{t+1} = \frac{\log R_{t\to t+1} + \log \beta}{\gamma} \simeq \frac{r_{t\to t+1} - \rho}{\gamma},$$

where $R = 1 + r$ and $\beta = e^{-\rho} \simeq \frac{1}{1+\rho}$.

[8] This relation is an equilibrium condition, and cannot be interpreted in a causal way (i.e., you can say either "interest rates are high because consumption growth is high", but you can equally well say the reverse).

This says that consumption growth is high when the interest rate is high, when the preference for the present is low, and when the IES $\frac{1}{\gamma}$ is high (assuming $r_{t\to t+1} > \rho$).

Agents all other things equal prefer a smooth path of consumption, both across time and across states, because of the concavity of the utility function u; however they are willing to accept some drift in consumption, if compensated by a high enough return. The return required to make consumption growth move depends on the curvature of the utility function u, i.e. the IES $\frac{1}{\gamma}$. If $\gamma = 0$, $u(c) = c$, the IES is infinite, and agents are willing to accept a lot of variation in consumption. If $\gamma \to \infty$, the IES is very low, and agents are adverse to consumption fluctuations. (To this day, the empirical debate about the strength of the IES persists.)

Note that under certainty all assets must have the same rate of return: for any asset i, taking out the expectations:

$$\frac{\beta u'(c_{t+1})}{u'(c_t)} = \frac{1}{R^i_{t\to t+1}} = \frac{1}{R^f_{t\to t+1}}.$$

Uncertainty: (1) Risk-free rate pricing

If R^f is an asset which return is known in advance (i.e. a risk-free asset), we have

$$\frac{1}{R^f} = E_t\left(\beta\left(\frac{c_{t+1}}{c_t}\right)^{-\gamma}\right).$$

Assume that $\Delta\log c_{t+1}$ is conditionally normal with mean $\mu = E_t\Delta\log c_{t+1}$ and conditional variance $\sigma^2 = V_t\Delta\log c_{t+1}$, I have, using the log-normal formula[9]:

$$\begin{aligned}\frac{1}{R^f} &= \frac{1}{1+r^f}\\ &= E_t e^{\log\beta-\gamma\Delta\log c_{t+1}}\\ &= e^{\log\beta-\gamma E_t\Delta\log c_{t+1}+\frac{\gamma^2}{2}V_t\Delta\log c_{t+1}}\\ &= 1+\log\beta-\gamma E_t\Delta\log c_{t+1}+\frac{\gamma^2}{2}V_t\Delta\log c_{t+1}\end{aligned}$$

Thus, if $\beta = e^{-\rho}$:

$$r^f \simeq \rho + \gamma E_t\left(\Delta\log c_{t+1}\right) - \frac{\gamma^2}{2}V_t\left(\Delta\log c_{t+1}\right).$$

A first important result.

Risk free rates can be high because:

- agents are impatient and want to consume now, so they borrow and push up the interest rate;

- agents expect consumption to be high in the future, so depending on the size of γ they borrow, pushing up the interest rate;

- agents expect little uncertainty about the future, so they do not engage in precautionary savings, pushing the last term up. Note this terms is weighted by γ^2 (for reasons that will become clear later.)

Uncertainty: (2) The Consumption Capital Asset Pricing Model (CCAPM)

[9] Important fact to remember: if X is normal (μ, σ^2), then $E\exp X = \exp(\mu + \sigma^2/2)$.

Assume now that $\left(\Delta \log c_{t+1}, \log R_{t+1}^{i}\right)$ is jointly normal. Then we can compute the (log) risk-premium exactly if utility is CRRA:

$$\begin{aligned}
1 &= E_t\left(\beta\left(\frac{c_{t+1}}{c_t}\right)^{-\gamma} R_{t+1}^{i}\right) \\
1 &= E_t e^{\log\beta-\gamma\Delta\log c_{t+1}+\log R_{t+1}^{i}} \\
1 &= e^{\log\beta-\gamma E_t\Delta\log c_{t+1}+E_t\log R_{t+1}^{i}+\frac{\gamma^2}{2}V_t\Delta\log c_{t+1}+\frac{1}{2}V_t\log R_{t+1}^{i}-\gamma Cov\left(\Delta\log c_{t+1},\log R_{t+1}^{i}\right)} \\
E_t\log R_{t+1}^{i} &= -\log\beta+\gamma E_t\Delta\log c_{t+1}-\frac{\gamma^2}{2}V_t\Delta\log c_{t+1}-\frac{1}{2}V_t\log R_{t+1}^{i}+\gamma Cov\left(\Delta\log c_{t+1},\log R_{t+1}^{i}\right)
\end{aligned}$$

V_t is the conditional variance: $V_t X_{t+1} = E_t X_{t+1}^2 - \left(E_t X_{t+1}\right)^2$.

Taking the difference between this last equation and the equation that gives us the risk-free rate, this yields:

$$E_t \log\left(\frac{R_{t+1}^{i}}{R_{t+1}^{f}}\right) + \frac{1}{2}V_t\left(\log R_{t+1}^{i}\right) = \gamma Cov\left(\Delta\log c_{t+1},\log R_{t+1}^{i}\right)$$

which can be also be rewritten:

$$\log\left(\frac{E_t R_{t+1}^{i}}{R_{t+1}^{f}}\right) = \gamma Cov\left(\Delta\log c_{t+1},\log R_{t+1}^{i}\right)$$

This equation is the *Consumer Capital Asset Pricing Model* (CCAPM). It states that the (log) expected return of an asset i is greater than the risk-free rate if and only if the asset return has a positive covariance with consumption growth. The size of the risk premium is given by the risk aversion γ times the covariance. This follows from the logic of insurance against aggregate (or macroeconomic) risk. Agents want assets that pay off (i.e. have high returns) when their consumption is low, because that's when marginal utility is largest. Hence they are willing to forego higher expected returns to get such assets. Inversely, assets that pay off when consumption is high have to pay a high expected return to compensate for the high risk. (Of course, because there is risk, these expectations will not be fulfilled in every period.)

Note 1: the variance of the asset return does not matter in general.

Note 2: idiosyncratic risk does not matter: if the asset return has some noise that is uncorrelated with aggregate consumption growth: $\log R_{t+1}^{i} + \varepsilon_{i,t}$, this does not change the expected return $E_t R_{t+1}^{i}$.

Finally, note that this condition is a statement about *equilibrium*: assets that have more risk (as measured by the covariance with consumption growth) must have higher expected returns, to compensate for the risk: otherwise investors would not hold them. Hence "bad" assets that are risky must have *high* expected returns.

Another way to state the CCAPM is

$$\begin{aligned}
E\left(r^i - r^f\right) &\approx \log\left(\frac{E\left(R^i\right)}{R^f}\right) = \beta^i \times \lambda, \\
\lambda &= \gamma V\left(\Delta\log c_{t+1}\right), \\
\beta^i &= \frac{Cov\left(\Delta\log c_{t+1},\log R_{t+1}^{i}\right)}{V\left(\Delta\log c_{t+1}\right)}.
\end{aligned}$$

$\lambda =$ "market price of risk" $=$ risk aversion times the variance of consumption growth

β^i is the "consumption beta" = slope coefficient of the time-series regression of asset return on consumption growth:

$$R^i_{t\to t+1} = \alpha_i + \beta_i \Delta \log c_{t+1} + \varepsilon_{i,t+1}.$$

The CAPM (Capital Asset Pricing Model)

You may be familiar with CAPM, which if often stated as

$$\begin{aligned} E\left(r^i - r^f\right) &= \beta^i \times \lambda, \\ \lambda &= E(R^m - R^f) \\ \beta^i &= \frac{Cov\left(R^m, R^i\right)}{Var\left(R^m\right)}, \end{aligned}$$

where R^m is the "market return" i.e. the return on the entire wealth of the agent. Here is one possible derivation of the CAPM. Write the budget constraint as $W_{t+1} = R^m_{t+1}(W_t - c_t)$ where R^m is the return on all the wealth (physical and human) of the agent. The Bellman equation is, assuming that the return is iid :

$$v(W) = \max_c \left\{u(c) + \beta EV\left(R^m(W-c)\right)\right\}.$$

If utility is quadratic, say $u(c) = -\frac{1}{2}c^2$, then a simple guess and verify method shows that V is quadratic too, and consumption is proportional to wealth: $c = \mu W$. Hence the Euler equation is now

$$\begin{aligned} E\left(\frac{\beta c_{t+1}}{c_t}\left(R^i_{t+1} - R^f_{t+1}\right)\right) &= 0 \\ E\left(\frac{\beta \mu W_{t+1}}{\mu W_t}\left(R^i_{t+1} - R^f_{t+1}\right)\right) &= 0 \\ E\left(\beta R^m_{t+1}(1-\mu)\left(R^i_{t+1} - R^f_{t+1}\right)\right) &= 0 \\ E\left(R^m_{t+1}\left(R^i_{t+1} - R^f_{t+1}\right)\right) &= 0 \end{aligned}$$

which implies

$$\begin{aligned} E\left(R^m_{t+1}R^i_{t+1}\right) &= E\left(R^m_{t+1}\right)R^f_{t+1} \\ Cov\left(R^m_{t+1}, R^i_{t+1}\right) &= E\left(R^m_{t+1}\right)\left(R^f_{t+1} - E\left(R^i_{t+1}\right)\right) \\ E\left(R^i_{t+1}\right) - R^f_{t+1} &= -\frac{Cov\left(R^m_{t+1}, R^i_{t+1}\right)}{Var\left(R^m_{t+1}\right)}\frac{Var\left(R^m_{t+1}\right)}{E\left(R^m_{t+1}\right)}. \end{aligned} \tag{5.3}$$

This relation holds for $i = m$, the market return, thus

$$E\left(R^m_{t+1}\right) - R^f_{t+1} = -\frac{Var\left(R^m_{t+1}\right)}{Var\left(R^m_{t+1}\right)}\frac{Var\left(R^m_{t+1}\right)}{E\left(R^m_{t+1}\right)} = -\frac{Var\left(R^m_{t+1}\right)}{E\left(R^m_{t+1}\right)}.$$

This implies I can rewrite (5.3) as

$$E\left(R^i_{t+1}\right) - R^f_{t+1} = \frac{Cov\left(R^m_{t+1}, R^i_{t+1}\right)}{Var\left(R^m_{t+1}\right)}\left(E\left(R^m_{t+1}\right) - R^f_{t+1}\right),$$

yielding the CAPM. (There are other possible derivations of the CAPM.)

