Macroeconomics Summary

Rebecca Sela

October 28, 2005

Most current macroeconomic research is based on dynamic, stochastic models which involve optimizing agents. We look for properties of the equilibrium outcomes, using dynamic programming.

1 Probability and Markov Chains

1.1 Markov Chains

Definition Let $\{x_t\}$ be a stochastic process in discrete time. $\{x_t\}$ satisfies the *Markov property* if $P(x_t \in A|x_{t-1}, ..., x_0) = P(x_t \in A|x_{t-1})$. That is, knowing x_{t-1} is sufficient for the entire history of the process.

Definition Let S be the space of column basis vectors, $e_i \in R^n$, where the i^{th} element of e_i is 1 and all the other elements are 0. We define e_i as being in *state* i. Let $\mathcal{S}$ be the set of all subsets of S. Then, $A \in \mathcal{S}$ is the event of being in a certain subset of states. We define a *transition matrix*, P, by $P_{ij} = P(x_t = e_j|x_{t-1} = e_i)$, which is the one-step probability of moving from state i to state j (this makes the process Markov). We assume that this matrix is time-invariant. Note that $\sum_{j=1}^{n} P_{ij} = 1$ and $P_{ij} \in [0, 1]$. We also define an initial density, π_0, which gives the probabilities that $x_0 = e_i$. The states, transition matrix, and initial condition define a *Markov chain*.

To find the probabilities of future states, we have $\pi_k' = \pi_0' P^k$.

Definition Let $\overline{y}$ be an $n \times 1$ vector. Define $y_t = \overline{y}' x_t$, where x_t is the current state. Then, y_t is a *Markov random variable*.

Then, we may take expectations of the random variable in each period:

$$\begin{aligned} E(y_0) &= \pi_0' \overline{y} \\ E(y_t|x_{t-1} = e_i) &= e_i' P \overline{y} \\ E(y_{t+k}|x_t) &= x_t' P^k \overline{y} \end{aligned}$$

Definition π is a *stationary distribution* of a Markov chain if $\pi' P = \pi'$. That is, the stationary distribution does not change from one period to the next. This is also called the ergodic distribution, the invariant distribution, the steady state distribution, and the unconditional distribution.

Definition A process is *asymptotically stationary* if, for all π_0, $\lim_{t\to\infty} \pi_0 P^t = \pi_\infty$, and π_∞ does not depend on π_0.

Note that π_∞ is a stationary distribution.

Definition A state is *transitory* if there is a positive probability of leaving the state and never returning.

Definition A set $E \subset S$ is *ergodic* if $P(x_t \in E | x_{t-1} \in E) = 1$ and no proper subset of E has this property. That is, once the Markov chain enters the set, it never leaves.

Theorem 1.1 *The state space of any Markov chain can be partitioned into a least one ergodic set and one (possibly empty) transitory set.*

Theorem 1.2 *Let $P^\infty = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} P^n$ (which always exists). Each row of P^∞ is an invariant distribution of the Markov chain, and any invariant distribution is a convex combination of the rows of P^∞.*

Theorem 1.3 *P has a unique ergodic set if there exists a state, j, such that for all i, $P^n_{ij} > 0$. That is, all states eventually lead to j. (Then, once you are at j, all future paths are identical; also, there are no absorbing sets that do not include j.)*

Theorem 1.4 *If all $P_{ij} > 0$ then there is a unique invariant distribution and one ergodic set. If $P^n_{ij} > 0$ for some n, then there is a unique steady state distribution.*

1.2 Basic Time Series

Definition An AR(1) process is defined as $Y_t = \mu + \rho Y_{t-1} + \epsilon_t$, where $\epsilon_t \sim Normal(0, \sigma^2)$ are independent and identically distributed, and $-1 < \rho < 1$. Note that Y_t is a continuous random variable with support on all of R.

The transition function is $Y_t | Y_{t-1} \sim Normal(\mu + \rho Y_{t-1}, \sigma^2)$. Note that knowing Y_{t-1} is enough to characterize the distribution conditional on all the information up to time $t-1$, and Y_t satisfies the Markov property.

Definition The *lag operator*, L, is defined by $LX_t = X_{t-1}$. The *forward operator*, F, is defined by $FX_t = X_{t+1}$. Note that $F = L^{-1}$.

The norm of a lag operator is 1, since it maps constants to themselves.
Some facts about AR(1) processes:

- $Y_t = \frac{\mu}{1-\rho L} + \frac{1}{1-\rho L}\epsilon_t = \frac{\mu}{1-\rho} + \sum_{j=0}^\infty \rho^j \epsilon_{t-j}$.
- $E(Y_t) = \frac{\mu}{1-\rho}$ and $Var(Y_t) = \frac{\sigma^2}{1-\rho^2}$.

- The process is mean-reverting. The importance of the initial value, Y_0, declines exponentially fast to 0 (this means that the process is *asymptotically stationary*).
- For forecasting: $E_t(Y_{t+n}) = \mu \sum_{j=0}^{n-1} \rho^j + \rho^n Y_t$.

Definition An AR(2) process is given by $Y_{t+1} = \mu + \rho_1 Y_t + \rho_2 Y_{t-1} + \epsilon_{t+1}$, with $\epsilon_{t+1} \sim Normal(0, \sigma^2)$.

This process is not Markov, because the distribution depends on two lags. However, we may rewrite it as a vector AR(1) process, using (assuming $\mu = 0$):

$$\begin{bmatrix} Y_{t+1} \\ Y_t \end{bmatrix} = \begin{bmatrix} \rho_1 & \rho_2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} Y_t \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \epsilon_{t+1}$$

Then, $Z_t = (Y_t, Y_{t-1})$ is Markov of the first order, with $Z_{t+1} = AZ_t + B\epsilon_{t+1}$.

For the process to be stationary, the eigenvalues of A must lie in the unit circle. $E(Z_t) = (I - A)^{-1} BE(\epsilon_t) = 0$. The variance can be computed recursively using the *Lyapunov equation*, $Var(Z_t) = AVar(Z_t)A' + B\sigma^2 B'$. For forecasting, $E_t(Z_{t+j}) = A^j Z_t$.

Definition An MA(1) process satisfies $Y_t = \epsilon_t - b\epsilon_{t-1}$ with $\epsilon_t \sim Normal(0, \sigma^2)$ iid, $|b| < 1$.

MA processes do not satisfy the Markov property. However, if we include ϵ_t as part of the information set, then we have the *companion form*:

$$\begin{bmatrix} Y_t \\ \epsilon_t \end{bmatrix} = \begin{bmatrix} 0 & -b \\ 0 & 0 \end{bmatrix} \begin{bmatrix} Y_{t-1} \\ \epsilon_{t-1} \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \epsilon_t$$

If we write this as $Z_t = AZ_{t-1} + C\epsilon_t$, we may compute:

$$\begin{aligned} E(Z_t) &= AE(Z_t) = 0 \\ Var(Z_t) &= AVar(Z_t)A' + \sigma^2 CC' \\ &= \begin{bmatrix} 1 + b^2 & 1 \\ 1 & 1 \end{bmatrix} \sigma^2 \\ E_t(Z_{t+j}) &= A^j Z_t \end{aligned}$$

Definition An ARMA(1,1) process is given by $Y_t = \mu + \rho Y_{t-1} + \epsilon_t - b\epsilon_{t-1}$ with $\epsilon_t \sim Normal(0, \sigma^2)$ independent and identically distributed.

