Appendix A

Mathematical Tools

A.1 Constrained Optimization

Consider a maximization problem with one equality constraint and one inequality constraint:

$$\max_x f(x) \text{ s.t. } g(x) = 0, h(x) \leq 0.$$

Functions f, g, h are continuous and differentiable and $x \in \mathbb{R}^n$. Form the Lagrangian $\mathcal{L}(x, \lambda, \mu) \equiv f(x) - \lambda g(x) - \mu h(x)$.

Suppose that $x^\star \in \mathbb{R}^n$ is a local maximizer of f on the constraint set. Suppose furthermore that the Jacobian matrix at $x^\star$ of the binding constraints has full rank. Then, there exist multipliers $\lambda^\star$ and $\mu^\star$ such that $\partial\mathcal{L}(x^\star, \lambda^\star, \mu^\star)/\partial x_i = 0$, $i = 1, \ldots, n$; $\mu^\star \geq 0$; $g(x^\star) = 0$; $h(x^\star) \leq 0$; and the complementary slackness condition $\mu^\star h(x^\star) = 0$ is satisfied.

Setting the derivatives of the Lagrangian with respect to the choice variables equal to zero thus yields necessary conditions for a local maximum. The result extends to the case with multiple inequality and/or equality constraints.

A.2 Infinite-Horizon Dynamic Programming

The *sequence problem* is defined by

$$V^\star(a_0) = \max_{\{a_{t+1}\}_{t\geq 0}} \sum_{t=0}^{\infty} \beta^t u(a_t R + w - a_{t+1}) \text{ s.t. } a_0 \text{ given, } a_{t+1} \in A(a_t). \tag{SP}$$

Rather than imposing a no-Ponzi-game condition we require the choice variable to lie in the set described by the correspondence $A(a_t)$. A lower bound on this set implies that debt cannot be rolled over forever (if interest rates are strictly positive) and thus, rules out Ponzi games.

The *Bellman equation* associated with the sequence problem reads

$$V(a_t) = \max_{a_{t+1} \in A(a_t)} \{u(a_t R + w - a_{t+1}) + \beta V(a_{t+1})\} \text{ for all } a_t \in \mathcal{A}. \tag{BE}$$

Since the problem is time-autonomous, the time indices of the state and the control variable in (BE) do not carry significance. $\mathcal{A}$ denotes the state space.

A.2.1 Principle of Optimality

We assume that the set $A(a_t)$ is nonempty for all $a_t \in \mathcal{A}$. We also assume that for all sequences $\{a_{t+1}\}_{t\geq -1}$ that start with $a_0 \in \mathcal{A}$ and satisfy $a_{t+1} \in A(a_t)$, the infinite sequence in (SP) exists and is finite. Under these conditions, $V^\star(a_t) = V(a_t)$ for all $a_t \in \mathcal{A}$. Moreover, a plan $\{a_{t+1}\}_{t\geq 0}$ conditional on $a_0 \in \mathcal{A}$ that attains $V^\star(a_0)$ in (SP) also solves (BE), and the reverse statement holds as well; this is the *Principle of Optimality*.

It follows from the fact that, if the infinite sum in (SP) exists and is finite, then it can be expressed as the sum of a contemporaneous payoff and a continuation payoff, similarly to the two terms on the right-hand side of (BE).

A.2.2 Uniqueness of V

If in addition, the set $\mathcal{A}$ is bounded and complete; $A(a_t)$ is nonempty, bounded, and complete for all $a_t \in \mathcal{A}$; and A and u are continuous, then a unique continuous and bounded function V satisfying (BE) as well as an optimal plan $\{a_{t+1}\}_{t\geq 0}$ solving (SP) or (BE) exist.

Uniqueness follows from results on *contractions*. Note that the right-hand side of (BE) constitutes an operator on the value function, $T(V)$ say: For any value function V on the right-hand side of (BE) the operator returns a value function on the left-hand side. The solution to (BE) then satisfies $V = T(V)$ and the function V constitutes a fixed point of the operator T.

Under the maintained assumptions, the maximization problem on the right-hand side of (BE) has a solution such that T exists and in fact, is continuous. The operator T therefore maps a set of continuous functions into the same set. Moreover, it satisfies Blackwell's sufficient conditions for a contraction.[1] But if an operator constitutes a contraction, then it has a unique fixed point. Moreover, repeated application of the operator generates a sequence of functions that converges to the fixed point. For an arbitrary continuous function V_0, repeated application of the operator thus generates a sequence of functions, $V_0, T(V_0), T(T(V_0)), T(T(T(V_0))), \ldots$, that converges to the fixed point V.

A.2.3 Properties of V

If in addition, u is concave and A convex, then the value function V in (BE) is strictly concave and the optimal plan $\{a_{t+1}\}_{t\geq 0}$ solving (SP) or (BE) is unique.

[1] A metric space $(\mathcal{M}, d)$ is a set $\mathcal{M}$ whose elements can be added, multiplied by scalars, and for pairs of which a norm or distance d is defined. An operator T that maps a metric space into itself is a contraction if there exists a $\rho \in [0, 1)$ such that $d(T(m), T(n)) \leq \rho d(m, n)$ for all $m, n \in \mathcal{M}$.