Note that $\beta^i = \frac{Cov\left(R^m, R^i\right)}{Var(R^m)}$ is now the slope of a time-series regression of the stock return on the market return, rather than consumption growth:

$$R^i_{t\to t+1} = \alpha + \beta_i R^m_{t\to t+1} + \varepsilon_{i,t+1},$$

and the "market price of risk" $\lambda = E\left(R^m_{t+1}\right) - R^f_{t+1}$ can be measured in the data without a reference to risk aversion. The CAPM is much more popular outside academia.

Price = Present Value of Dividends, and the Random walk theory of stock prices

Since $R_{t\to t+1} = \frac{P_{t+1}+D_{t+1}}{P_t}$, I can rewrite the Euler equation as:

$$P_t = E_t\left(\frac{\beta u'(c_{t+1})}{u'(c_t)}\left(P_{t+1} + D_{t+1}\right)\right).$$

Iterating forward yields "price equals the present *discounted* value of dividends":

$$P_t = E_t\left(\sum_{j\geq 1} \frac{\beta^j u'(c_{t+j})}{u'(c_t)} D_{t+j}\right).$$

For empirical work dividing by D_t yields a stationary series, which is better:

$$\frac{P_t}{D_t} = E_t\left(\sum_{j\geq 1} \frac{\beta^j u'(c_{t+j})}{u'(c_t)} \frac{D_{t+j}}{D_t}\right).$$

Note that a special case of this formula is if $u(c) = c$, i.e. the agent is risk-neutral. Under this condition we have

$$P_t = E_t \sum_{j\geq 1} \beta^j D_{t+j},$$

i.e. price = present value of dividends with a *constant* discount rate β. In the general case when u is concave (risk-aversion), the future dividends must be discounted not only because they are paid in the future (the β part) but also because they may be paid in periods when marginal utility is high (the $u'(c_{t+j})$ part).

If utility is linear (i.e. risk-neutrality), then we get that

$$E_t\left(R_{t+1}\right) = \frac{1}{\beta},$$

so the expected return on any asset is constant across assets and across time (since agents do not care about risk). In particular, for a small interval of time, without dividends, and with $\beta \simeq 1$, we have

$$P_t \simeq E_t\left(P_{t+1}\right),$$

the *random walk theory of stock prices*. You can see the heavy assumptions required to obtain this: risk neutrality and a short horizon. The RW theory works well at very short horizon, but it is violated (as we will see, you can forecast stock prices or returns to some extent.)

Solving for the price of an asset with $D_t = C_t$

Start from the general formula:

$$\frac{P_t}{D_t} = E_t\left(\frac{\beta u'(c_{t+1})}{u'(c_t)}\left(\frac{P_{t+1}}{D_{t+1}} + 1\right)\frac{D_{t+1}}{D_t}\right)$$

Specializing to our case of CRRA utility and $c_t = D_t$:

$$\frac{P_t}{c_t} = E_t\left(\beta\left(\frac{c_{t+1}}{c_t}\right)^{1-\gamma}\left(\frac{P_{t+1}}{c_{t+1}}+1\right)\right).$$

Case A: consumption growth is log-normal and iid.

Assumption:

$$\log\left(\frac{c_{t+1}}{c_t}\right) \rightsquigarrow N(\mu,\sigma^2)$$

Let's guess and verify that $\frac{P_t}{D_t}$ is constant equal to q. (Can prove this: price = PDV of future consumption growth, since consumption growth is iid this conditional expectation is constant..)

$$\begin{aligned} q &= \beta\left(1+q\right)E_t\left(\frac{c_{t+1}}{c_t}\right)^{1-\gamma} = \beta(1+q)e^{(1-\gamma)\mu+\frac{(1-\gamma)^2}{2}\sigma^2} \\ \implies q &= \frac{\beta e^{(1-\gamma)\mu+\frac{(1-\gamma)^2}{2}\sigma^2}}{1-\beta e^{(1-\gamma)\mu+\frac{(1-\gamma)^2}{2}\sigma^2}}. \end{aligned}$$

Expected return on equity:

$$\begin{aligned} E_t\left(R_{t\to t+1}\right) &= E_t\left(\frac{P_{t+1}+C_{t+1}}{P_t}\right) \\ &= E_t\left(\frac{\left(\frac{P_{t+1}}{C_{t+1}}+1\right)}{\frac{P_t}{C_t}}\frac{C_{t+1}}{C_t}\right) \\ &= \frac{1+q}{q}e^{\mu+\frac{\sigma^2}{2}} \\ &= \frac{1}{\beta}e^{\gamma\mu+\frac{2\gamma-\gamma^2}{2}\sigma^2} \end{aligned}$$

Volatility of Equity Return:

$$R_{t\to t+1} = \frac{P_{t+1}+C_{t+1}}{P_t} = \frac{q+1}{q}\frac{C_{t+1}}{C_t},$$

so the volatility of the ex-post return is equal to $\frac{q+1}{q}$ times the volatility of consumption growth. In practice q is the price-dividend ratio, which is about 20 or 25, so $\frac{q+1}{q}$ is approximately 1 and returns are as volatile as dividend growth (= consumption growth in this case).

Risk-free rate:

$$R^f_{t\to t+1} = \frac{1}{E_t\left(\beta\left(\frac{c_{t+1}}{c_t}\right)^{-\gamma}\right)} = \frac{1}{\beta}e^{\gamma\mu-\frac{\gamma^2}{2}\sigma^2}$$

(Geometric) Risk premium:

$$\begin{aligned} \frac{E_t\left(R_{t\to t+1}\right)}{R^f_{t\to t+1}} &= \frac{\frac{1}{\beta}e^{\gamma\mu+\frac{2\gamma-\gamma^2}{2}\sigma^2}}{\frac{1}{\beta}e^{\gamma\mu-\frac{\gamma^2}{2}\sigma^2}} = e^{\gamma\sigma^2} \\ E_t\left(R_{t\to t+1}\right)-R^f_{t\to t+1} &\simeq \gamma\sigma^2. \end{aligned}$$

In US annual data, σ is about 2%, so the risk premium should be 0.04% with log utility. To put it another way, to match an average LHS of 6% per year, we need $\gamma = 6 \times 25 = 150$. This is the "equity premium puzzle": given the smoothness of aggregate consumption, we need a very high (implausible) coefficient of risk aversion to match the risk premia we see in the data.[10] Our number actually underestimates the problem because in the data, dividends are not perfectly correlated with consumption.

Case B: Mehra-Prescott 1985: consumption growth is a Markov chain.

This is more general: assume $g_t = \frac{c_{t+1}}{c_t}$ shifts between N states $g_t = \lambda_1, ..., \lambda_N$ according to the $N \times N$ transition matrix $P = (P_{i,j})_{i,j=1...N}$.

Iterating forward shows that P_t/c_t depends only on future consumption growth. Since consumption growth is Markov, it depends only on today's consumption growth: $\frac{P_t}{c_t} = \Phi(\lambda)$ where λ is the current state.

Substituting gives an equation for Φ :

$$\forall i = 1...N : \Phi(\lambda_i) = \sum_{j=1}^{N} P_{i,j} \beta \lambda_j^{1-\gamma} \left(\Phi\left(\lambda_j\right) + 1\right).$$

This is a linear system of N equations in N unknowns $\Phi\left(\lambda_i\right)$. Mehra and Prescott choose a simple Markov chain to represent US consumption growth and solve this model.

Exercise/Reading: write the formulas for the expected equity return and the expected risk-free rate, given today state is i; also give the formulas for the unconditional averages (using the invariant distribution).

Some empirical puzzles

It is interesting to list the empirical failures of the standard model (i.e. CRRA utility and iid consumption growth):

(1) equity premium too low for reasonable risk aversion ("equity premium puzzle");

(2) if I increase the risk aversion γ to match the equity premium, I need $\beta > 1$ to match the risk-free rate ("risk-free rate puzzle"), which sounds strange (but is possible);

(3) the stock return volatility is too low in the model (data: $\sigma(R) = 17\%$ and model: $\sigma(R) = \sigma(\Delta \log C) = 2\%$).

(4) constant price-dividend ratio in the model;

(5) the risk premium on equity is constant over time; this contrasts with the fact that returns are forecastable in regressions of the type:

$$R^e_{t \to t+k} = \alpha + \beta x_t + \varepsilon_{t+k},$$

where x_t is a variable known at time t.

(6) failure in cross-section data: the CCAPM doesn't explain well average returns by consumption betas, as it should if $E\left(R_i - R_f\right) = \beta_i \lambda$.

Euler Equation tests

One important empirical test is (using either individual or aggregate data) whether the Euler equation holds, i.e. test for

$$E_t \left(\beta \left(\frac{c_{t+1}}{c_t} \right)^{-\gamma} \left(R^i_{t \to t+1} - R^f_{t \to t+1} \right) \right) = 0.$$

[10] Note the similarity of the equity premium formula with the welfare cost of business cycles formula of Lucas. This motivated Alvarez and Jermann (2004) to infer the welfare cost of business cycles from asset price data.