The companion form for the Markov property to hold is:

$$\begin{bmatrix} Y_t \\ \epsilon_t \end{bmatrix} = \begin{bmatrix} \rho & -b \\ 0 & 0 \end{bmatrix} \begin{bmatrix} Y_{t-1} \\ \epsilon_{t-1} \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \epsilon_t + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \mu$$

The one-step-ahead forecast is $E_t(Y_{t+1}) = \mu + \rho Y_t - b\epsilon_t$.

2 Dynamic Programming

Definition A *correspondence* is a mapping $\Gamma : X \to S(Y)$, where $S(Y)$ is the set of subsets of Y.

Dynamic programming depends on the following elements:

- A space of *exogenous variables*, Z, which are determined outside the system, and evolve according to a Markov process with probabilities $q(z, z')$. Note that Z may include lags of other variables, as long as everything evolves according to a Markov process.
- A space of *endogenous (control) variables*, X, which are determined in the system.
- A *payoff function* at date t, given by $F(x_t, x_{t+1}, z_t)$.
- A *constraint correspondence*, $x_{t+1} \in \Gamma(x_t, z_t)$.
- A *discount factor*, β, for future payoffs.

We wish to choose a *policy function*, $\pi_t : Z^{t-1} \to X$, in which we choose the next value of the control variable given all of the previous values of the exogenous variable (and, therefore, implicitly, all of the previous choices for the control variable).

Definition A policy, π, is *feasible* if $\pi_{t+1} \in \Gamma(\pi_t, z_t)$ for all t.

The *sequence problem* is to maximize:

$$V(x_0, z_0) = \sup_{\pi_{t+1} \in \Gamma(\pi_t, z_t)} \left(\sum_{t=0}^{\infty} \beta^t E(F(\pi_t, \pi_{t+1}, z_t) | x_0, z_0) \right)$$

That is, we wish to maximize the present value of the payoffs.

According to *Bellman's Principle*, the optimal policy has the property that whatever the initial state and actions, the remaining decisions constitute an optimal policy with regard to the state resulting from the initial decision. That is, one should make the best decision today given that one will make the best decision tomorrow. Using this principle (and that fact that the payoff function and constraint correspondence depend only on the current state), we have the *Bellman equation* for the sequence problem:

$$V(x, z) = \max_{x' \in \Gamma(x,z)} (F(x, x', z) + \beta E(V(x', z') | z))$$

We search for a value function, V, that will satisfy this relationship. We then find a policy function, $x' = \pi(x, z)$, to maximize the resulting expression.

Theorem 2.1 *Bellman's recursive formulation is equivalent to the original sequence problem if V does not change over time.*

Proof (For the case with certainty.) In this case:

$$\begin{aligned} V(x_0) &= \sup_{x_1 \in \Gamma(x_0)} F(x_0, x_1) + \beta V(x_1) \\ &= ... \\ &= \sup_{\{x_t\} \text{ feasible}} F(x_0, x_1) + \beta F(x_1, x_2) + \beta^2 F(x_2, x_3) + ... + \beta^k V(x_k) \end{aligned}$$

If $\beta^k V(x_k) \to 0$, as would happen if $V(x_k)$ does not grow over time, then the two problems are identical. ∎

Definition Let (S, ρ) be a metric space. Let $T : S \to S$ be a function. T is a *contraction mapping* with modulus β if, for some $\beta \in (0, 1)$,

$$\rho(Tx, Ty) \leq \beta \rho(x, y)$$

for all $x, y \in S$.

Theorem 2.2 Contraction Mapping Theorem. *If (S, ρ) is a complete metric space and $T : S \to S$ is a contraction mapping of modulus β, then T has exactly one fixed point, $v \in S$, and, for all $v_0 \in S$, $\rho(T^n v_0, v) \leq \beta^n \rho(v_0, v)$.*

Proof Let v_0 be any point in the space. Then, $\{T^n v_0\}$ defines a Cauchy sequence, since $\rho(T^n v_0, T^{n-1} v_0) \leq \beta^n \rho(v, T v_0)$ (and we may then apply the triangle inequality for any m, n). Because the metric space is complete, this sequence converges to some v which must be a fixed point.

The fixed point is unique, since if both v_0 and v_1 are fixed, then $\rho(v_0, v_1) = \rho(T v_0, T v_1) \leq \beta \rho(v_0, v_1)$, and the distance between them must be 0. ∎

We may use such iterations to compute the fixed point. For any v_0, we bound the distance to the fixed point (repeatedly using the triangle inequality):

$$\begin{aligned} \rho(T^n v_0, v) &\leq \sum_{i=0}^{\infty} \rho(T^{n+i} v_0, T^{n+i+1} v_0) \\ &\leq \sum_{i=0}^{\infty} \beta^i \rho(T^n v_0, T^{n+1} v_0) \\ &= \frac{1}{1-\beta} \rho(T^n v_0, T^{n+1} v_0) \end{aligned}$$

Thus, we can look at the stepsize between iterations to decide when the iterates are close enough to the fixed point.

Corollary 2.3 *If $S' \subset S$ is closed and $T(S') \subset S'$, then $v \in S'$. If $T(S') \subset S'' \subset S$, then $v \in S''$.*

Theorem 2.4 Blackwell's Conditions for a Contraction Mapping. *Let $X \subset R^k$. Let $B(X)$ be the space of bounded functions on X with the* sup *norm (that is, $\rho(f, g) = \sup_{x \in X} |f(x) - g(x)|$). Suppose T is a mapping on $B(X)$ that satisfies:*

- Boundedness*: If $f \in B(X)$ than $Tf \in B(X)$.*
- Monotonicity*: If $f, g \in B(X)$ and $f(x) \leq g(x)$ for all $x \in X$, then $Tf(x) \leq Tg(x)$ for all $x \in X$.*
- Discounting*: There exists $\beta \in (0,1)$ such that $(T(f+a))(x) \leq (Tf)(x) + \beta a$, for all $f \in B(X)$, constants $a \geq 0$, and $x \in X$.*

Then T is a contraction mapping on $B(X)$.