If in addition, u is strictly increasing in the state and A is monotone, then the value function V in (BE) is strictly increasing.

If in addition, u is continuously differentiable on the interior of its domain, then the value function V in (BE) is differentiable.

A.3 Systems of Linear Difference Equations

Consider a system of two difference equations in the variables x_t and y_t,

$$z_{t+1} = Mz_t \text{ with } z_t \equiv \begin{bmatrix} x_t \\ y_t \end{bmatrix}, \; M \equiv \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

If M is diagonal (i.e., $b = c = 0$) then the two equations are uncoupled: we can solve them independently of each other, yielding $x_t = a^t x_0$ and $y_t = d^t y_0$ for arbitrary x_0, y_0. If M is not diagonal, we can use eigenvalues and -vectors to transform the system and render the equations uncoupled.

Suppose that M has two distinct and real eigenvalues, ρ_1 and ρ_2, with associated eigenvectors, v_1 and v_2, satisfying

$$M[v_1 \; v_2] = [v_1 \; v_2] \begin{bmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{bmatrix} \text{ or } MV = VP.$$

Pre-multiplying the original system $z_{t+1} = Mz_t$ by V^{-1} then yields a transformed, uncoupled system in the vector $Z \equiv V^{-1}z$ with diagonal matrix entries equal to the eigenvalues of M:

$$Z_{t+1} = \begin{bmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{bmatrix} Z_t \text{ and therefore } Z_t = \begin{bmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{bmatrix}^t Z_0.$$

Using $z_t = VZ_t$, the latter equation can be transformed back to yield

$$z_t = V \begin{bmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{bmatrix}^t Z_0 = V \begin{bmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{bmatrix}^t V^{-1} z_0.$$

Note that M^t is given by VP^tV^{-1}. Letting $\varphi_0 \equiv V^{-1}z_0$ we conclude that

$$z_t = V \begin{bmatrix} \rho_1^t & 0 \\ 0 & \rho_2^t \end{bmatrix} \varphi_0 \text{ or } z_t = \varphi_{0[1]}\rho_1^t v_1 + \varphi_{0[2]}\rho_2^t v_2.$$

This first-order difference equation system in z_t has a family of solutions with two degrees of freedom, corresponding to the two elements of z_0 or φ_0. For a definite solution, we need two restrictions. An initial condition for an element of z_0 constitutes such a restriction. If an eigenvalue is unstable then the requirement that system dynamics be stable also implies a restriction; for example, $\rho_1 > 1$ and system stability imply $\varphi_{0[1]} = 0$.

Similar solution strategies are available for cases where the eigenvalues of M are not distinct or where they are complex. The extension to the case with more than two variables is immediate.

A.4 Bibliographic Notes

Simon and Blume (1994, 18, 19), Mas-Colell et al. (1995, M.K), and Acemoglu (2009, A.11) review Lagrangian methods.

Stokey and Lucas (1989, 3, 4) and Acemoglu (2009, 6) cover dynamic programming. Acemoglu (2009, example 6.5) covers the saving problem and discusses an approach to guarantee compactness of $\mathcal{A}$ in that program.

Simon and Blume (1994, 23) cover linear difference equations, including the case of repeated or complex eigenvalues.

Related Topics Stokey and Lucas (1989, 7–9) and Acemoglu (2009, 16) cover dynamic programming under risk.

Appendix B

Technical Discussions

B.1 Transversality Condition in Infinite-Horizon Saving Problem

The household's program is given by

$$\max_{\{a_{t+1}\}_{t\geq 0}} \sum_{t=0}^{\infty} \beta^t u(a_t R_t + w_t - a_{t+1}) \text{ s.t. } a_0 \text{ given, } \lim_{T\to\infty} q_T a_{T+1} \geq 0.$$

Let $\hat{a} \equiv \{\hat{a}_{t+1}\}_{t\geq 0}$ denote a plan that satisfies the Euler equation at all times as well as $\lim_{T\to\infty} q_T \hat{a}_{T+1} = 0$. Let $\bar{a} \equiv \{\bar{a}_{t+1}\}_{t\geq 0}$ denote an alternative plan that satisfies $\lim_{T\to\infty} q_T \bar{a}_{T+1} \geq 0$. We want to show that the former dominates the latter.

For brevity, let $\hat{u}_t \equiv u(\hat{a}_t R_t + w_t - \hat{a}_{t+1})$ and $\hat{u}'_t \equiv u'(\hat{a}_t R_t + w_t - \hat{a}_{t+1})$ and similarly for $\bar{u}_t$ and $\bar{u}'_t$. Strict concavity of u and positive marginal utility imply $\hat{u}_t + R_t \hat{u}'_t(\bar{a}_t - \hat{a}_t) - \hat{u}'_t(\bar{a}_{t+1} - \hat{a}_{t+1}) > \bar{u}_t$ if $\hat{a}_t \neq \bar{a}_t$ or $\hat{a}_{t+1} \neq \bar{a}_{t+1}$. If $\hat{a}_{t+1} \neq \bar{a}_{t+1}$ for some $t \in \{0, \ldots, T\}$, we thus have