This can be done by first multiplying by any time t variable z_t; and then simply taking averages, and finding the parameters that make these sample moments

$$\frac{1}{T}\sum_{t=1}^{T}\left(z_t\beta\left(\frac{c_{t+1}}{c_t}\right)^{-\gamma}\left(R^i_{t\to t+1}-R^f_{t\to t+1}\right)\right)$$

as close to zero as possible.. That's GMM as in Hansen and Singleton (1983), who also reject the CRRA utility model.

Extensions

(1) Dividends do not equal consumption (Easy, you can modify some of the examples yourself).

(2) Changing the utility function, but keep a representative agent. E.g. habits: $u(c_t - \theta h_t)$ with $h_t = (1-\delta)h_{t-1} + \delta c_{t-1}$; or Epstein-Zin utility; or nonseparable utility with leisure, durables, housing...

(3) Incomplete markets. Much more complicated: need to compute the equilibrium when people can trade some assets, but cannot insure their consumption. Much work on this.

(4) Catastrophic events, i.e. a small probability of very large negative events, ...

Handout 4: Welfare cost of business cycles

Lucas (1987) asked the question, How much is the US representative agent willing to pay to get rid of business cycles? This is an upper bound on the benefits of stabilization policies.

Assume a standard (= expected discounted time-separable CRRA) utility function:

$$E\sum_{t\geq 0}\beta^t\frac{c_t^{1-\gamma}}{1-\gamma},$$

and assume that consumption grows along a deterministic trend, plus some stochastic fluctuations:

$$\log c_t = \mu t - \frac{\sigma^2}{2} + \varepsilon_t,$$

with ε_t normal $(0,\sigma^2)$; the constant $-\sigma^2/2$ is added so that $E(c_t) = e^{\mu t}$ for all $t \geq 0$.

Note 1: we do not make the assumption that ε_t is iid; for instance it could be an AR(1) process.

Note 2: this assumption of a deterministic trend was still standard at the time that Lucas wrote, but most people now prefer to use a random walk as benchmark for the time series process of consumption: $\Delta\log c_t = \mu + \sigma\varepsilon_t$. See your problem set for more on this.

Recall the log-normal formula (you *have to* know this): if X is $N(\mu,\sigma^2)$, then $Ee^X = e^{\mu+\frac{\sigma^2}{2}}$. Thus:

$$Ec_t^{1-\gamma} = Ee^{(1-\gamma)\left(\mu t-\frac{\sigma^2}{2}+\varepsilon_t\right)} = e^{(1-\gamma)\left(\mu t-\frac{\sigma^2}{2}\right)+\frac{(1-\gamma)^2}{2}\sigma^2} = e^{(1-\gamma)\mu t-\frac{\gamma(1-\gamma)}{2}\sigma^2}$$

The expected utility as of time 0 is:

$$\sum_{t\geq 0}\beta^t\frac{Ec_t^{1-\gamma}}{1-\gamma} = \frac{e^{-\frac{\gamma(1-\gamma)}{2}\sigma^2}}{1-\gamma}\frac{1}{1-\beta e^{(1-\gamma)\mu}}.$$

One way to measure the willingness to pay is to ask: what % of consumption *at all future dates* are you willing to give up to get rid of the volatility ε_t? i.e. what is the η such that $c_t = \widetilde{c}_t = (1-\eta)\exp(\mu t)$ *for sure* and

$$\sum_{t\geq 0}\beta^t\frac{\widetilde{c}_t^{1-\gamma}}{1-\gamma} = \frac{e^{-\frac{\gamma(1-\gamma)}{2}\sigma^2}}{1-\gamma}\frac{1}{1-\beta e^{(1-\gamma)\mu}}$$

Simple algebra yields:

$$\frac{(1-\eta)^{1-\gamma}}{1-\gamma}\frac{1}{1-\beta e^{(1-\gamma)\mu}} = \frac{e^{-\frac{\gamma(1-\gamma)}{2}\sigma^2}}{1-\gamma}\frac{1}{1-\beta e^{(1-\gamma)\mu}}$$

$$e^{-\frac{\gamma(1-\gamma)}{2}\sigma^2} = (1-\eta)^{1-\gamma}$$

$$-\gamma\frac{(1-\gamma)^2\sigma^2}{2} = (1-\gamma)\log(1-\eta)$$

$$\eta \cong \frac{\gamma\sigma^2}{2}$$

Key result: Welfare cost = risk aversion × variance of consumption "shock" ε_t over 2.

For the US economy, σ is estimated to be small, about 2% = 0.02 for annual data. Thus

$$\eta \cong \frac{\gamma \times 0.0004}{2},$$

and for a γ of say 1 (log utility), we have $\eta \cong 0.02\%$: the US representative consumer is willing to give up only 0.02% of his consumption every year to get rid of fluctuations. This appears small; it is at odds with the view that recessions are "bad events".

This result takes as given the observed fluctuations in consumption, and does not need to ask where the volatility ε_t comes from (monetary shocks, or technology shocks, or...).

Note that in contrast, if we can increase the trend growth rate, the welfare benefit is very large. Let's do the similar computation for a change from μ_0 to $\mu > \mu_0$:

$$\frac{(1-\eta)^{1-\gamma}}{1-\gamma}\frac{1}{1-\beta e^{(1-\gamma)\mu}} = \frac{1}{1-\gamma}\frac{1}{1-\beta e^{(1-\gamma)\mu_0}}$$

$$\begin{aligned}
(1-\eta)^{1-\gamma} &= \frac{1-\beta e^{(1-\gamma)\mu}}{1-\beta e^{(1-\gamma)\mu_0}} \\
(1-\gamma)\log(1-\eta) &\simeq \log\left(\frac{1-\beta e^{(1-\gamma)\mu}}{1-\beta e^{(1-\gamma)\mu_0}}\right) \\
(1-\gamma)\log(1-\eta) &\simeq \log\left(1+\frac{\beta\left(e^{(1-\gamma)\mu_0}-e^{(1-\gamma)\mu}\right)}{1-\beta e^{(1-\gamma)\mu_0}}\right) \\
(1-\gamma)\log(1-\eta) &\simeq \frac{\beta\left(\mu_0-\mu\right)(1-\gamma)e^{(1-\gamma)\mu_0}}{1-\beta e^{(1-\gamma)\mu_0}} \\
\eta &\simeq (\mu-\mu_0)\frac{\beta e^{(1-\gamma)\mu_0}}{1-\beta e^{(1-\gamma)\mu_0}}
\end{aligned}$$

How much consumption do you give up to go from $\mu_0 = 0.02$ to $\mu = 0.03$? Example: $\gamma = 1$,

$$\eta \simeq (\mu-\mu_0)\frac{\beta}{1-\beta}$$

so if $\beta = 0.95$ then $\eta \simeq 19\%$. A "small" difference in the growth rate matters a lot.

Two key criticisms of the business cycle calculation:

(1) The representative agent assumption: it is justified with complete markets. But markets are incomplete, and it may be that some agents suffer a lot from recessions, even though most agents don't. Two issues:

- The welfare cost may be unequally distributed, with some $fixed$ categories of agents (e.g. unskilled) suffering more from every recession;

- It could also be that in any given recession, a few agents ("randomly" chosen, i.e. not a fixed category) would suffer a lot. Since risk-averse agents dislike "big shocks" this would make the welfare cost bigger.

(2) As we will see later the CRRA utility function has problems to match the risk premium on equity, which appears to be closely related to the welfare cost. This should make us dubious of this result.

Hence the Lucas calculation has started a large literature (see Lucas 2003 AER for a review):

- what if we use a different utility function?

- what is we have a different consumption process?

- what if we have incomplete markets?

(Note: this method of measuring welfare costs can also be applied to measuring the welfare effect of taxes.)

Handout 6: Permanent Income Hypothesis

This is the first incomplete market model we see in this class. The motivation is that the complete markets model with CRRA utility is rejected: it implies full consumption insurance, which does not appear to be true in consumption data, and asset prices data similarly suggests a rejection of this model. Hence we need to think of models where agents cannot insulate their consumption from their income completely. One extreme example is 'autarky' i.e. your income equals your consumption. A more interesting example is the PIH: people can save or borrow in a risk-free asset (a 'savings account'), but they cannot buy or sell state-contingent commodities (Arrow-Debreu assets).

Two key assumptions up front:

- The agent can borrow or save at gross risk-free rate $R = 1 + r$. This is the only asset; in particular there is no asset which can insure him against his labor income risk.

- We impose only the present-value budget constraint. Or to put it another way, we do not impose that consumption must be positive.[11] This means that the consumer can, and will always repay (asymptotically), if necessary by making his consumption negative.

Consumer problem:

$$\max_{\{c_t, b_{t+1}\}_{t=0}^{\infty}} E \sum_{t \geq 0} \beta^t U(c_t) \tag{5.4}$$

$$s.t. \quad : \quad b_{t+1} = Rb_t + y_t - c_t, \tag{5.5}$$

$$b_0 \text{ given.}$$

b_t = asset at end of period $t-1$. (Be careful about the timing, some authors use a different one.) Iterating on the budget constraint yields the present-value budget constraint:

$$\sum_{j \geq 0} \frac{c_j}{R^j} = Rb_0 + \sum_{j \geq 0} \frac{y_j}{R^j}. \tag{5.6}$$

Euler equation (assuming interior solution):

$$u'(c_t) = R\beta E_t u'(c_{t+1}). \tag{5.7}$$

Two additional assumptions now:

- $R\beta = 1$;
- utility is quadratic, i.e.

$$u(c) = -\frac{1}{2}(c - c^*)^2.$$

Valid only for $c < c^*$.

Then we get the (certainty-equivalent) *PIH*:

$$\mathbf{c}_t = \mathbf{E}_t \mathbf{c}_{t+1}. \tag{5.8}$$

[11] As we will see later in the class, imposing that consumption must be positive puts a bound on the maximum amount the agent can borrow, since he must repay *for sure* (it is a *risk-free* asset), so there is always a borrowing constraint.