Proof Given $f, g \in B(X)$, set $a = \sup_{x \in X} |f(x) - g(x)|$. Then, $(f-a)(x) \leq g(x) \leq (f+a)(x)$ for all $x \in X$. By monotonicity, $T(f-a) \leq Tg \leq T(f+a)$. By discounting,

$$\begin{aligned} T(f+a)(x) &\leq Tf(x) + \beta a \\ Tf(x) &\leq T(f-a)(x) + \beta a \\ T(f-a)(x) &\geq Tf(x) - \beta a \end{aligned}$$

(Two steps are needed in the second case because discounting is only defined for $a \geq 0$.) Then, we have bounds on Tg:

$$Tf(x) - \beta a \leq T(f-a)(x) \leq Tg(x) \leq T(f+a)(x) \leq Tf(x) + \beta a$$

This shows that $\rho(Tf, Tg) \leq \beta a = \beta\rho(Tf, Tg)$. Thus, these conditions ensure that T is a contraction. ∎

For Bellman's equation, we consider the mapping:

$$Tf(x,z) = \sup_{y \in \Gamma(x,z)} F(x,y,z) + \beta E(f(y,z')|z)$$

If F is bounded, this mapping satisfies boundedness, since the maximum of two bounded functions is bounded as well. This mapping satisfies discounting, because $E(f(y,z') + a|z) = E(f(y,z')|z) + a$, so that:

$$\begin{aligned} T(f+a)(x,z) &= \sup_{y \in \Gamma(x,z)} F(x,y,z) + \beta E((f+a)(y,z')|z) \\ &= \sup_{y \in \Gamma(x,z)} F(x,y,z) + \beta E(f(y,z')|z) + \beta a \\ &= Tf(x,z) + \beta a \end{aligned}$$

This mapping satisfies monotonicity, because when $f \leq g$,

$$\begin{aligned} Tf(x,z) &= \sup_{y \in \Gamma(x,z)} F(x,y,z) + \beta E(f(y,z')|z) \\ &\leq \sup_{y \in \Gamma(x,z)} F(x,y,z) + \beta E(g(y,z')|z) \\ &= Tg(x,z) \end{aligned}$$

since $f \leq g$ implies that $E(f(y',z)|z) \leq E(g(y',z)|z)$ for any y, z. Thus, this mapping is a contraction, with the value function as a fixed point.

Definition A correspondence, $\Gamma : X \rightarrow Y$, is *lower hemi-continuous* at x if $\Gamma(x) \neq \emptyset$ and if for all $y \in \Gamma(x)$ and all $x_n \rightarrow x$, there exists $N \geq 1$ and a sequence $\{y_n\}$ such that $y_n \rightarrow y$ with $y_n \in \Gamma(x_n)$ for all $n \geq N$. (This means that new choices cannot appear discontinuously.)

Definition A compact-valued correspondence, $\Gamma : X \rightarrow Y$, is *upper hemi-continuous* at x if $\Gamma(x) \neq \emptyset$ if, for all sequences $x_n \rightarrow x$ and every sequence $\{y_n\}$ such that $y_n \in \Gamma(x_n)$ for all n, there exists a convergent subsequence of $\{y_n\}$ whose limit point is in $\Gamma(x)$. (This means that choices do not disappear discontinuously.)

Roughly, an upper hemi-continuous correspondence has a closed graph.

Definition If a correspondence is both upper and lower hemi-continuous, then we say that it is *continuous.*

Theorem 2.5 Theorem of the Maximum. *Let $X \subset R^l$, $Y \subset R^m$. Let $f : X \times Y \rightarrow R$ be a continuous function. Let $\Gamma : X \rightarrow Y$ be a compact-valued, continuous correspondence. Define $h : X \rightarrow R$ by $h(x) = \max_{y \in \Gamma(x)} f(x, y)$. Define the correspondence $G : X \rightarrow Y$ by $G(x) = \{y \in \Gamma(x) | f(x, y) = h(x)\}$. Then, h is continuous, and G is nonempty, compact-valued, and upper hemi-continuous.*

Note that there may be more than one optimal choice for any given x (this is why G is a correspondence), but a new one will always appear before the old one disappears, by upper hemi-continuity. Since h is constant across the two choices, it is still continuous in x.

Corollary 2.6 *If the constraint set, Γ, is convex and the function f above is strictly concave, then the optimal correspondence is single-valued and therefore is a continuous function.*

Proposition 2.7 *If F and Γ are continuous in x, then the value function is continuous in x.*

Proof Let g be continuous. Then, $F(x, y, z) + \beta E(g(x', z')|x, z)$ is the sum of continuous functions. Applying the Theorem of the Maximum shows that $Tg(x, z)$, which is the maximum, must be continuous (in x) as well. Note that the set of bounded, continuous functions on a compact set is closed under the sup norm and is mapped to itself under this contraction. Thus, the fixed point must be bounded and continuous as well. ∎

Proposition 2.8 *Suppose $F(x, y, z)$ is increasing in x and that Γ is monotonic in x (that is, whenever $x_1 \leq x_2$, $\Gamma(x_1, z) \subseteq \Gamma(x_2, z)$). Then, the value function is increasing in x.*

Proof Suppose f is increasing and $x_2 > x_1$. Let $y(x_1)$ and $y(x_2)$ be the optimal choices for x_1, x_2. Since $\Gamma(x, z)$ is monotonic, $y(x_1) \in \Gamma(x_2, z)$. Because F is monotonic in x and $y(x_2)$ is the optimal choice, we must have:

$$\begin{aligned} Tf(x_1, z) &= F(x_1, y(x_1), z) + \beta E(f(y(x_1), z')|z) \\ &\leq F(x_2, y(x_1), z) + \beta E(f(y(x_1), z')|z) \\ &\leq F(x_2, y(x_2), z) + \beta E(f(y(x_2), z')|z) \\ &= Tf(x_2, z) \end{aligned}$$

and Tf is increasing. Thus, T maps increasing functions to increasing functions, so the fixed point is increasing.

If F is strictly increasing in x, then any non-decreasing function, f, will be mapped to a strictly increasing function. Thus, the fixed point must be strictly increasing, since it cannot be non-decreasing. ∎

Proposition 2.9 *Suppose $F(x, y, z)$ is concave in x and y (that is, for any $\theta \in (0, 1)$, $F(\theta x_1 + (1-\theta)x_2, \theta y_1 + (1-\theta)y_2, z) \geq \theta F(x_1, y_1, z) + (1-\theta)F(x_2, y_2, z)$) and that $\Gamma(x, z)$ is convex. Then, the value function is concave.*

Proof Let g be any concave function. Consider $Tg(x_1, z)$, $Tg(x_2, z)$ and the associated maximizers, y_1, y_2. Define

$$\hat{T}g(x, z) = F(x, \alpha y_1 + (1-\alpha)y_2, z) + \beta E(g(\alpha y_1 + (1-\alpha)y_2, z')|z)$$

for $x = \alpha x_1 + (1-\alpha)x_2$. This assigns the choice which is the average of the choices for x_1 and x_2; this is feasible because Γ is convex. Since the average is a linear function, $\hat{T}g$ is concave. Note that $\hat{T}g(x_1, z) = Tg(x_1, z)$, $\hat{T}g(x_2, z) = Tg(x_2, z)$, and $\hat{T}g(x, z) \leq Tg(x, z)$ in the interior (because T chooses the optimal value and therefore cannot do worse than $\hat{T}$). Thus, T is greater than a concave function on the interior of the interval, and is therefore concave. In fact, if F and therefore $\hat{T}$ are strictly concave, then T is strictly concave as well.