$$\sum_{t=0}^{T} \beta^t (\bar{u}_t - \hat{u}_t) < \sum_{t=0}^{T} \beta^t \{R_t \hat{u}'_t(\bar{a}_t - \hat{a}_t) - \hat{u}'_t(\bar{a}_{t+1} - \hat{a}_{t+1})\} = \beta^T \hat{u}'_T(\hat{a}_{T+1} - \bar{a}_{T+1}),$$

where we use $\hat{a}_0 = \bar{a}_0$ and $\hat{u}'_t = \beta R_{t+1} \hat{u}'_{t+1}$. From the Euler equation, $\beta^T \hat{u}'_T = \hat{u}'_0 q_T$. If $\hat{a}$ is not identical to $\bar{a}$, it follows that

$$\lim_{T\to\infty} \sum_{t=0}^{T} \beta^t (\bar{u}_t - \hat{u}_t) < \lim_{T\to\infty} \hat{u}'_0 q_T (\hat{a}_{T+1} - \bar{a}_{T+1}) = \lim_{T\to\infty} -\hat{u}'_0 q_T \bar{a}_{T+1} \leq 0$$

such that $\sum_{t=0}^{\infty} \beta^t \hat{u}_t > \sum_{t=0}^{\infty} \beta^t \bar{u}_t$. Satisfying $\lim_{T\to\infty} q_T \hat{a}_{T+1} = 0$ therefore is optimal.

B.2 Representative Household

According to the *Sonnenschein-Mantel-Debreu theorem* aggregate demand functions possess less structure than the underlying demand functions of individual households.

While both aggregate and individual demand functions are homogeneous of degree zero and satisfy Walras' Law the former generally do not satisfy the restrictions imposed by individual rationality, e.g., consistency with the weak axiom of revealed preference and symmetry of the Slutsky matrix. This is a consequence of the fact that aggregate demand blends household specific wealth effects (of price changes) which may differ when endowments and preferences are heterogeneous.

Since aggregate demand functions cannot generally be viewed as representing the demand of a single, fictional decision maker we need to impose additional assumptions if we want to interpret macroeconomic aggregates through the lens of demand theory. One strategy is to assume that all households in the economy literally are alike such that average outcomes are identical to the individual choices of the typical, *representative household*. Other strategies require that we restrict household preferences or the wealth distribution. We consider two examples.

Suppose first that preferences of each household h admit an *indirect utility function*, v^h, of the *Gorman form*,

$$v^h(p, w^h) = a^h(p) + b(p)w^h,$$

where p denotes the price vector; w^h denotes household wealth; and a^h and b are differentiable functions. Using Roy's Identity household h's (vector valued) demand function, $x^h(p, w^h)$, satisfies

$$x^h(p, w^h) = -\frac{\partial v^h(p, w^h)/\partial p}{\partial v^h(p, w^h)/\partial w^h} = -\frac{(\partial a^h(p)/\partial p) + (\partial b(p)/\partial p)w^h}{b(p)}.$$

Accordingly, aggregate demand is given by

$$\sum_h x^h(p, w^h) = -\frac{\sum_h \partial a^h(p)/\partial p}{b(p)} - \frac{\partial b(p)/\partial p}{b(p)} w,$$

where $w = \sum_h w^h$.

Note that aggregate demand does not depend on the distribution of wealth because the function $b(p)$ is identical across households that is, for a fixed price vector and for each good, all households have parallel, straight Engel curves.[1] As a consequence, aggregate demand can be viewed as solving the program of a fictional, rational household—a *positive representative household*—with indirect utility function

$$v(p, w) = \sum_h a^h(p) + b(p)w.$$

By Roy's Identity, the demand of this fictional households satisfies

$$-\frac{\partial v(p, w)/\partial p}{\partial v(p, w)/\partial w} = -\frac{\sum_h \partial a^h(p)/\partial p}{b(p)} - \frac{\partial b(p)/\partial p}{b(p)} w,$$

[1]The reverse implication holds as well: If aggregate demand is to be independent of the wealth distribution, for any price vector, then households must have preferences that admit indirect utility functions of the Gorman form with identical $b(p)$. Identical, homothetic or quasilinear preferences satisfy this restriction.

which exactly corresponds to aggregate demand.

Suppose next that each household h has generalized CIES preferences and maximizes

$$\sum_{t=0}^{\infty} \beta^t \frac{(\phi + \chi c_t^h)^{1-\sigma}}{1-\sigma},$$

where $\sigma > 0, \phi \geq 0, \chi > 0$; β denotes the discount factor; and c_t^h denotes consumption of household h at date t. Let p_t denote the price of consumption at date t with $p_0 = 1$. The Euler equation and intertemporal budget constraint of household h then imply

$$\begin{aligned}(\phi + \chi c_t^h) &= \left(\beta^{-t} p_t\right)^{-\frac{1}{\sigma}} (\phi + \chi c_0^h), \\ \sum_{t=0}^{\infty} c_t^h p_t &= w^h.\end{aligned}$$

Due to market completeness and the functional form assumption for preferences the equilibrium conditions can be summed across households.[2] Letting $c_t \equiv \sum_h c_t^h$ we have

$$\begin{aligned}(\phi + \chi c_t) &= \left(\beta^{-t} p_t\right)^{-\frac{1}{\sigma}} (\phi + \chi c_0), \\ \sum_{t=0}^{\infty} c_t p_t &= w,\end{aligned}$$

which constitute the optimality conditions of a positive representative household that chooses aggregate consumption subject to the aggregate budget constraint.