The first key result. Comments:

(a) The underlying idea is consumption smoothing: the consumer wants to have a smooth flow of consumption, so he keeps his consumption on average flat (since $R\beta = 1$, no reason to "tilt" the path of consumption).

(b) Consumption is a martingale. in particular there is no way to forecast next period's consumption better than using simply current consumption.

(c) Consumption reacts on the *news*: agents are forward looking so they form expectations about their future income, and increase their consumption by the expected amount as soon as they learn about some future extra income.

Take expectations of the budget constraint (5.6) and use (5.8) to derive the consumption function:

$$\begin{aligned} c_t &= (R-1)\, b_t + \frac{R-1}{R} E_t \left(\sum_{j \geq 0} \frac{y_{t+j}}{R^j} \right), \\ &= r \left(b_t + \frac{1}{1+r} E_t \left(\sum_{j \geq 0} \frac{y_{t+j}}{R^j} \right) \right). \end{aligned} \tag{5.9}$$

The right-hand side is the "permanent income" of the consumer: the interest on (physical) wealth plus the interest on future expected human wealth. This can be succintly written:

$$c_t = rW_t = r(b_t + H_t).$$

Uncertaintly about future income does not affect the consumption decision, which depends only on the mean of future income (i.e., the variance of future income doesn't play a role). There is *certainty equivalence*, i.e. agents' behavior is the same, whatever the amount of uncertainty they face.[12]

Friedman viewed permanent income as stable (i.e. W_t is smooth), which explained that consumption, which is proportional to permanent income, is relatively smooth, and certainly smoother than income[13]:

$$\sigma(c_t) < \sigma(y_t).$$

Implicitely, this assumes that if income increases today, it's only transitory, so the consumer will save most of it. One way to think about it is in terms of MPC =marginal propensity to consume. The MPC out of wealth, i.e. the increase in consumption if your wealth increases by 1\$ (a windfall), is r.

On the other hand, if your income increases by 1\$ today and at all future dates, the present value of your income increases by $\frac{1}{r}$ and your consumption increases by $r \times \frac{1}{r} = 1\$$. This shows that consumption reacts differentlty to a permanent shock than to a transitory shock. Implicitely Friedman assumed that a good deal of the shocks where transitory, so that wealth and permanent income (and thus consumption) were smoother than current income.

[12] Mechanically, this is because decision rules are linear (because the objective function is quadratic and the constraints are linear), so you can take expectations in the decision rules and see that only means (i.e. averages, or expectations) matter, and not higher moments.

[13] You need to take away some trend from both c_t and y_t before computing this standard deviation. E.g. look at $\Delta \log c_t$ vs. $\Delta \log y_t$.

In his book he showed how this kind of reasoning (though both more elaborate and less mathematically precise) could explain a wide variery of cross-section and time-series facts: the difference in the slope of farmers vs. non farmers consumption function as a function of current income; or why blacks do not "oversave"; etc. (see below).

Apply the operator $E_t - E_{t-1}$ to (5.9), this cancels all the terms known at $t-1$, hence:

$$c_t - E_{t-1}c_t = c_t - c_{t-1} = \frac{R-1}{R}\left(E_t - E_{t-1}\right)\left(\sum_{j\geq 0}\frac{y_{t+j}}{R^j}\right). \tag{5.10}$$

This shows that the innovation (=unexpected movement) in consumption is given by the innovation in the present discounted value of future income, aka the revision in current and future labor income.

Empirical evidence with Aggregate Time Series

Hall (1978, JPE) seminal paper launched the Euler equation methodology. He checked the restriction

$$E_t\left(c_{t+1} - c_t\right) = 0,$$

which implies that no variable can forecast $c_{t+1} - c_t$ or that in any regression of $c_{t+1} - c_t$ on X_t, a variable known at t, the coefficient should be zero:

$$c_{t+1} = \alpha + \beta c_t + \gamma X_t + \varepsilon_{t+1} :$$

Prediction: $\beta = 1$, $\gamma = 0$. (This model is run in logs...)

He found that aggregate consumption is surprisingly hard to forecast, so $\log c_t$ follows approximately a random walk. Empirically it is a good rough approximation that consumption growth $= \log\left(\frac{c_{t+1}}{c_t}\right)$ is *iid*. Hall found that only stock prices could forecast consumption. Further research has found that this result is not robust. In particular there are two important violations of PIH:

Excess sensitivity

Lagged income forecasts consumption, i.e.

$$C_t = C_{t-1} + \alpha + \beta Y_{t-1} + \varepsilon_t,$$

with $\beta \neq 0$. Excess sensitivity is a violation of the martingale property.

Excess smoothness

If income is an AR(1), then Friedman's intuition that permanent income is smoother than current income is true, justifying that consumption is smoother than income. But in the data income growth may well be positively autocorrelated:

$$\Delta y_t = \rho\Delta y_{t-1} + \mu + \varepsilon_t,$$

and in this case the PIH predicts that consumption should be more volatile than income. (See Pb Set 5.) That's the "excess smoothness" puzzle.

Individual data

Researchers (e.g. Attanasio, Blundell, Deaton), have also applied the PIH to micro-level data. Some additional issues that come up in this case include (a) different goods, (b) household composition and life-cycle issues, (c) borrowing constraints, (d) quality of the consumption data.

The PIH tends to be rejected, though it is not a bad description of some features of the data: for instance, the cross-sectional variance of consumption rises linearly with age (See Pb Set 6).

Windfalls

People have also used (unanticipated) windfalls or tax rebates to see if households behave according to the PIH. One criticism of this work is that the welfare gain to not smoothing consumption, for a small variation in consumption, is not very large, so inaction is nearly optimal.

Multiple shocks and information

An important possibility is that households face an income process with more than one shock, one of which is permanent and the other transitory. This would lead them to have smooth consumption despite having some persistence in income growth (see Quah). This assumes that the consumers are able to distinguish the two shocks.

More generally, you want to worry whether what we measure as a shock in the data, is really a surprise (i.e. unexpected) by the agent, who is presumably better informed. In principle, consumption reveals to you the expectation of the agent regarding his future income (see e.g. Richard Blundell's work for some nice applications).

Cross-Equations Restrictions and the Lucas critique

The PIH is a good example of the usefulness of rational expectations. You could formulate PIH for any set of expectations: consumption depends on expected future income. To bring this theory to the data, you need to specify a model for expected future income (one equation), and deduce its implications for consumption (a second equation). The power of rational expectations is that you obtain cross-equations restrictions, i.e. some parameters enter *both* equation. Concrete example: PIH with an AR(1) income process (say $\rho > 0$):

$$\text{1st equation: } y_{t+1} - \overline{y} = \rho\left(y_t - \overline{y}\right) + \varepsilon_{t+1}.$$

$$\text{2nd equation: } c_t = r\left(b_t + \frac{1}{1+r}E_t\left(\sum_{j\geq 0}\frac{y_{t+j}}{R^j}\right)\right).$$

Under rational expectations, the present value sum of the second equation *must* be consistent with the first equation:

$$\begin{aligned} E_t\left(\sum_{j\geq 0}\frac{y_{t+j}}{R^j}\right) &= E_t\left(\sum_{j\geq 0}\frac{y_{t+j}-\overline{y}}{R^j}\right) + \frac{\overline{y}}{1-\frac{1}{R}} \\ &= \left(\sum_{j\geq 0}\frac{\rho^j}{R^j}\right)\left(y_t - \overline{y}\right) + \frac{\overline{y}}{1-\frac{1}{R}} \\ &= \frac{y_t-\overline{y}}{1-\frac{\rho}{R}} + \frac{\overline{y}}{1-\frac{1}{R}}, \end{aligned}$$

and hence

$$c_t = rb_t + \overline{y} + \frac{y_t - \overline{y}}{R - \boldsymbol{\rho}}. \tag{5.11}$$

Hence, rather than having a free parameter for expected future income, the parameter is overdetermined, since ρ shows up both in the consumption equation and in the income equation. These additional restrictions are useful from an econometric point of view since they allow to test these models.

Lucas' critique is tightly related to this cross-equation restriction. Suppose that an econometrician runs a regression of c_t on y_t and b_t, and assume there is some classical measurement error in equation (5.11). The econometrician will recover coefficients γ_0, γ_1 and γ_2 such that $c_t = \gamma_0 + \gamma_1 b_t + \gamma_2 y_t + \varepsilon_t$. This equation will however not predict consumption's response to an additional increase in income (e.g., a government tax rebate). This is because this equation embodies the assumption that income is an AR(1), and the [high] coefficient on y_t in this regression is due to the empirical correlation between y_t and future $y's$ (because it is an AR(1)). Hence, a one-time tax rebate would not satisfy the process for forecasting y in the first equation, and would lead to much smaller consumption response.

Friedman's PIH

Friedman's book *A Theory of Consumption Function* introduced the permanent income theory, which is the basis for the PIH, but Friedman had in mind a more general model than the one I presented. (The model I presented, due probably to R. Hall (1978 JPE) is what is meant by PIH today.) Friedman's book was the first to use both cross-section and time-series data and develop a theory that could account for a variety of observations.

Here's a version of Friedman's theory:

(1) consumption is proportional to permanent income:

$$c_t = k \times y_t^P,$$

(2) current income is permanent income plus transitory income:

$$y_t = y_t^P + y_t^T,$$

with y_t^T independent of y_t^P, so that $Cov(y_t^P, y_t^T) = 0$.

The marginal propensity to consume, measured as the regression coefficient of current consumption on current income is, assuming no measurement error:

$$mpc = \frac{Cov(c_t, y_t)}{Var(y_t)} = \frac{kVar(y_t^P)}{Var(y_t^P) + Var(y_t^T)} < k,$$

so that the mpc is less than k (the average propensity to consume). The difference between the two depends on the importance of transitory income.