Since (strictly) concave functions are mapped to themselves, the fixed point must be (strictly) concave. The set of concave functions is closed under the sup norm. Thus, the value function is concave. ∎

Theorem 2.10 Benvenesti-Schenkman. *Consider $V(x, z) = \max_{y \in \Gamma(x)} F(x, y, z) + \beta E(V(y, z')|z)$. If V is continuous and concave, Γ is continuous, and F is continuous, concave, and differentiable, then V is differentiable in x, if the optimal policy lies in the interior of $\Gamma(x, z)$.*

Proof There must be a hyperplane that lies above V and touches it at only one point (because V is concave). Let y^* be the optimal choice for x^*, z^*. Suppose we have some nearby x, z. By continuity and the fact that $y^* \in Interior(\Gamma(x^*, z^*))$, y^* is feasible. However, y^* need not be optimal. Thus, $F(x, y^*, z) + \beta E(V(y^*, z')|Z) \leq V(x, z)$, with equality only if y^* is still optimal. Thus, $V(x, z)$ is sandwiched between two functions that are differentiable in x, so it must be differentiable in x as well. ∎

This yields the following sequence of steps for dealing with value functions:

- Suppose the payoff function, $F(x, y, z)$, is bounded and continuous. Then, the mapping based on the value function is a contraction by *Blackwell's Theorem*, and the value function exists and is bounded.
- If $\Gamma(x, z)$ is continuous as well, then we apply the *Theorem of the Maximum* to show that the value function is continuous.
- Apply *Benevesti-Schenkman*'s theorem to show that the value function is differentiable and the other theorems to show that the value function is concave and increasing.

2.1 The Euler Equation

Because we are maximizing the control variable as part of evaluating the value function, we have a first order condition that the control variable next period must satisfy if it is in the interior:

$$F_2(x, y, z) + \beta E(V_1(y, z')|z) = 0$$

where F_2 is the partial derivative of F with respect to the second argument and V_1 is the partial derivative of V with respect to the first argument. (This means that we wish to balance the marginal payoff today with the discounted marginal value tomorrow.) If there is more than one control variable (in x or just in the one-period maximization), then there will be more first order conditions of the same form.

Theorem 2.11 Envelope Theorem. *If* $V(x, z) = \max_{y \in \Gamma(x,z)} F(x, y, z) + \beta E(V(y, z')|y)$ *and* y *is an interior point, then*

$$V_1(x, z) = F_1(x, y, z)$$

(where the V_1 *and* F_1 *are partial derivatives with respect to the first arguments).*

Proof *(Sketch.)* We may write y as a function of x, and we then use the chain rule to compute:

$$V_1(x, z) = F_1(x, y, z) + F_2(x, y, z)\frac{\partial y}{\partial x} + \beta E(V_2(y, z')|z)\frac{\partial y}{\partial x}$$

The last two terms sum to 0 if y is an interior point by the first order conditions. ∎

To use the envelope theorem, we plug in all the equality constraints, so that the value function is written in terms of only the current states and the maximizer. We then take the derivative(s) of the value function with respect to the control variables by: (1) dropping the fact that we are maximizing, (2) treating all the other states as constants, (3) treating the variable that we have maximized as constant, and (4) taking the derivative with respect to the variable of interest (probably applying the chain rule).

We may use the envelope theorem to calculate $V_1(y, z')$ to substitute into the first order conditions (moving all variables one period forward). This yields the *Euler equation*:

$$F_2(x, y, z) + \beta E(F_1(y, y', z')|z) = 0$$

where z' is the value of the exogenous state variable next period and y' is the corresponding choice of the control variable. The same idea works out if there are multiple endogenous variables.

Much analysis happens using only the Euler equation. This will not give unique solutions because it is only a first order condition, describing relative changes in the variables. To completely understand the solution, we also care about the constraints and the levels of the variables; these are the boundary conditions for the first order condition and are in the value function.

2.2 Steady States

In a non-stochastic problem, one may want to show the existence of a steady state, where $x_t = x_{t-1}$. (One can solve for this by just plugging $x_{SS} = y_{SS}$ into all the equations.) We then want to show if the steady state is *stable* by examining the derivatives to see if small perturbations will be pushed back toward the steady state.

In a stochastic problem, one could find an analog of the steady state by setting all the exogenous variables equal to their means and solving for the control variables as before.

2.3 Calculus of Variations

We use the Euler equation to study the effect of small perturbations on the system. At the margin, the loss of a $F_2(x, y, z)\epsilon$ today should be balanced by a gain of $\beta E(V_1(y, z')|z)\epsilon$ tomorrow. This shows the effect of small perturbations on the optimal path.

One may string together a sequence of Euler equations to find the effect of a perturbation over multiple periods. (For example, taking less payoff now and then having more value in two periods.)

2.4 Solution Methods

2.4.1 Numerical Dynamic Programming

To compute the value function numerically, we may start with any bounded function, I_0, and then use the mapping, T, to iterate to the value function:

$$T(I_k) = \sup_{y \in \Gamma(x,z)} F(x, y, z) + \beta E(I_{k-1}(y', z')|z)$$

where the values of I_k are computed over a grid (which is chosen to include all possible values that the control variable might take in the problem; this suggests that we should find bounds for the problem overall), and the maximum is found by searching over the grid of $F(x,y,z)+\beta E(I_{k-1}(y',z')|z)$, since I_{k-1} is known from the previous iteration. Then,

$$V(x,z)=\lim_{n\to\infty}T^n(I_0)$$

(one stops iterating once the functions are "close enough," based on the formula above for the distance to the fixed point).

2.4.2 Finding a closed form

Given a functional form for the payoff function, one may be able to guess a relationship between x and $y(x)$ that will satisfy the Euler equation (that is, if one replaces y by $y(x)$ in the Euler equation, everything will cancel, leaving some constants that are defined by the resulting equation).

2.4.3 Log-Linearization

For any value of x_t, we may write $x_t = x_{SS}(1+\hat{x}_t)$, where x_{SS} is the steady-state value and $\hat{x}_t$ is the percentage change from the steady state. We may wish to study how (percentage) deviations from the steady state are propagated, perhaps using stochastic difference equations.