An economy admits a *normative representative household* when it admits a positive representative household whose preference relation constitutes a meaningful welfare measure. Suppose that there exists a wealth distribution rule which maps prices and aggregate wealth into household wealth, $\{w^h(p, w)\}_h$, such that aggregate demand, $\sum_h x^h(p, w^h(p, w))$, solves the program of a positive representative household. Moreover, suppose that for some social welfare function, W, the wealth distribution is optimal that is, for any (p, w) the wealth distribution rule solves the program

$$\max_{\{w^h\}_h} W\left(\{v^h(p, w^h)\}_h\right) \text{ s.t. } \sum_h w^h = w.$$

The positive representative household then is a normative representative household and the social welfare function is the utility function of the normative representative household.

When preferences admit indirect utility functions of the Gorman form then the positive representative household is a normative one. This follows from the fact that any wealth distribution is optimal with respect to the *utilitarian social welfare function* which gives equal weight to all households.[3]

[2] Market completeness implies that all households share the same marginal rates of substitution. The functional form assumption implies that the Euler equation is linear in consumption. The budget constraint also is linear in consumption.

[3] Since $b(p)$ is the same for all households changes in the wealth distribution do not affect social welfare when the social welfare function is utilitarian.

B.3 Transversality Condition in Infinite-Horizon Planner Problem

Let $g(k_t) \equiv k_t(1-\delta) + f(k_t, 1)$ with f denoting the neoclassical production function. Note that the function g is strictly concave. The program is given by

$$\max_{\{k_{t+1}\}_{t\geq 0}} \sum_{t=0}^{\infty} \beta^t u(g(k_t) - k_{t+1}) \text{ s.t. } k_0 \text{ given}, \ \ k_{t+1} \geq 0.$$

Let $\hat{k} \equiv \{\hat{k}_{t+1}\}_{t\geq 0}$ denote a plan that satisfies the Euler equation and complementary slackness condition, $\hat{u}'_t = \beta g'(\hat{k}_{t+1})\hat{u}'_{t+1} + \hat{\mu}_t$ and $\hat{\mu}_t \hat{k}_{t+1} = 0$ respectively, at all times, as well as $\lim_{T\to\infty} \beta^T \hat{u}'_T \hat{k}_{T+1} = 0$. Let $\bar{k} \equiv \{\bar{k}_{t+1}\}_{t\geq 0}$ denote an alternative plan that satisfies $\lim_{T\to\infty} \beta^T \bar{u}'_T \bar{k}_{T+1} \geq 0$. Here, we let $\hat{u}_t \equiv u(g(\hat{k}_t) - \hat{k}_{t+1})$ and $\hat{u}'_t \equiv u'(g(\hat{k}_t) - \hat{k}_{t+1})$ and similarly for $\bar{u}_t$ and $\bar{u}'_t$. We want to show that $\hat{k}$ dominates $\bar{k}$.

Strict concavity of u and g and positive marginal utility imply $\hat{u}_t + \hat{u}'_t g'(\hat{k}_t)(\bar{k}_t - \hat{k}_t) - \hat{u}'_t(\bar{k}_{t+1} - \hat{k}_{t+1}) > \bar{u}_t$ if $\hat{k}_t \neq \bar{k}_t$ or $\hat{k}_{t+1} \neq \bar{k}_{t+1}$. If $\hat{k}_{t+1} \neq \bar{k}_{t+1}$ for some $t \in \{0, \ldots, T\}$, we thus have

$$\begin{aligned} \sum_{t=0}^{T} \beta^t (\bar{u}_t - \hat{u}_t) &< \sum_{t=0}^{T} \beta^t \{\hat{u}'_t g'(\hat{k}_t)(\bar{k}_t - \hat{k}_t) - \hat{u}'_t(\bar{k}_{t+1} - \hat{k}_{t+1})\} \\ &= \beta^T \hat{u}'_T(\hat{k}_{T+1} - \bar{k}_{T+1}) - \sum_{t=0}^{T-1} \beta^t \hat{\mu}_t \bar{k}_{t+1}, \end{aligned}$$

where we use $\hat{k}_0 = \bar{k}_0$, the Euler equation, and the complementary slackness condition. If $\hat{k}$ is not identical to $\bar{k}$, it follows that

$$\begin{aligned} \lim_{T\to\infty} \sum_{t=0}^{T} \beta^t (\bar{u}_t - \hat{u}_t) &< \lim_{T\to\infty} \left\{ \beta^T \hat{u}'_T(\hat{k}_{T+1} - \bar{k}_{T+1}) - \sum_{t=0}^{T-1} \beta^t \hat{\mu}_t \bar{k}_{t+1} \right\} \\ &= -\lim_{T\to\infty} \left\{ \beta^T \hat{u}'_T \bar{k}_{T+1} + \sum_{t=0}^{T-1} \beta^t \hat{\mu}_t \bar{k}_{t+1} \right\} \leq 0, \end{aligned}$$

since marginal utility, the capital stock, and the multiplier are all weakly positive. We conclude that $\sum_{t=0}^{\infty} \beta^t \hat{u}_t > \sum_{t=0}^{\infty} \beta^t \bar{u}_t$. Satisfying $\lim_{T\to\infty} \beta^T \hat{u}'_T \hat{k}_{T+1} = 0$ therefore is optimal.