Friedman showed this model could explain a variety of puzzles, e.g.:[14]

(1) The cross-sectional Marginal Propensity to Consume (XMPC) is less than the Average Propensity

to Consume out of Permanent Income,

(2) Expected Consumption Growth is not Positively Correlated with Income in the Cross-Section,

(3) Farmers, and other Self-Employed, Have a Lower XMPC, and a smaller ratio XMPC/k,

(4) Time Aggregation Increases the XMPC,

(5) XMPC is Smaller When Estimated in more Homogeneous Cross-Sections, with application to cross-country differences,

(6) Same XMPC for Blacks and Whites but In a Group of People with the Same Income, Blacks Consume Less than do Whites,

(7) Time Series Aggregate Marginal Propensity to Consume (TMPC) is Less Than k.

[14] The list is taken from Casey Mulligan.

Handout 7: Precautionary Savings

In Handout 6, we studied the PIH, which relies on the following assumptions:
(1) there is only a risk-free asset to save or borrow (no state-contingent assets);
(2) there is no borrowing constraint;
(3) the utility function is quadratic;
(4) $\beta R = 1$.

The first assumption is important if we want to get away from complete markets: we need to assume that some assets are not available. The fourth one is not very important, it merely simplifies slightly the analysis by removing the trends. This leaves #2 and #3. In handout 8, we will relax #2. In this handout, we'll look at the consequence of changing #3.

Euler Equation Approximation

Start from the Euler equation, for a risk-free asset. Since there are no borrowing constraints, it holds with equality:

$$u'(c_t) = \beta R E_t \left(u'(c_{t+1})\right).$$

Do a 2nd order Taylor approximation:

$$u'(c_{t+1}) \simeq u'(c_t) + u''(c_t) \times (c_{t+1} - c_t) + \frac{1}{2} u'''(c_t) \times (c_{t+1} - c_t)^2.$$

Plug this in the Euler equation:

$$\frac{1}{\beta R} \simeq E_t \left(1 + \frac{u''(c_t)c_t}{u'(c_t)} \times \frac{c_{t+1} - c_t}{c_t} + \frac{1}{2} \frac{u'''(c_t)c_t}{u''(c_t)} \frac{u''(c_t)c_t}{u'(c_t)} \times \left(\frac{c_{t+1} - c_t}{c_t} \right)^2 \right)$$

Hence, if I define $\gamma = -\frac{u''(c_t)c_t}{u'(c_t)}$ and $\eta = -\frac{u'''(c_t)c_t}{u''(c_t)}$ (=risk aversion) and approximating $\frac{c_{t+1}-c_t}{c_t}$ by $\Delta \log c_{t+1}$:

$$\frac{1}{\beta R} - 1 \simeq -\gamma E_t \Delta \log c_{t+1} + \frac{1}{2} \eta \gamma \times E_t \left(\Delta \log c_{t+1}^2 \right)$$

$$E_t \Delta \log c_{t+1} \simeq \frac{1}{\gamma} \left(1 - \frac{1}{\beta R} \right) + \frac{1}{2} \eta \times Var_t \left(\Delta \log c_{t+1}^2 \right),$$

hence consumption growth is high (people are saving a lot) when either

- the interest rate is high, or
- the discount factor is high, or
- $\frac{1}{\gamma}$ (the intertemporal elasticity of substitution of consumption) is low (assuming $\beta R < 1$), or
- $\eta > 0$ and the variance of consumption growth is large.

Note that η involves the third derivative of the utility function, for which we have little intuition. While γ = risk aversion, η is called the *prudence coefficient*. If $u''' > 0$, a higher uncertainty in consumption induces higher savings.

Note that we always assume risk-aversion: higher uncertainty will thus reduce the utility of the agent. But how the uncertainty affects the agent's behavior is unclear unless we make the additional assumption of prudence.

Closed-form solution with CARA utility

There's a nice special case for which a closed-form solution exists. Assume that income y follows an AR(1) process:

$$y_{t+1} - \overline{y} = \rho\left(y_t - \overline{y}\right) + \sigma\varepsilon_{t+1},$$

with ε_{t+1} iid $N(0,1)$. Assume that the utility function has constant absolute risk aversion:

$$u(c) = -e^{-\delta c},$$

where $\delta =$ coefficient of absolute risk aversion.

The state variables of this problem are assets a and income y, because they are sufficient to summarize the agent's situation.

There are several ways to solve for the optimal consumption rule.

- Using the Bellman equation:

$$\begin{aligned} V(a,y) &= \max_{c,a'} \left\{u(c) + \beta E_{y'|y} V(a',y')\right\} \\ a' &= Ra + y - c. \end{aligned}$$

You can guess and verify that $V(a,y) = -e^{k_0 a + k_1 y + k_2}$, where k_0, k_1, k_2 are coefficients to be determined. Once you have the value function, you will also have the consumption function $c = c(a,y)$.

- Alternatively, you can use the Euler equation:

$$e^{-\delta c_t} = \beta R E_t \left\{e^{-\delta c_{t+1}}\right\},$$

and you can guess and verify that $c_t = c(a_t, y_t) = k_3 a_t + k_4 y_t + k_5$, where k_3, k_4, k_5 are coefficients to determined. Let's find these coefficients, i.e. find k_3, k_4, k_5 so that. the following equality holds for any (a,y) :

$$e^{-\delta(k_3 a + k_4 y + k_5)} = \beta R E_{y'|y} e^{-\delta(k_3 a' + k_4 y' + k_5)}$$

$$e^{-\delta(k_3 a + k_4 y + k_5)} = \beta R E_{y'|y} e^{-\delta(k_3 Ra + k_3 y - k_3(k_3 a + k_4 y + k_5) + k_4 y' + k_5)}$$

$$e^{-\delta(k_3 a + k_4 y + k_5)} = \beta R e^{-\delta(k_3 Ra + k_3 y - k_3(k_3 a + k_4 y + k_5) + k_4(\rho(y-\overline{y}) + \overline{y}) + k_5) + \frac{1}{2}\delta^2 k_4^2 \sigma^2}$$

$$-\delta\left(k_3 a + k_4 y + k_5\right) = \log\left(\beta R\right) - \delta \left(\begin{array}{c} k_3 Ra + k_3 y - k_3\left(k_3 a + k_4 y + k_5\right) \\ + k_4\left(\rho\left(y - \overline{y}\right) + \overline{y}\right) + k_5 \end{array}\right) + \frac{1}{2}\delta^2 k_4^2 \sigma^2$$

Since this equality must hold for any a, y, we must have that the coefficients in front of a (resp. y, resp. the constant) are equal on the LHS and on the RHS:

$$-\delta k_3 = -\delta k_3 R + \delta k_3^2$$

$$-\delta k_4 = -\delta k_3 + \delta k_3 k_4 - \delta k_4 \rho$$

$$-\delta k_5 = \log\left(\beta R\right) + \delta k_3 k_5 - \delta k_4 \overline{y}(1-\rho) - \frac{1}{2}\delta^2 k_4^2 \sigma^2 - \delta k_5$$

The first equation implies that

$$R - 1 = k_3$$

The second equation then implies that

$$k_4\left(1 + k_3 - \rho\right) = k_3$$

$$k_4 = \frac{R-1}{R-\rho},$$

and finally

$$k_5 = \frac{-\log(\beta R) - \delta k_4 \overline{y}(1-\rho) - \frac{1}{2}\delta k_4^2 \sigma^2}{\delta(R-1)}.$$

Putting all this together,

$$c(a,y) = (R-1)\,a + \frac{R-1}{R-\rho}\,(y-\overline{y}) + \overline{y} - \frac{\log(\beta R)}{\delta(R-1)} - \frac{1}{2}\frac{(R-1)}{(R-\rho)^2}\sigma^2.$$

Note how the first three terms are exactly the PIH as we know it from the quadratic model:

- MPC out of wealth is $r = R-1$;

- term in front of income is $\frac{R-1}{R-\rho}$ for an AR(1) income; if $\rho = 0$ this yields $r = R-1$ and if $\rho = 1$ this yields 1;

- $\overline{y}$ is the average income;

New terms:

- not-so-new: discounting/interest rate term $\frac{\log(\beta R)}{\delta(R-1)}$;

- really new term: precautionary savings: $\frac{1}{2}\frac{(R-1)}{\delta(R-\rho)^2}\sigma^2$: a higher uncertainty reduces consumption. The effect is stronger if the persistence of the income process ρ is greater since the consumer faces more risk then.

Interpretation of the 3 components: impatience; 'saving for a rainy day'; precautionary savings.

Gourio

Handout 8: A Savings Problem with Borrowing Constraints

In this handout we will study the effect of borrowing constraints on savings behavior. We still assume that an agent faces some uncertainty about his future income, and he uses a risk-free asset to smooth consumption. But now the agent may be constrained in how much he can borrow. First, from a theoretical point of view, we will see that you need to have some borrowing constraints to ensure that people will repay their debts. Second, from a practical point of view, borrowing constraints seem to allow to solve some of the puzzles created by the PIH model or the models with precautionary savings.

There's a wide debate in all of macroeconomics about the importance of borrowing constraints (or more generally market incompleteness), including how exactly they should be modelled, whether they matter at the micro level, and whether they matter at the macro level.

1- Introductory examples

Two-period model with certainty, with borrowing constraint and with a kink in the interest-rate. Graphical presentations.

Case A: no borrowing constraint:

$$\begin{aligned} & \max_s \{u(c_1) + \beta u(c_2)\} \\ s.t. & \quad : \\ c_1 + s & = y_1 \text{ and } c_2 = y_2 + Rs \end{aligned}$$

i.e. Present-Value Budget Constraint:

$$c_1 + \frac{c_2}{R} = y_1 + \frac{y_2}{R}.$$

Euler equation holds: $u'(y_1 - s) = R\beta u'(y_2 + Rs)$.

Implication: MPC out of y_1 =MPC out of $y_2/R \rightarrow$ only the total wealth matters.