By Taylor's Theorem, we may approximate any function by a linear function. We use *log-linearization*, using the fact that the percentage change in x is approximately $\ln(x)-\ln(x_0)$ and that:

$$\begin{aligned} f(x) &= f(e^{\ln x}) \approx f(e^{\ln(x_0)}) + f'(e^{\ln x_0})e^{\ln x_0}(\ln x - \ln x_0) \\ &= f(x_0) + f'(x_0)x_0(\ln x - \ln x_0) \end{aligned}$$

In two variables, this works out to:

$$f(x,y) = f(x_0,y_0) + f_x(x_0,y_0)x_0(\ln x - \ln x_0) + f_y(x_0,y_0)y_0(\ln y - \ln y_0)$$

Given an equation (such as an Euler equation) or a constraint, we may then apply this function term-by-term (since Taylor expansions add), expanding about the steady state values. Because the constraints and Euler equation must hold at the steady states, we then cancel the level terms, leaving linear equations that approximate the dynamics about the steady state. For example, if $f(x,y)=g(z)$, then we compute:

$$\begin{aligned} f(x_0,y_0) + f_x(x_0,y_0)x_0(\ln x - \ln x_0) + f_y(x_0,y_0)y_0(\ln y - \ln y_0) &= g(z_0) + g_z(z_0)z_0(\ln z - \ln z_0) \\ f_x(x_0,y_0)x_0(\ln x - \ln x_0) + f_y(x_0,y_0)y_0(\ln y - \ln y_0) &= g_z(z_0)z_0(\ln z - \ln z_0) \\ f_x(x_0,y_0)x_0\hat{x} + f_y(x_0,y_0)y_0\hat{y} &= g_z(z_0)z_0\hat{z} \end{aligned}$$

Alternatively, we may take logs of both sides of the equation first, and subtract the log of the equation at the steady state. For example, if $Z = XY$, then

$$\begin{aligned}
\log(Z) &= \log(X) + \log(Y) \\
\log(Z_{SS}) &= \log(X_{SS}) + \log(Y_{SS}) \\
z_t = \log(Z) - \log(Z_{SS}) &= x_t + y_t
\end{aligned}$$

We can also divide two equations first, and then take logs.

2.4.4 Linear-Quadratic Models and Stochastic Difference Equations

Another way to make everything linear is for the payoff to be quadratic (we need the resulting derivative to be linear in both x and y) and for all constraints and evolutions of variables to be linear. (Note that any increasing and concave quadratic payoffs will have a maximum, called the *bliss point* or the *satiation point*; we generally assume that point is outside the range of possible values.)

Then, the Euler equation can be written as a stochastic difference equation in terms of the expectations of the endogenous and exogenous variables. This can be re-written in a form like:

$$E_t(x_t + \alpha_1 x_{t+1} + \alpha_2 x_{t+2}) = E_t(z_t + \gamma_1 z_{t+1})$$

We may write the left-hand side with lag polynomials:

$$E_t((B^2 + \alpha_1 B + \alpha_2) x_{t+2}) = E_t(z_t + \gamma_1 z_{t+1})$$

We may factor the lag polynomial into $(B - \delta_1)(B - \delta_2)$. If the problem is reasonable, then $|\delta_1| < 1$ and $|\delta_2| \geq 1$. We take the term with δ_1 "into the future"(since zeroes less than one correspond to time series that are stationary when written in terms of future shocks) and the term with δ_2 "into the past". To do this, we rewrite $B - \delta_1 = (1 - \frac{\delta_1}{B})B$, which yields the difference equation:

$$E_t((B - \delta_2)(1 - \frac{\delta_1}{B}) x_{t+1}) = E_t(z_t + \gamma_1 z_{t+1})$$

(since $B x_{t+2} = x_{t+1}$).

Divide both sides by the term to be taken into the future:

$$E_t((B - \delta_2) x_{t+1}) = E_t\left(\frac{z_t + \gamma_1 z_{t+1}}{1 - \frac{\delta_1}{B}}\right)$$

Expand the fraction to write the right-hand side in terms of future values of the exogenous variable and solve for x_{t+1}:

$$x_{t+1} = x_t + E_t\left(\sum_{j=0}^{\infty} \frac{\delta_1^j}{B^j}(z_t + \gamma_1 z_{t+1})\right) = x_t + E_t\left(\sum_{j=0}^{\infty} \delta_1^j (z_{t+j} + \gamma_1 z_{t+j+1})\right)$$

This writes the control variable for the next period in terms of the previous value and the expectations of future shocks.

2.5 Testing the Model

One way to test a model is to estimate the model parameters directly from data (using GMM; any impulse response function is potentially a moment condition). This is then used to identify the shocks and see how well the model fits the data. (This is the Chicago method.)

Alternatively, one can calibrate the model, using parameter values determined in other studies (for example, a risk aversion parameter from experiments), instead of directly from the data. Then, one finds predictions from the model (such as the standard deviation of GDP growth) and compares them to what is actually seen in macro data.

3 Growth Models

3.1 Non-stochastic Growth

Definition The *Inada conditions* for a production function are:

1. $\lim_{K\to 0} f'(K) = \infty$ (the first unit of capital is infinitely productive)
2. $\lim_{K\to\infty} f'(K) = 0$ (eventually, additional capital is useless)

Definition In a *putty-putty model*, capital from a previous period may be consumed in the current period. (This is called "eating the capital".) In a *putty-clay model*, capital from a previous period cannot be consumed in the next period.

One way to bound everything when there is growth in the economy (such as population and technology growth) is to divide everything by the growing variable. For example, we may divide by effective labor, AL:

$$\frac{Y}{AL} = F(\frac{K}{AL}, 1)$$

(We assume constant returns to technology in this equation.)

Assumptions for this model:

- A consumer maximizes $\sum_{i=0}^{\infty} \beta^i u(C_{t+i})$, where u is strictly increasing, strictly concave, and continuous.
- The output function, $Y_t = f(K_t)$, is strictly increasing, strictly concave, has $f(0) = 0$, and satisfies the Inada conditions.
- Capital Accumulation: $K_{t+1} = (1-\delta)K_t + I_t$, where δ is the depreciation rate.
- The resource constraints: $Y_t = C_t + I_t$, $0 \le C_t \le (1-\delta)K_t + Y_t$. (This is a putty-putty model.)

The *Social Planner's Problem* is to maximize the utility of a representative agent.

The Bellman equation (written in terms of next period's capital as the control variable) is:

$$v(K) = \max_{K' \in [0,(1-\delta)K+f(K)]} u(f(K) - (1-\delta)K - K') + \beta v(K')$$

In this case, there is a natural upper bound for K, found by solving the equation $\overline{K} = (1-\delta)\overline{K} + f(\overline{K})$. (Such a solution exists because $\frac{d}{dk}(1-\delta)\overline{K} + f(\overline{K})$ is greater than 1 at 0 and asymptotically $(1-\delta)$ by the Inada conditions; the solution occurs when $(1-\delta)\overline{K} + f(\overline{K})$ crosses the 45-degree line.) This ensures that the constraint correspondence is compact and that f, u, and v are bounded as well. We find $v(K)$ as the fixed point of the mapping on bounded functions on $[0, \overline{K}]$, since the mapping satisfies Blackwell's conditions.