B.4 Non-Expected Utility

Let U_t denote the utility derived from a consumption sequence starting at date t, $\{c_t(\epsilon^t), c_{t+1}(\epsilon^{t+1}), \ldots\}$. For legibility, we omit the reference to histories when there is no danger of confusion. We consider *recursive preference* specifications of the form

$$U_t = \mathcal{A}\left(c_t, \mathcal{B}^{-1}\left(\mathbb{E}_t[\mathcal{B}(U_{t+1})]\right)\right),$$

where the function $\mathcal{A}$ is homogeneous of degree one and the functions $\mathcal{A}$ and $\mathcal{B}$ are increasing and concave. Note that $\mathcal{B}^{-1}(\mathbb{E}_t[\mathcal{B}(U_{t+1})])$ is the *certainty equivalent* of the random variable U_{t+1}, conditional on information at date t. That is,

$$\mathcal{B}(\text{CE}_t[U_{t+1}]) = \mathbb{E}_t[\mathcal{B}(U_{t+1})],$$

where CE_t denotes the certainty equivalent operator.

We adopt the functional form assumptions

$$\begin{aligned}\mathcal{A}(c, C) &\equiv \left((1-\beta)c^{1-\sigma} + \beta C^{1-\sigma}\right)^{\frac{1}{1-\sigma}}, \ \beta \in (0,1), \sigma > 0, \sigma \neq 0, \\ \mathcal{B}(U) &\equiv \frac{U^{1-\gamma}}{1-\gamma}, \ \gamma > 0, \gamma \neq 0,\end{aligned}$$

implying

$$U_t = \left((1-\beta)c_t^{1-\sigma} + \beta\left(\mathbb{E}_t\left[U_{t+1}^{1-\gamma}\right]\right)^{\frac{1-\sigma}{1-\gamma}}\right)^{\frac{1}{1-\sigma}}. \tag{B.1}$$

Since $\mathcal{B}^{-1}(\mathbb{E}_t[\mathcal{B}(U_{t+1})])$ is the certainty equivalent of U_{t+1}, γ represents a measure of risk aversion. In contrast, σ is an (inverse) measure of the intertemporal elasticity of substitution. To see this suppose that the consumption sequence is deterministic. Equation (B.1) then reduces to

$$U_t = \left((1-\beta)c_t^{1-\sigma} + \beta U_{t+1}^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$$

or, defining $\bar{U}_t \equiv U_t^{1-\sigma}$,

$$\bar{U}_t = (1-\beta)c_t^{1-\sigma} + \beta\bar{U}_{t+1} = (1-\beta)\sum_{s=t}^{\infty}\beta^{s-t}c_s^{1-\sigma}.$$

Absent risk, U_t thus exhibits a constant intertemporal elasticity of substitution, σ^{-1}.

When $\gamma = \sigma$ we recover standard expected utility. Equation (B.1) then implies

$$U_t = \left((1-\beta)c_t^{1-\sigma} + \beta\mathbb{E}_t\left[U_{t+1}^{1-\sigma}\right]\right)^{\frac{1}{1-\sigma}}$$

or

$$\bar{U}_t = (1-\beta)\sum_{s=t}^{\infty}\beta^{s-t}\mathbb{E}_t\left[c_s^{1-\sigma}\right].$$

We conclude that relative to the benchmark specification with $\gamma = \sigma$, the more general *non-expected utility* specification (B.1) decouples risk aversion and the intertemporal elasticity of substitution.

Differences between γ and σ imply a preference for early or late resolution of risk. To establish this, we consider an environment with a single source of risk namely a permanent shock to consumption at date $t = 1$. Specifically, we assume that consumption at date $t = 0$ is certain and equals c while consumption at dates $t \geq 1$ equals either x or y, with equal probability. We compare two scenarios. In the first scenario (early resolution), risk is resolved at date $t = 0$ that is, the household learns at date $t = 0$ whether consumption starting at date $t = 1$ equals x or y. In the second scenario (late resolution), the household only learns this at date $t = 1$.