Case B: Cannot save less than $\overline{s}$. Then write Lagrangean, get FOC: either $u'(y_1 - \overline{s}) > R\beta u'(y_2 + R\overline{s})$, then $s = \overline{s}$ and you are constrained, o/w interior solution as before. Draw the budget set. Bunching at the constraint. If you are constrained, $dC_1/dY_1 = 1$ and $dC_1/dY_2 = 0$.

Case C: kink with different interest rates for borrowers and for lenders: $R^L < R^B$. Draw the budget set. Bunching at 0. Three possible cases. (To show in class.) People at the kink have a high MPC, while the other people have a lower MPC.

Case D: add uncertainty about future income $y_{2,s}$. First no borrowing constraint. Then

$$\begin{aligned} & \max_b \left\{ u(c_1) + \beta \sum_s \pi_s u(c_{2s}) \right\} \\ s.t. & \quad : \\ c_1 + b & = y_1 \text{ and } c_{2s} = y_{2s} + Rb \end{aligned}$$

If the solution is interior, the Euler equation holds:

$$u'(y_1 - b) = R\beta \sum_s \pi_s u'(y_{2s} + Rb).$$

Note however that in order for consumption to be weakly positive tomorrow, we must have $y_{2s} + Rb \geq 0$ for every state of nature s. Hence, this means the agent cannot borrow more than $-b \leq \min\left\{\frac{y_{2s}}{R}\right\}$, where the min is taken over all the possible states. In particular, if there is one state where income tomorrow is zero, the agent cannot borrow today, since he might not be able to repay tomorrow.[15]

2- The natural borrowing constraint

Example D shows that there is always some bound on how much you can borrow: the maximum amount that you can borrow today is the *minimum* possible present-value of your income in the future. This is called the natural borrowing constraint:

$$a \geq -\min \sum_{t=1}^{\infty} \frac{y_t}{(1+r)^t} = a^{natural}.$$

Since the PIH does not take this into account, it is somewhat internally inconsistent (i.e., you need to assume that people sometimes have negative consumption to repay their debt - which is not a problem with quadratic utility, but would be with different preferences).

If there is a lowest possible income level $\underline{y}$, and this income level is reached with positive probability in each future period, then

$$a^{natural} = -\sum_{t=1}^{\infty} \frac{\underline{y}}{(1+r)^t} = -\frac{\underline{y}}{1+r} \frac{1}{1-\frac{1}{1+r}} = -\frac{\underline{y}}{r}.$$

In particular, if $\underline{y}= 0$, then the natural borrowing constraint is zero: the agent cannot borrow anything, because he might not be able to repay.

Moreover, if the utility function satisfies a Inada condition, i.e. $\lim_{c \to 0} u'(c) = +\infty$, then the agent will never choose to have zero assets, i.e. to have a constraint which exactly binds: this is because in this case,

Hence, in an environment with , we would never observe the borrowing constraint

3- The general savings problem

This work is associated (mostly) with the names of Bewley, Aiyagari, Caroll, and Deaton. I highly recommend S-L (beginning of chapter 17) or Huggett (1993) as alternative expositions. While some of the results here can be proved, I will not propose proofs.

The model takes as given a borrowing constraint, which could be the natural borrowing constraint, or could be a (tighter) constraint, e.g. $a \geq 0$.

(1) Utility: expected discounted separable utility: $E \sum_{t \geq 0} \beta^t u(c_t)$

(2) Income shocks: each agents has some income y each period. y follows a Markov chain with values $\{y_1, ..., y_N\}$ and transition matrix P :

$$\forall i, j = 1 \text{ to } N : \Pr(y_{t+1} = y_j \mid y_t = y_i) = P_{i,j}.$$

(3) Budget constraint: let a_t = assets at the end of $t-1$. Then

$$a_{t+1} = Ra_t + y_t - c_t.$$

$$R = \text{ gross risk-free rate } = 1 + r$$

[15] We use a risk-free asset, so it must be repaid in all states of nature. An interesting extension is to allow agents to default in some states, i.e. not to repay their debt.

[This budget constraint can be rewritten in terms of the assets at the *beginning* of the period Ra_t w/o the income y_t; or also including the income y_t, i.e. $Ra_t + y_t$.]

(4) Borrowing constraint: assets must be greater than $\underline{a}$ at any time.

Let $v(a, y)$ the max utility you can get if start from assets a and income y. The Bellman equation is:

$$\begin{aligned} v(a, y) &= \sup_{c \geq 0, a' \geq \underline{a}} \left\{ u(c) + \beta \sum_{y'=1}^{N} P_{y,y'} v(a', y') \right\} \\ s.t. &: a' = Ra + y - c. \end{aligned}$$

$$v(a, y) = \sup_{a' \geq \underline{a}} \left\{ u(Ra + y - a') + \beta \sum_{y'=1}^{N} P_{y,y'} v(a', y') \right\}.$$

First-order condition:

$$u'(c) \geq \beta \sum_{y'=1}^{N} P_{y,y'} v_a(a', y'),$$

with equality if $a' > \underline{a}$.

Envelope condition:

$$v_a(a, y) = Ru'(c).$$

Combining the two:

$$u'(c) \geq \beta E_{y'/y} v_a(a', y') = \beta R. E_{y'/y} u'(c'),$$

with equality if $a' > \underline{a}$. The Euler equation is a strict inequality if the borrowing constraint binds.

General properties

Obtained from the theorems of dynamic programing of Handout #2:

- v is increasing in a and concave in a.

- if $y_i \leq y_j$ implies that the distribution $P_{y_j,.}$ FOSD $P_{y_i,.}$, then v is increasing in y.

- the policy functions $c(a, y)$ and $a'(a, y)$ are also increasing in a and in y under the same condition.

["Technicality": first need to prove that assets remain bounded, i.e. $a'(a, y) \leq a$ for $a \geq a^*$ and all y.] Draw typical plots: c is concave not linear in a.

Certainty case

To discuss in class (or see SL chapter 16). After some time, consumption will be perfectly smooth.

The iid case

In the iid case $P_{y,y'}$ does not depend on y. Call it $\pi(y')$. The Bellman equation is:

$$v(a, y) = \sup_{a' \geq \underline{a}} \left\{ u(Ra + y - a') + \beta \sum_{y'=1}^{N} \pi(y') v(a', y') \right\}.$$

Define $z = Ra + y$. Let $w(z)$ the value if you start with z. The new Bellman equation is

$$w(z) = \max_{s \geq \underline{a}} \left\{ u(z - s) + \beta \sum_{y'=1}^{N} \pi(y') w(Rs + y') \right\}.$$

Note that by redefining the state variable, we have simplified the problem since our value function now depends only on one variable, z, instead of both a and y. The FOC and Envelope conditions are now

$$\begin{aligned} u'(z-s) &\geq \beta R \sum_{y'=1}^{N} \pi(y') w'(Rs+y'), \\ w'(z) &= u'(z-s). \end{aligned}$$

Result: the borrowing constraint binds iff $z \leq z^*$, for some z^*. Proof to sketch. Graph of LHS and RHS.

Empirical results

This model departs strongly from complete markets, and has generally less insurance than the PIH: people sometimes cannot borrow as much as they would like to. Borrowing constraints can thus explain some features of the data, e.g. why young people are not consuming much relative to their income. Depending on parameter values, the model behaves either approximatively like the complete markets model, or more like autarky. Clearly, if there is little risk, (and income has no upward trend), then the borrowing constraint will not matter too much in the long run. Hence, when the persistence of the income process is not too big, or the standard deviation of shocks is not too big, we get closer to the PIH. When agents are patient enough, they accumulate a large stock of assets and they can then use these assets to smooth consumption if they have a sequence of bad income shocks.

Variants of this model have been used and estimated by Gourinchas and Parker (2002) to explain life-cycle patterns in consumption. Krusell and Smith (1998) used it to study business cycles with heterogeneous agents, and Aiyagari (1995) studied capital taxation. There are now many more applications.

Handout 9: Stationary Distributions and Macroeconomic Equilibrium in models with idiosyncratic shocks

This handout explains how to obtain aggregate implications from the savings problem that we discussed in handout 8. Basically, we will compute the savings choice of each individual, as a function of his assets and income, using the last handout, then we will aggregate these choices by summing across agents with different levels of assets or income, and finally we will make sure that markets clear, i.e. supply equals demand, by adjusting the price - in our savings case, the interest rate.

The method is very general, hence this handout is also a guide to how one can build *equilibrium* models where heterogeneous individual agents make choices in a dynamic, uncertain environment. Indeed, the last section of this handout gives examples where I apply this framework to study labor demand, firm dynamics (i.e. entry and exit), and investment.

These models can allow for a large amount of heterogeneity across agents, either in ex-ante fixed characteristics (preferences, endowed skill, ...), in ex-post random events (health, income shocks, ...) and in the choices implied by these characteristics and events (e.g. savings, schooling choice, ...). Hence these models can be seen as 'macroeconomics without the representative agent' in that we allow for many differences between the agents.[16]

1 - Stationary distributions

In Handout 2, we saw that some Markov chains have a well-defined unique "stationary distribution" (aka invariant distribution, ergodic distribution). For instance, for the 2-state Markov chain with states $x_t = 1$ or 2 and with transition matrix between these two states:

$$P = \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix},$$

for $0 < p < 1$ and $0 < q < 1$, there is a unique invariant distribution

$$\pi = \begin{pmatrix} \frac{1-q}{2-p-q} \\ \frac{1-p}{2-p-q} \end{pmatrix}$$

since $\pi = P'\pi$ (this was computed in PS2). There are two interpretations of this computation:

- *Time-Series*: pick any initial value $x_0 \in \{1, 2\}$, then draw $\{x_t\}$ according to the Markov chain implied by the transition matrix P. On average over long periods of time, x_t will spend a fraction $\frac{1-q}{2-p-q}$ of the time in state 1 and the remainder (i.e. $\frac{1-p}{2-p-q}$) in state 2. Hence the stationary distribution π^* tells us how the values of x_t will be distributed on average in the long run.