The first order conditions are:

$$u'(C) = \beta v'(K')$$

where C is defined as a function of K' by the resource constraints. Using the envelope theorem,

$$v'(K) = u'(C)((1-\delta) + f'(K))$$

Combining the two yields the *consumption Euler equation*:

$$u'(C) = \beta((1-\delta) + f'(K))u'(C')$$

This shows that optimal consumption depends on the product of capital and the rate of depreciation. (The equilibrium rate of return on capital can be defined as $1 + r = 1 - \delta + f'(K)$.)

Suppose $u(C) = \ln(C)$, $f(K) = K^\alpha$ for some $\alpha \in (0,1)$, and $\delta = 1$. In this case, the Euler equation is:

$$\frac{1}{C_t} = \beta(\alpha(K_{t+1})^{\alpha-1})\frac{1}{C_{t+1}}$$

We guess that $C_t = (1-s)K_t^\alpha$, and therefore $K_{t+1} = sK^\alpha$. This satisfies the Euler equation:

$$\begin{aligned} \frac{1}{(1-s)K_t^\alpha} &= \beta\alpha(K_{t+1})^{\alpha-1}\left(\frac{1}{(1-s)(K_{t+1})^\alpha}\right) \\ s &= \beta\alpha \end{aligned}$$

and therefore this must be the optimal policy. This also defines the optimal saving rate.

In the steady state, $K_{SS} = sK_{SS}^\alpha$, and $K_{SS} = s^{\frac{1}{1-\alpha}}$. If $K_t < K_{SS}$, then $\frac{K_{t+1}}{K_t} = sK_t^{\alpha-1} > 1$, and capital moves up toward the steady state. In addition, if $K_t > K_{SS}$, then capital moves toward the steady state. Thus, the steady state is stable. The interest rate in any period is $1 + r = \alpha(K_{t+1})^{\alpha-1}$. In the steady state, this is $1 + r_{SS} = \frac{\alpha}{s} = \frac{1}{\beta}$. If $K > K_{SS}$, then $r < r_{SS}$; if $K < K_{SS}$ then $r > r_{SS}$.

Suppose we instead have $u(C) = \frac{C^{1-\sigma}-1}{1-\sigma}$ (this is constant relative risk aversion, $\sigma = \frac{Cu''}{u'}$, so that utility is more concave with larger values of σ; this also assumes that consumers treat risk across time and risk across space the same, which may not be right). We use log linearization on the constraints, where the lower case is the log deviation from steady state:

$$\begin{aligned}
K_{SS}k_{t+1} &= (1-\delta)K_{SS}k_t + I_{SS}i_t \\
Y_{SS}y_t &= C_{SS}c_t + I_{SS}i_t \\
Y_{SS}y_t &= \alpha K_{SS}^{\alpha-1}K_{SS}k_t \\
y_t &= \alpha k_t
\end{aligned}$$

(The last equation follows because $Y_{SS} = K_{SS}^{\alpha}$.) Log-linearizing the Euler equation:

$$-\sigma C_{SS}^{-\sigma}c_t = -\sigma C_{SS}^{-\sigma}c_{t+1}\beta\left((1-\delta) + \alpha K_{SS}^{\alpha-1}\right) + \beta C_{SS}^{-\sigma}\alpha(\alpha-1)K_{SS}^{\alpha-1}k_{t+1}$$

We may then substitute to find a difference equation for capital:

$$\frac{1}{\beta}k_t + \left(\frac{C_{SS}\beta\alpha(\alpha-1)K_{SS}^{\alpha-2}}{\sigma} - 1 - \frac{1}{\beta}\right)k_{t+1} + k_{t+2} = 0$$

Using this difference equation and initial guesses for k_t, k_{t+1}, we may use *forward induction* to work out k_{t+2} and future values. These values should return to the steady state; if they diverge, then one should adjust k_{t+1}, because that was not the optimal choice for capital in the next period. Also, one can choose k_{t+2}, k_{t+1} near the steady state and use *backward induction* to find a path that would have come to that point.

3.2 Stochastic growth

Assumptions:

- The consumer maximizes $\sum_{i=0}^{\infty} \beta^i E(u(C_{t+i}, L_{t+i}))$, where C_t is consumption and L_t is leisure.
- Production is defined by $Y_t = A_t F(K_t, N_t)$, where A_t is a Markov chain describing total factor productivity, K_t is capital, and N_t is labor.
- Capital grows according to $K_{t+1} = (1-\delta)K_t + I_t$, and we must have $K_{t+1} \in [0, (1-\delta)K_t + Y_t]$.
- We have resource and time constraints: $Y_t = C_t + I_t$ and $N_t + L_t = 1$.

The state variables are K_t (endogenous) and A_t (exogenous). In each period we choose N_t as well (which then specifies L_t, I_t, C_t, Y_t as we choose K_{t+1}), but we do not need N_t to describe future states. This yields the Bellman equation:

$$v(K, A) = \max_{K' \in [0,(1-\delta)K+F(K,N)], N \in [0,1]} u((1-\delta)K + AF(K,N) - K', 1-N) + \beta E(v(K', A')|A)$$

Note that Γ is a joint constraint correspondence on (K', N).

There are now two first order conditions (one for each control variable) and one application of the envelope theorem:

$$\begin{aligned} u_1(C, L) &= \beta E(V_1(K', A')|A) \\ u_1(C, L)AF_2(K, N) &= u_2(C_t, L_t) \\ V_1(K, A) &= u_1(C, L)(1 - \delta + AF_1(K, N)) \end{aligned}$$

where $C = (1-\delta)K + AF(K, N) - K'$ and $L = 1 - N$. Note that the first order condition based on K' is intertemporal; it describes the tradeoff of more consumption today versus having more capital tomorrow. In contrast, the first order condition based on N is intratemporal; it describes the tradeoff between more leisure today and more consumption (or capital) today. The first and third equations can be combined to find the Euler equation for consumption:

$$u_1(C_t, L_t) = E(u_1(C_{t+1}, L_{t+1})(1 - \delta + A_{t+1}F_1(K_{t+1}, N_{t+1}))|A_t)$$

Other combinations can be use to find intertemporal leisure substitution and other possible substitutions.

We may log-linearize this model, in order to find the effect of technology shocks, for example. The movements of variables will depend on both the Euler equations and the constraints.

Using log-linearization, we note that:

$$\begin{aligned} Y_{SS}y_t &= C_{SS}c_t + I_{SS}i_t \\ y_t &= \frac{C_{SS}}{Y_{SS}}c_t + \frac{I_{SS}}{Y_{SS}}i_t \end{aligned}$$

This shows that the percentage deviation in output is a weighted average of the consumption and investment deviations.

In the steady state, $I_{SS} = \delta K_{SS}$.