We determine first the utility at date $t = 1$ conditional on knowing that current and future consumption equals x, U^x say. Due to stationarity,

$$U^x = \left((1-\beta)x^{1-\sigma} + \beta(U^x)^{1-\sigma}\right)^{\frac{1}{1-\sigma}},$$

implying $U^x = x$. Similarly, utility conditional on knowing that consumption equals y is given by $U^y = y$. Turning next to the utility at date $t = 0$, consider first the scenario with *early resolution of risk*. Utility at date $t = 0$ conditional on learning that consumption will equal x or y, respectively, is given by

$$\begin{aligned} U_0|x &= \left((1-\beta)c^{1-\sigma} + \beta x^{1-\sigma}\right)^{\frac{1}{1-\sigma}}, \\ U_0|y &= \left((1-\beta)c^{1-\sigma} + \beta y^{1-\sigma}\right)^{\frac{1}{1-\sigma}}, \end{aligned}$$

and utility before information is revealed equals the certainty equivalent of the risky utility U_0,

$$U^{\text{early}} = \left(\frac{(U_0|x)^{1-\gamma} + (U_0|y)^{1-\gamma}}{2}\right)^{\frac{1}{1-\gamma}}.$$

In contrast, utility at date $t = 0$ with *late resolution of risk* equals

$$U^{\text{late}} = \left((1-\beta)c^{1-\sigma} + \beta\left(\frac{x^{1-\gamma} + y^{1-\gamma}}{2}\right)^{\frac{1-\sigma}{1-\gamma}}\right)^{\frac{1}{1-\sigma}}.$$

When $\gamma = \sigma$ (the expected utility case) the timing of the resolution of risk does not affect utility, $U^{\text{early}} = U^{\text{late}}$. When $\gamma \neq \sigma$, in contrast, the timing does have welfare consequences since

$$\mathcal{A}(c, \text{CE}_0[U_1]) \neq \text{CE}_0[\mathcal{A}(c, U_1)].$$

Figure B.1 plots the level curves of $U^{\text{early}} - U^{\text{late}}$ against γ and σ; darker areas indicate higher values. Early (late) resolution is preferred when $\gamma > (<)\sigma$.

Returning to the general specification (B.1), we compute the stochastic discount factor at date t, history ϵ^t. Differentiating U_t with respect to $c_t(\epsilon^t)$ and $c_{t+1}(\epsilon^{t+1})$, respectively, yields

$$\begin{aligned} \frac{\partial U_t}{\partial c_t} &= U_t^{\sigma}(1-\beta)c_t^{-\sigma}, \\ \frac{\partial U_t}{\partial c_{t+1}} &= U_t^{\sigma}\beta\left(\text{CE}_t(U_{t+1})\right)^{\gamma-\sigma}\left(U_{t+1}\right)^{-\gamma} h(\epsilon^{t+1}|\epsilon^t)\frac{\partial U_{t+1}}{\partial c_{t+1}}, \end{aligned}$$

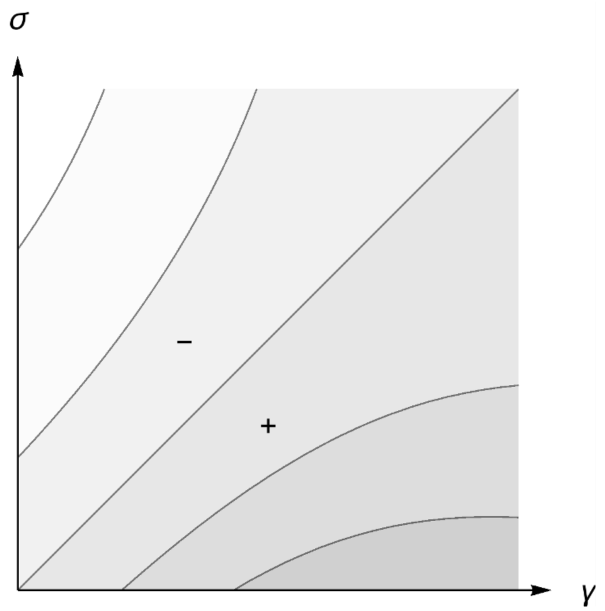

Figure B.1: Preference for early or late resolution of risk: $U^{\text{early}} - U^{\text{late}}$.

where c_{t+1} and U_{t+1} denote consumption and continuation utility in history ϵ^{t+1}, and $h(\epsilon^{t+1}|\epsilon^t)$ denotes the probability (density) of ϵ^{t+1} conditional on ϵ^t.

Define the stochastic discount factor, $m_{t+1}(\epsilon^{t+1})$, (as usual) as

$$\begin{aligned} m_{t+1}(\epsilon^{t+1}) &\equiv \frac{\partial U_t/\partial c_{t+1}}{\partial U_t/\partial c_t}\frac{1}{h(\epsilon^{t+1}|\epsilon^t)} = \frac{U_t^\sigma \beta\, (\text{CE}_t(U_{t+1}))^{\gamma-\sigma}\, (U_{t+1})^{-\gamma}\, \frac{\partial U_{t+1}}{\partial c_{t+1}}}{U_t^\sigma (1-\beta) c_t^{-\sigma}} \\ &= \frac{\beta\, (\text{CE}_t(U_{t+1}))^{\gamma-\sigma}\, (U_{t+1})^{-\gamma}\, U_{t+1}^\sigma (1-\beta) c_{t+1}^{-\sigma}}{(1-\beta)c_t^{-\sigma}} \\ &= \beta \left(\frac{c_{t+1}}{c_t}\right)^{-\sigma} \left(\frac{\text{CE}_t(U_{t+1})}{U_{t+1}}\right)^{\gamma-\sigma}. \end{aligned}$$

When $\gamma = \sigma$ this expression reduces to

$$m_{t+1}(\epsilon^{t+1}) = \beta \left(\frac{c_{t+1}}{c_t}\right)^{-\sigma}.$$

Otherwise the stochastic discount factor and thus, asset prices are a function not only of consumption in the two periods but also of the continuation value and its certainty equivalent.