- *Cross-Section:* consider following a large number N of processes x_{it}, each of which follows the Markov chain P; assume that these processes are statistically independent, so that the draw for x_{it+1} is independent of the draw for x_{jt+1} for any i and j. Assume furthermore that the initial

[16] Of course, when markets are complete it is easy to incorporate heterogeneity (in wealth, in preferences...) since we can aggregate and obtain a representative agent utility function. What we are doing here is substantially harder, since we relax the assumption of complete markets, so we do not have a representative agent, and we must compute the equilibrium that will arise given the market structure (in this case, a bond market).

condition x_{i0} is drawn according to the invariant distribution π^*, i.e. a fraction $\frac{1-q}{2-p-q}$ of these N processes start in state 1 and the remainder start in state 2.[17] Then, at any future date t, the cross-sectional distribution of x_{it} is π^*, i.e. there is a fraction $\frac{1-q}{2-p-q}$ of processes that are in state 1, and a fraction $\frac{1-p}{2-p-q}$ that are in state 2. Hence, the average across all the process $(\frac{1}{N}\sum_i x_{it})$ is constant, even though there are fluctuations at the individual level. In this case, π^* is the cross-sectional distribution.

2 - Markov chain for assets induced by the optimal individual behavior

Recall the savings problem we studied in Handout 8:

$$V(a,y) = \sup_{a' \geq \underline{a}} \left\{ u(Ra + y - a') + \beta \sum_{y'=1}^{N} P_{y,y'} V(a', y') \right\},$$

and for simplicity assume that assets a and a' must lie within a grid, i.e. a finite set of values A (which can be as large as you want):$a \in A$. For this problem the state vector is two dimensional, it is (a, y) : we need to know current assets and current income to know

The model delivers a policy function $a' = a'(a, y)$ which gives the optimal choice of assets tomorrow given assets today. The ongoing shocks to y make assets evolve randomly. However, this is not any kind of uncertainty: because assets tomorrow depend only on assets today and income today, the vector $X_t = \begin{bmatrix} a_t \\ y_t \end{bmatrix}$ follows a Markov process. Why? $a_{t+1} = a'(a_t, y_t)$ is a (deterministic) function of the vector X_t and y_{t+1} is drawn from the probability distribution $P(y_{t,}.)$. Hence both variables only depend stochastically (or determinically, which is a special case) only on the last period's two variables. Formally, the law of motion for the distribution over assets and income is:

$$\begin{aligned} \Pr\left(X_{t+1} = (\widetilde{a}, \widetilde{y})\right) &= \sum_{y=1}^{N} \sum_{a \in A} \Pr\left(y_{t+1} = \widetilde{y} \mid y\right) \times 1_{a'(a,y)=\widetilde{a}} \times \Pr(y_t = y, a_t = a) \\ &= \sum_{y=1}^{N} \sum_{a \in A} \Pr\left(y_{t+1} = \widetilde{y} \mid y\right) \times 1_{a'(a,y)=\widetilde{a}} \times \Pr(X_t = (a, y)), \end{aligned}$$

where $1_{a'(a,y)=\widetilde{a}}$ is a characteristic function, i.e. it is one if $a'(a, y) = \widetilde{a}$ and 0 otherwise.

What is this equation? We are simply counting the different ways of arriving to the state $(\widetilde{a}, \widetilde{y})$. To do this, we add up all the probabilities of going from any (a, y) to $(\widetilde{a}, \widetilde{y})$ in one period, and then multiply them by the probability of being in (a, y) at time t. This law of motion shows that we can compute the probability of being in state $(\widetilde{a}, \widetilde{y})$ tomorrow given the probability of being in state (a, y) today only, so this is a Markov process.

Just like in Section 1 of this handout we looked for a stationary distribution, we will look for a stationary distribution for the vector X. (For instance, think of stacking all the possible combinations: the enumeration of possible states, rather than being (low y, high y), is now have (low y, assets $= a_1$), (high y, assets $= a_1$), (low y, assets $= a_2$), (high y, assets $= a_2$), and so on up to a_N.)

[Write stacked system as in Sargent-Ljungqvist]

[17]Note: if you do not start with the distribution π^*, as we know from handout 2, you will under some conditions converge to it, so that our statement will still be true, in the long run.

Once we have found this stationary distribution for X, we have the *joint* distribution of assets and income in this economy, call it $\pi^*(a,y)$. Given this π^*, I can compute the mean assets $Ea = \sum_{y=1}^{N}\sum_{a\in A} a\pi^*(a,y)$, and for instance the covariance b/w assets and income $Cov\,(a,y) = \sum_{y=1}^{N}\sum_{a\in A}(a-Ea)(y-Ey)\,\pi^*(a,y)$, and any other statistic about the cross-sectional relations or the micro-level time series relations in this economies.

But given that this distribution is invariant, aggregates will be constant over time: Ea does not depend on calendar time t even though each agent's assets a_{it} will vary over time.

3 - Equilibrium with No Outside Supply of Bonds ("pure credit economy")

I now consider an equilibrium where agents trade with each other this risk-free bond. I assume that the only source of uncertainty in this economy is the idiosyncratic shocks that each agent faces. Faced with good income shocks, some agents will want to lend, while some other agents will want to borrow. I then search an equilibrium where the total amount of net credit in this economy is zero, i.e. the bond market clears. (This of course assumes that no one is supplying net credit: in some context you may think that the government or the rest of the world are actually supplying some net credit.) This assumption is called the *no outside bond supply,* or the *zero-net supply bond.*

The construction of the equilibrium proceeds as follows. For a given interest rate r, I can solve the Bellman equation (numerically) and find the policy functions $a'(a,y;r)$. This policy function depends on the interest rate r. This allows me to define the transition matrix for the vector X (i.e. assets, income) $Q^{(r)}$ This in turn allows to find the ergodic distribution of the process X, let's call it $\pi^{*(r)}$. As a result I can find if total assets

$$Ea(r) \overset{def}{=} g(r) = \sum_{a\in A}\sum_{y=1}^{N} a\pi^{*(r)}(a,y).$$

I have thus defined a mapping that takes a r and find the aggregate excess demand for assets $g(r)$. To find an equilibrium I only need to adjust r such that $g(r)=0$. (It is possible to prove under some conditions that this will work theoretically, i.e. the mapping g is monotonic and continuous; and this is not hard to do numerically.) This is what Huggett (1993 JEDC) does in his paper. One key result is that in equilibrium, $r<\rho=\frac{1}{\beta}-1$. Hence this equilibrium has a "low" risk-free rate, i.e. lower than the representative agent model with the same preferences.

4 - Equilibrium with Neoclassical Production Function

Another way to define an equilibrium is to embed this savings problem in the neoclassical growth model (This is Ayagari 1994 QJE). Agents are going to borrow or lend, but now there is some capital that constitutes an outside supply of assets. Reinterpret y as the number of efficiency units of labor that the agent has, i.e. he provides y units of labor to the market and makes yw out of it.

An equilibrium is then defined by:

(1) Given w and r, each agent solves:

$$\begin{aligned} v(a,y;w,r) &= \sup_{c\geq 0,a'\geq \underline{a}}\left\{u(c)+\beta\sum_{y'=1}^{N}P_{y,y'}v(a',y')\right\} \\ s.t. &: \quad a'=(1+r)a+wy-c. \end{aligned}$$

(2) The firm chooses capital and labor to maximize profits:

$$\max_{K,N}\left\{F(K,N)-wK-(r+\delta)\,K\right\}.$$

(3) Equilibrium:

$$\begin{aligned} K &= \sum_{a\in A}\sum_{y=1}^{N}\pi^*(a,y)a, \\ N &= \sum_{a\in A}\sum_{y=1}^{N}\pi^*(a,y)a \\ F(K,N)-I &= F(K,N)-\delta K = C = \sum_{a\in A}\sum_{y=1}^{N}\pi^*(a,y)c(a,y) \end{aligned}$$

To find an equilibrium proceed as in the section above. Start with a guess for r. Deduce from the Firm's FOC the implied w. Solve the Bellman equation for the consumer problem. Deduce the policy rule $a'(a,y)$. Find the transition probabilities for the vector $X=(a,y)$. Compute the ergodic distribution π^*. Check that the equilibrium conditions hold. If not, adjust your guess for r until they do. This tatonement process works reasonably well in practice.

5 - Additional Examples of Equillibria with idiosyncratic shocks

Example A: Dynamic Labor Demand

Consider a measure one of firms which are ex-ante identical but are hit ex-post by productivity shocks z. These shocks follow a Markov process with transition P and values in the set Z. The firm can hire labor at wage w, however it must incur an adjustment cost $c(N,N')$ when it changes its employment from N to N'. The profit function is $\pi(z,N)=zN^\alpha$ for some $0<\alpha<1$. The Bellman equation for a given wage w is:

$$V(N,z;w)=\max_{N'\geq 0}\left\{zN'^\alpha - wN' - c(N,N') + \beta E_{z'|z}V(N',z';w)\right\},$$

where $\beta E_{z'|z}V(N',z') = \beta\sum_{z'\in Z}P(z,z')V(N',z')$. Note that the timing is such that you can adjust your employment right after observing the shock. Note that if there are no adjustment costs, the employment decision becomes a static choice that you can solve by pencil and paper. Finally remark that I am not making any assumption on the nature of adjustment costs (quadratic vs. fixed vs. asymmetric, gross vs. net adjustment costs).

The solution to this Bellman equation is a value function $V(N,z;w)$ and the associated policy function $N'=N'(N,z;w)$. Since z is a Markov process, the vector $\begin{pmatrix} N \\ z \end{pmatrix}$ is itself a Markov process with a transition matrix induced by the transition matrix for z and the policy function. Assume that this Markov process has a unique invariant distribution μ. Then total labor demand by firms is $L^d(w)=\int\int N\mu(dN,dz)=\sum_{z\in Z}\sum_{n\in N}n\mu(n,z)$.