4 Consumption Models

Assumptions:

- Consumers receive utility, $u(C_t)$, in each period and have a discount factor $\beta < 1$. Consumers want to maximize the present discounted value of utility.
- Consumers begin with initial assets A_0.
- The rate of return, $R_t = 1 + r_t$, and the income process, Y_t, are Markov. Rates of return from this period to the next may be known this period (like bonds) or may be unknown (like stocks).
- The budget constraint depends on whether the consumer can borrow, the effect of uncertainty on borrowing (and the enforcement mechanism for repayment), and whether the problem has a finite or infinite horizon.

If there is a two-period horizon with certainty and no borrowing, then we know Y_1, Y_2, R. If there is no borrowing, the constraints are:

$$\begin{aligned} 0 &\leq C_1 \leq A_1 + Y_1 \\ 0 &\leq C_2 \leq R(A_1 + Y_1 - C_1) + Y_2 \end{aligned}$$

On the other hand, if there is borrowing (and we note that the second constraint will always be binding if u is increasing), then we have the constraint $C_1 + \frac{C_2}{R} = A_1 + Y_1 + \frac{Y_2}{R}$. We see that consumption over time is the sum of initial assets and discounted human capital. In general, consumption must always be less that total wealth, $W_t = A_t + H_t$, where H_t is the present discounted value of human capital. This requires that $\frac{A_{t+1}}{R_{t+1}} > 0$ and $\lim_{T \to \infty} \frac{A_{T+1}}{R_{T+1}} \geq 0$.

If there is uncertainty in income, then a reasonable borrowing constraint is that one must be able to pay back everything in any state of the world. This means that one can only borrow against the discounted human capital in the worst case: $H_{min} = \sum_{j=0}^{\infty} \frac{1}{R_{max}^j} Y_{min}$ (where Y_{min} and R_{max} are well-defined in a Markov chain).

If we assume no borrowing and that R is known at time t, then the Bellman equation is:

$$v(A, Y, R) = \max_{C \in [0, A+Y]} u(C) + \beta E(v(R(A + Y - C), Y', R')|RY)$$

We must assume that $u(C)$ is bounded on $[0, A + Y]$. We could also bound the problem by considering the control variable $\frac{C}{A}$ instead, by finding an upper bound for A (as we did for capital in the growth model) or by just assuming everything will work out.

We may write the Bellman equation in terms of next period's assets as the control variable as well. If R is not known ahead of time (so that assets tomorrow are $R_{t+1}(A_t + Y_t - C_t)$, which is unknown at time t), then we may use the alternative state variable, $S_t = A_t + Y_t - C_t$, which keeps the endogenous and exogenous state variables separate.

The first order condition (no matter which state variables are used in the Bellman equation) is:

$$u'(C) = \beta E(RV_1(R(A + Y - C), R', Y')|R, Y)$$

Using the envelope theorem (most easily calculated with A' or S' as the choice variable):

$$V_1(A, R, Y) = u'(C)$$

This yields the Euler equation, $u'(C) = \beta E(Ru'(C')|R, Y)$. If the consumer is constrained at $C = Y + A$ (this is a *liquidity constraint*), then this becomes:

$$u'(C) \begin{cases} = \beta RE(V(A', Y', R')|Y, R) & A' \geq 0 \\ > \beta RE(V(A', Y', R')|Y, R) & A' = 0 \end{cases}$$

If the constraint is binding, then $A' = 0$ for small changes in A as well. This means that $E(V_1(0, Y', R')|R, Y)$ is constant for the consumer (because nothing changes), and we will still have $V_1(A,R) = u'(C)$ even when the constraint binds. Thus, we have the Euler equation for the constrained consumer, $u'(C) > \beta RE(u'(C'))$.

4.1 Linear-Quadratic Model of Consumption

Suppose $u(C)$ is quadratic, $R = 1 + r$ is known, Y_t is ARIMA, and $\beta R = 1$. Then, the Euler equation is just $C_t = E_t(C_{t+1})$, or, in terms of A_t:

$$\begin{aligned}
A_t - \frac{A_{t+1}}{R} &= E_t(A_{t+1} + Y_{t+1} - \frac{A_{t+2}}{R}) \\
A_t - (1+\frac{1}{R})A_{t+1} + E_t(\frac{A_{t+2}}{R}) &= -Y_t + E_t(Y_{t+1})
\end{aligned}$$

We find paths for the stochastic difference equation using lag polynomials:

$$\begin{aligned}
E_t((B^2 - (1+\frac{1}{R})B + \frac{1}{R})A_{t+2}) &= E_t(-Y_t + Y_{t+1}) \\
E_t((B-1)(B-\frac{1}{R})A_{t+2}) &= E_t(-Y_t + Y_{t+1}) \\
E_t((B-1)(1-\frac{1}{BR})A_{t+1}) &= E_t(-Y_t + Y_{t+1}) \\
E_t((B-1)A_{t+1}) &= E_t\left(\frac{-Y_t + Y_{t+1}}{1-\frac{1}{BR}}\right) \\
A_t - A_{t+1} &= E_t\left(\sum_{j=0}^{\infty}(\frac{1}{BR})^j(Y_t - Y_{t+1})\right) \\
A_t &= A_{t+1} + E_t\left(\sum_{j=0}^{\infty}\frac{1}{R^j}(Y_{t+j} - Y_{t+j+1})\right)
\end{aligned}$$

We may also substitute this expression into the definition of consumption:

$$\begin{aligned}
C_t &= A_t + Y_t - \frac{A_{t+1}}{R} \\
&= A_t + Y_t - \frac{1}{R}\left(A_t - E_t\left(\sum_{j=0}^{\infty}\frac{1}{R^j}(Y_{t+j} - Y_{t+j+1})\right)\right) \\
&= A_t(1-\frac{1}{R}) + (1-\frac{1}{R})E_t\left(\sum_{j=0}^{\infty}\frac{1}{R^j}Y_{t+j}\right) \\
&= \frac{r}{1+r}\left(A_t + E_t\left(\sum_{j=0}^{\infty}\frac{1}{(1+r)^j}Y_{t+j}\right)\right)
\end{aligned}$$

This shows that one consumes from expected total assets as an annuity, which is Milton Friedman's *permanent income hypothesis*. Note that, in this model, the optimal policy is the same whether one is certain or uncertain about future income; this is called *certainty equivalence*, and happens in linear-quadratic models because marginal utility is linear and not concave.

This expression for consumption allows us to calculate the response to an unexpected one unit change in income. If $Y_t = \rho Y_{t-1} + \epsilon_t$, then $\Delta c_t = \frac{r}{1+r}(\epsilon_t + \frac{\rho}{1+r}\epsilon_t + \frac{\rho^2}{(1+r)^2}\epsilon_t + ...) = \frac{r}{1+r-\rho}\epsilon_t$, and consumption is predicted to be a random walk. More complicated forms of Y_t may make consumption more volatile than income.