B.5 Linear Rational Expectations Models

B.5.1 Single Equation Model

Backward- and Forward-Solution of a Difference Equation

Consider the difference equation

$$x_{t+1} = \alpha x_t + s_t,$$

where the variable x_t is endogenous and predetermined and the variable s_t is exogenous. The sequence $\{s_t\}_{t\geq 0}$ is bounded. Iterating the equation backward yields

$$x_{t+1} = s_t + \alpha s_{t-1} + \alpha^2 s_{t-2} + \ldots + \alpha^t s_0 + \alpha^{t+1} x_0.$$

Iterating forward implies

$$x_{t+T} = \alpha^T x_t + \alpha^{T-1} s_t + \alpha^{T-2} s_{t+1} + \ldots + \alpha s_{t+T-2} + s_{t+T-1}.$$

When $|\alpha| < 1$, then $\lim_{T\to\infty} x_{t+T}$ is bounded irrespective of the value x_t. When $|\alpha| > 1$, in contrast, this does not hold true. To see this, define

$$x_t^\star \equiv - \lim_{T\to\infty} \sum_{j=1}^{T} \frac{s_{t+j-1}}{\alpha^j}.$$

Since $\{s_t\}_{t\geq 0}$ is bounded and $|\alpha| > 1$, $x_t^\star$ is well defined. For $x_t = x_t^\star$, $\lim_{T\to\infty} x_{t+T}$ equals zero. But for $x_t = x_t^\star + \Delta$ with $\Delta \neq 0$, x_{t+T} diverges as T increases. We conclude that for $|\alpha| > 1$ there exists a unique value for x_t, namely $x_t^\star$, such that $\{x_t\}_{t\geq 0}$ is bounded.

Rational Expectations Model

Consider next a stochastic (but still bounded) sequence $\{s_t\}_{t\geq 0}$.[4] The model is given by two conditions. First, the difference equation

$$\mathbb{E}_t[y_{t+1}] = \alpha y_t + s_t \quad \text{or} \quad y_{t+1} = \alpha y_t + s_t + \delta_{t+1},$$

where the variable y_t is endogenous and non-predetermined, and δ_{t+1} denotes a forecast error satisfying $\mathbb{E}_t[\delta_{t+1}] = 0$. And second, the restriction that $\lim_{T\to\infty} \mathbb{E}_t[y_{t+T}]$ be bounded.

Iterating forward and using the law of iterated expectations yields

$$\mathbb{E}_t[y_{t+T}] = \alpha^T y_t + \alpha^{T-1} s_t + \mathbb{E}_t[\alpha^{T-2} s_{t+1} + \ldots + \alpha s_{t+T-2} + s_{t+T-1}].$$

When $|\alpha| < 1$, then $\lim_{T\to\infty} \mathbb{E}_t[y_{t+T}]$ is bounded irrespective of the value y_t. That is, the model restrictions do not determine the actual realization of y_t, but only its expectation, $\mathbb{E}_{t-1}[y_t] = \alpha y_{t-1} + s_{t-1}$. The model thus leaves room for a *sunspot shock* to affect y_t.

[4] To simplify the notation, we do not index variables by history.

When $|\alpha| > 1$, in contrast, boundedness of $\lim_{T\to\infty} \mathbb{E}_t[y_{t+T}]$ requires that y_t equals

$$y_t^\star \equiv -\lim_{T\to\infty} \mathbb{E}_t \sum_{j=1}^{T} \frac{s_{t+j-1}}{\alpha^j},$$

for parallel reasons as in the deterministic case. The model determines the expectation $\mathbb{E}_{t-1}[y_t] = \alpha y_{t-1} + s_{t-1}$, and it also determines the actual realization, $y_t = y_t^\star$. As in the case with $|\alpha| < 1$, the forecast error $\delta_t = y_t^\star - \mathbb{E}_{t-1}[y_t]$ is unpredictable. But unlike in that case, it only reflects the effect of new information about $\{s_{t+j}\}_{j\geq 0}$ on $y_t^\star$; the model leaves no room for a sunspot shock to affect y_t.

B.5.2 Multiple Equation Model

The model consists of the system of difference equations

$$\begin{bmatrix} x_{t+1} \\ \mathbb{E}_t[y_{t+1}] \end{bmatrix} = M \begin{bmatrix} x_t \\ y_t \end{bmatrix} + N s_t \tag{B.2}$$

or equivalently,

$$\begin{bmatrix} \underset{(n_x\times 1)}{x_{t+1}} \\ \underset{(n_y\times 1)}{y_{t+1}} \end{bmatrix} = \underset{(n\times n)}{M} \begin{bmatrix} \underset{(n_x\times 1)}{x_t} \\ \underset{(n_y\times 1)}{y_t} \end{bmatrix} + \underset{(n\times n_s)}{N} \underset{(n_s\times 1)}{s_t} + \begin{bmatrix} \underset{(n_x\times 1)}{0} \\ \underset{(n_y\times 1)}{\delta_{t+1}} \end{bmatrix}.$$

There are n_x predetermined variables (including, for example, the capital stock), denoted by x_t; n_y non-predetermined variables (e.g., consumption), denoted by y_t; and n_s exogenous bounded variables (e.g., productivity), denoted by s_t. The number of endogenous variables equals $n = n_x + n_y$, and δ_t denotes a vector of forecast errors. The model imposes the additional restriction that the endogenous variables be bounded.