To make this model an equilibrium, there are several possibilities. The simplest is to assume that there is a labor supply function $L^s(w)$ which is increasing in the wage. Then we look for the wage w^* that clears the market: $L^d(w^*)=L^s(w^*)$. A general equilibrium modelling is to assume a representative consumer with utility $U(C,N)$; the consumer receives firms' profits and wages and spends it on consumption. Then we can again find the wage that will equate labor supply and labor demand. Hopenhayn and Rogerson (JPE 1993) apply this setup to evaluate the welfare effects of firing costs.

Example B: Entry and Exit

Consider a setup similar to Example A, except that there are no adjustment costs. There is a measure one of firms in an industry. The industry has a downward-sloping demand $Q=D(P)$.

Each firm has the production function $Y = zN^\alpha$, where z still follows a Markov chain with transition matrix P. Firms can hire labor at the fixed wage w. Firms also need to pay a fixed cost each period f. Firms can decide to exit the industry if their productivity (stochastically) becomes too low; in this case they receive $\overline{v}$ upon exit. New firms can enter the industry by paying a sunk cost c; once they have paid the sunk cost, they draw an initial productivity z from a distribution λ. Thereafter their productivity evolves according to P.

The individual Bellman equation for a continuing firm is now:

$$V(z;P) = \max\left\{\overline{v}, \max_N \left\{zPN^\alpha - wN - f + \beta E_{z'|z}\left(V(z';P)\right)\right\}\right\},$$

where the first max reflects the decision to stay in the industry or to exit, and the second max reflects the optimal labor choice. Clearly, owing to the fixed cost, firms will exist if their productivity is low enough (and they do not expect it to rebound too quickly[18]). That is, there is a threshold z^* such that if $z \leq z^*$, the firm decides to exit.

For new firms, the expected value is $\int V(z)\lambda(dz)$, or in the discrete case $\sum_{z\in Z}\lambda(z)V(z)$. Free entry implies that this expected value is equal to the cost of entering, as long as entry is positive: $\int V(z)\lambda(dz) \leq c$, with equality if there is entry.

The distribution of firms over z evolves exogenously according to P, except for the selection (low productivity firms exit) and new entry:

$$\mu_{t+1}(z') = \sum_{z\in Z}\mu_t(z)1_{z>z^*} + M_t \times \lambda(z'),$$

where $M_t = \#$ of new entrants at time t.

An equilibrium for this industry is then a price P, entrant mass M, threshold z^*, distribution $\mu(z)$, such that

(1) each firm solves its Bellman equation and this determines z^* and output supply $y(z;P)$;

(2) μ is invariant: $\forall z' \in Z : \mu(z') = \sum_{z\in Z}\mu(z)1_{z>z^*} + M \times \lambda(z')$;

(3) the free entry condition is satisfied: $\int V(z)\lambda(dz) = c$;

(4) and output supply equals output demand:

$$\int y(z;P)\mu(dz) = Q^d(P).$$

This example is based on Hopenhayn (1992 Econometrica).

[18] i.e. the transition matrix P is monotonic in the sense of FOSD.

Example C: Investment

Similarly, we can consider the choice of capital by firms. Let $\pi(K, z)$ be the variable profit function when capital is K and technology z. Then we can write the problem

$$\begin{aligned} V(K, z; r, w) &= \max_{I} \left\{ \pi(K, z; r, w) - I - C(I, K) + \frac{1}{1+r} E_{z'|z} V(K', z') \right\} \\ s.t. &: K' = (1-\delta)K + I, \end{aligned}$$

where C is an adjustment cost function. This yields a policy function $I(K, z; r, w)$ or equivalently $K'(K, z; r, w)$.

To make this an equilibrium, assume a representative agent with utility $U(C, N)$, who supplies labor to firms, receives wages and dividends. Then we can have the interest rate and the wage rate clearing the goods market and the labor market.

Brief Notes on Compensating Differentials, Sorting, and Hedonic Markets

Compensating Differential

Consider $N \geq 2$ identical workers who have to choose between living in two cities A and B. Workers have utility over how much income they receive and in which city they live: $u(x, y)$, where $x =$ income and $y = A$ or B. One city, say A, is 'more enjoyable' than the other one, i.e. $\forall x \geq 0, u(x, A) \geq u(x, B)$. Clearly, we would like everyone to live in A if possible. However it might be that the number of possible people living in A is limited (e.g. there is a fixed number of houses). In any case, if some workers choose to live in city A and others choose to live in city B, workers must be indifferent to moving, so that the income of someone living in A is

$$u(x_A, A) = u(x_B, B).$$

The income differential $x_B - x_A > 0$ exactly compensates people for their location choice.

For instance, suppose there are only M_A houses in A and $M_B > N - M_A$ houses in B, and suppose that everyone has income y. The price for a house in B is then 0 since at the margin, there is no demand for it (more houses than people). The price of a house in A will then be such that

$$u(y - p_A, A) = u(y, B).$$

Example 1: labor markets: think of jobs as having different attributes: wage and some 'benefit' e.g. riskiness, dirtiness, comfort...

Exercise 2: General Equilibrium. Assume that the production function in each city is $X = m^\alpha$ where $m = \#$ of workers in the city, with $0 < \alpha < 1$. Work out the equilibrium. (No need for houses here!)

Exercise 3: An important question in labor economics is the effect of the minimum wage on employment. In particular, a puzzle is that if people are paid a wage equal to their productivity, and if the distribution of productivity shows "no bunching" (i.e. is atomless, or continuous), there should be only very few people paid *exactly* the minimum wage. In the data, this is not the case - there's a big mass of people who are paid exactly the minimum wage (e.g. in France it is around 15%). How does the compensating differential view of the labor market affect this question?

Question for thought: Clearly firms can compete by offering different non-pecuniary benefits rather than wages. Why then, might it be better that firms compete in prices rather than in non-pecunary benefits?

Note: of course, the reason I mention the compensating differential is that, as mentioned in class, the CAPM or CCAPM can be viewed as 'compensating differentials' where the characteristics are (mean return, risk) instead of (city, income) or (wage, benefit).

Sorting

Now suppose that people are different. They might differ in their taste, or in their income. Here I consider the case of income; so let's say income y is distributed according to the distribution F.

In our housing example, if p_A is the price of a house in A and p_B is the price of a house in B, then $p_B = 0$ since there are more houses than people; and p_A will be determined so that the marginal agent is indifferent b/w living in A and in B :

$$M_A = N \times (1 - F(y^*)),$$

$$U(y^* - p_A, A) = U(y^*, B).$$

Hence, people sort themselves between the two cities depending on their income (and tastes). In our case, this will tend to exacerbate the observed compensating differential because people who choose to go to A are those who value A most.

(Heterogeneity is hence a rationale for downward-sloping demand curves: customers who buy the good first are those with the highest valuation, hence as we lower the price we reach consumers with lower valuations.)

Question 1: how does sorting influence our interpretation of the CAPM?

Question 2: Suppose that some people prefer A but some people prefer B. Plot the price of the house in A and B that clears the market, as a function of the number of houses in A.

For thought: consider the market for houses. Houses are different in many dimensions (e.g. basement vs. sunny, neigborhood, square footage, etc.). Which of these characteristics affect the price of the house?

Hedonic Markets

Let's consider in general the choice of quality. Let's make the assumption that the quality can be represented by a number q (i.e. it is one-dimensional). We abstract from quantity choice by assuming that people have to decide only about which quality to buy, and they will buy only one unit. Assume markets are competitive.

(A) Firms are identical and Consumers are identical

Assume that the utility function is quasilinear[19]: the utility of the consumer is $U(q) - p$ where p is how much he pays for his unit. Also assume that firms have constant return to scale, so that the unit cost (average cost = marginal cost) of producing one unit of good of quality q is $c(q)$.

Taking as given the prices $p(q)$ of one unit of each quality, we find the supply curve and demand curve from the optimization of firms and households. With constant return, firms will price at average cost = marginal cost, i.e. $p(q) = c(q)$.

Households pick quality q to maximize

$$U(q) - p = U(x(q)) - c(q).$$

In equilibrium, only one quality is produced, the one which the consumer prefers:

$$U'(q^*) = c'(q^*).$$

(With q on the x-axis and p on the y-axis, plot of tangency of the marginal cost of quality curve and the iso-u curve, which determines q^*.)

[19] A reasonable approximation for partial-equilibrium analysis...

(B) Consumer heterogeneity, Firms are identical

Now I assume that consumers differ in their utility of the good. Let $U(q,\theta)-p$ be the utility of a consumer of type θ. Let θ be distributed according to F in the population. Given $p(q)=c(q)$ still holds with constant returns, we can find the consumer θ demand: he will buy the good $q^*(\theta)$ where

$$U_q(q^*(\theta),\theta)=c'(q(\theta)).$$

The firm will offer many qualities, each of which appeals to a given consumer. (Plot of various iso-u curves tangent to the *same* marginal cost of quality.)

(C) Firm heterogeneity, Consumers are identical

Let's go back to the case where households are identical, but firms differ in their cost: the cost function is $c(q,x)$. Each firm produces the quality $q^*(x)$ such that $U'(q^*(x))=c_q(q^*(x),x)$. The consumer is indifferent among these qualities. (Plot of various marginal cost curves, tangent to the same iso-u curve.)

(D) Firm heterogeneity and Consumers heterogeneity

In this case, there is an equilibrium price function $p(q)$ such that (1) households pick the quality they like best, based on their $\theta: U_q(q^*(\theta),\theta)=p(q^*(\theta))$, and (2) firms produce the qualities that maximize their profits, i.e. $p(\widetilde{q}(x))=c(\widetilde{q}(x),x)$, and (3) markets clear for each quality. [Add plot of the indifference curves and cost curves in the price-quality space.]

See Sherwin Rosen, JPE 1974 or AER 2003 for more.