This form of the model can be tested against micro or macro data. The accepted wisdom is that consumption has *excess smoothness*, because it does not respond enough to innovations in permanent income, and *excess sensitivity* to variables that should already be in the information set. Some of this can be explained by liquidity constraints (because people cannot respond to expected future income shocks until their receive the income) or delays in learning information. Another test of the model (Zeldes, 1981) assumes that only poor people have liquidity constraints and then tests that equality holds for rich people and only sometimes holds for poor people.

Hall (1978) also tested this model, using only the Euler equation form, by estimating $C_{t+1} = \alpha_0 + \alpha_1 C_t + \gamma X_t + \epsilon_{t+1}$, where X_t is any other information known at time t; the coefficient on this other information should be 0. This is a reduced form test of the previous model.

4.2 Other applications of the Euler equation

(Hansen and Singleton, 1988) If the model is correct, we may estimate β, σ (assuming constant relative risk aversion for the utility function). One way to estimate this is by deriving moment conditions from the Euler equation, $E_t((\beta R_{t+1}\frac{C_{t+1}^{-\sigma}}{C_t^{-\sigma}} - 1)X_t) = 0$, where X_t is any information known at time t. (They defined R_t using stock returns.)

5 McCall Search Models

Suppose we have an unemployed worker searching for a job. In each period, $t = 1, 2, ...$, the worker receives a single wage offer, w_t, from a known distribution, $F(w)$, on bounded support, $[0, B]$. Each period, the worker may accept the offer and work forever at the wage, w_t, or may reject the offer, receive unemployment insurance, c, that period, and draw again next period. We assume that future earnings are discounted at $\beta < 1$ and draws from $F(w)$ are independent and identically distributed. The worker wants to maximize the present discounted value of income. We want to find the optimal decision for each wage draw.

Let $V(w)$ be the expected discounted value of future earnings if the optimal decision (for draw w) is made. Then, the Bellman equation is:

$$\begin{aligned} V(w) &= \max\{accept, reject\} \\ &= \max\{\frac{w}{1-\beta}, c + \beta E(V(w'))\} \end{aligned}$$

The value function is bounded because $0 \leq V(w) \leq \max\{\frac{B}{1-\beta}, \frac{c}{1-\beta}\}$. We write the associated mapping of functions as

$$Tg(w) = \max\{w + \beta g(w), c + \beta E(g(w'))\}$$

(The value function just simplifies the first term, using the fact that the worker stays at a job forever.) T is bounded, and preserves continuity, since the maximum of two continuous functions is continuous. T also satisfies Blackwell's conditions, so V exists and is continuous.

This can also be written as two value functions (and therefore two mappings), one for the case where the worker already has a job and the other for the case where the worker is still unemployed:

$$\begin{aligned} V_{job}(w) &= w + \beta V_{job}(w) \\ V_{un}(w) &= \max\{w + \beta V_{job}(w), c + \beta E(V_{un}(w'))\} \\ T\begin{bmatrix} g_0 \\ g_1 \end{bmatrix}(w) &= \begin{bmatrix} \max\{w + \beta g_1(w), c + \beta E(g_0(w'))\} \\ w + \beta g_1(w) \end{bmatrix} \end{aligned}$$

We maximize the value function by accepting when $\frac{w}{1-\beta} \geq c + \beta E(V(w))$; this identifies $\overline{w}$ such that we should accept any $w > \overline{w}$. This is the first order condition for $\overline{w}$:

$$\frac{\overline{w}}{1-\beta} = c + \beta E(V(w))$$

Since $V(w)$ does not depend on time,

$$\overline{w} + \frac{\beta}{1-\beta}\overline{w} = c + \beta\left(\int_0^{\overline{w}} \frac{\overline{w}}{1-\beta} dF(w) + \int_{\overline{w}}^{B} \frac{w}{1-\beta} dF(w)\right)$$

(the first term is if we reject the offer, the second is if we accept the offer). This equation implicitly defines $\overline{w}$, so we could calculate it for a specified F.

Definition A distribution, $F_2(w)$, is a *mean-preserving spread* of $F_1(w)$ if $\int_0^B (F_2(w) - F_1(w))dw = 0$ and $\int_0^x (F_2(w) - F_1(w))dw \geq 0$ for all $x \in (0, B)$. That is, the two distributions have the same mean, but F_2 has more weight on extreme observations than F_1.

Proposition 5.1 *Let F_2 be a mean-preserving spread of F_1. Then the cutoff wage, $\overline{w}$, is higher for F_2 than F_1.*

Proof Recall that $E(w) = \int_0^B w dF(w) = \int_0^B (1 - F(w))dw$.

Then,

$$\begin{aligned}\int_{\overline{w}}^{B}(w'-\overline{w})dF(w') &= E(w')-\overline{w}-\int_0^{\overline{w}}(w'-\overline{w})dF(w')\\ &= E(w')-\overline{w}+\int_0^{\overline{w}}F(w')dw'\end{aligned}$$

(the last step uses integration by parts).

Using our previous solution, we find that

$$\begin{aligned}\overline{w}-c &= \frac{\beta}{1-\beta}\int_{\overline{w}}^{B}(w'-\overline{w})dF(w')\\ &= \frac{\beta}{1-\beta}(E(w')-\overline{w}+\int_0^{\overline{w}}F(w')dw')\\ &= \beta(E(w')-c+\int_0^{\overline{w}}F(w')dw')\end{aligned}$$

Under a mean-preserving spread, the first two terms do not change, but the third term increases. This means that $\overline{w}$ must increase as well. ∎

Notice that:

$$\begin{aligned}\frac{\beta}{1-\beta}\int_0^{B}(w'-0)dF(w') &= \frac{\beta}{1-\beta}E(w')\\ \frac{\beta}{1-\beta}\int_B^{B}(w'-B)dF(w') &= 0\end{aligned}$$

Since $\frac{\beta}{1-\beta}\int_k^B(w'-k)dF(w')$ is continuous and decreasing in k, it must cross the line $k-c$ at some point; this crossing occurs at $k=\overline{w}$.

Suppose c increases. Then, the point where the two lines cross will be higher, which means that $\overline{w}$ increases. This means that unemployment insurance reduces the incentive to take lower paying jobs.

Extensions of the model include:

- Quits: A worker might be able to quit and search again. If the underlying F has not changed, though, this would never be worth it, since one will still take only jobs with a wage over $\overline{w}$. If one can search on the job with no cost, however, one would take the first job with $w>c$ and then keep looking for better jobs.

- Search with recall: Suppose that a worker may take any previous wage offer. Then, the value function is $V(x_t)$, with $x_t=\max\{w_t,x_{t-1}\}$. Note that $V(X_t)=\max\{\frac{1}{1-\beta}x_t, c+\int V(x')dF^*\}$, where F^* is a modification of the distribution (because of the maximum) and will change from one period to the next. However, note that the distribution changes only in the rejection region. This means that if a wage offer is rejected in the past, it will still be rejected in the future. So the problem has not changed dramatically.