Using the notation of appendix A.3, matrix M can be represented as the product of matrices that contain its eigenvectors and eigenvalues,

$$M = VPV^{-1} \equiv \underset{(n\times n)}{[v_1\ v_2 \cdots v_n]} \underset{(n\times n)}{\begin{bmatrix} \rho_1 & & 0 \\ & \ddots & \\ 0 & & \rho_n \end{bmatrix}} \underset{(n\times n)}{V^{-1}} \equiv V \begin{bmatrix} \underset{(n_<\times n_<)}{P_{<<}} & \underset{(n_<\times n_>)}{0} \\ \underset{(n_>\times n_<)}{0} & \underset{(n_>\times n_>)}{P_{>>}} \end{bmatrix} V^{-1},$$

where $n_<$ ($n_>$) denotes the number of eigenvalues whose absolute value is weakly smaller than (exceeds) unity, respectively. We assume that the eigenvalues are distinct and real and ordered in ascending absolute value, $|\rho_1| < |\rho_2| < \ldots < |\rho_n|$. Let

$$Z_t \equiv \begin{bmatrix} \underset{(n_<\times 1)}{Z_{<t}} \\ \underset{(n_>\times 1)}{Z_{>t}} \end{bmatrix} \equiv V^{-1} \begin{bmatrix} x_t \\ y_t \end{bmatrix} \text{ and } \begin{bmatrix} \underset{(n_<\times n_s)}{N_<} \\ \underset{(n_>\times n_s)}{N_>} \end{bmatrix} \equiv V^{-1}N. \tag{B.3}$$

Premultiplying equation (B.2) by V^{-1} yields $\mathbb{E}_t[Z_{t+1}] = PZ_t + V^{-1}Ns_t$. Iterating forward, we arrive at

$$\mathbb{E}_t[Z_{t+T}] = P^T Z_t + \sum_{j=0}^{T-1} P^{T-1-j} V^{-1} N \mathbb{E}_t[s_{t+j}].$$

Boundedness ($\lim_{T\to\infty} \mathbb{E}_t[Z_{t+T}] = 0$) thus requires

$$\lim_{T\to\infty} \left((P_{>>})^T Z_{>t} + \sum_{j=0}^{T-1} (P_{>>})^{T-1-j} N_> \mathbb{E}_t[s_{t+j}] \right) = 0$$

or

$$Z_{>t} = -\sum_{j=0}^{\infty} (P_{>>})^{-1-j} N_> \mathbb{E}_t[s_{t+j}].$$

That is, the requirement that system dynamics be stable imposes $n_>$ restrictions on Z_t. The initial conditions for the predetermined variables x_t impose n_x additional restrictions. We may distinguish three cases:

No Solution, $n_y < n_>$

When $n_y < n_>$ (and thus, $n_x > n_<$) then the stability requirement imposes more restrictions than there are non-predetermined variables that could adjust to satisfy them. In general, the model has no solution in this case.

Determinacy, $n_y = n_>$

When $n_y = n_>$ then the stability requirement imposes as many restrictions as there are non-predetermined variables. All endogenous variables thus are uniquely pinned down (as long as the relevant submatrix of V^{-1} is invertible such that a given $Z_{>t}$ implies a unique y_t in (B.3)). The model dynamics satisfy

$$\begin{bmatrix} x_{t+1} \\ y_{t+1} \end{bmatrix} = V \begin{bmatrix} P_{<<} & 0 \\ 0 & 0 \end{bmatrix} V^{-1} \begin{bmatrix} x_t \\ y_t \end{bmatrix} + V \begin{bmatrix} N_< \\ 0 \end{bmatrix} s_t + V \begin{bmatrix} \underset{(n_< \times n_>)}{0} \\ \underset{(n_> \times n_>)}{I} \end{bmatrix} Z_{>t+1}. \tag{B.4}$$

Typically, $y_t \neq \mathbb{E}_{t-1}[y_t]$. The forecast error, δ_t, reflects the effect of new information about $\{s_{t+j}\}_{j\geq 0}$ on $Z_{>t}$.

Indeterminacy, $n_y > n_>$

When $n_y > n_>$ then the stability requirement does not pin down the non-predetermined variables; there are $n_y - n_>$ degrees of freedom. Suppose, for example, that $n_x = n_s = 0$ such that the model reduces to

$$y_{t+1} = My_t + \delta_{t+1}$$

and $Z_{>t} = 0$. In this case, $n_<$ elements of y_t can freely be chosen in each period without triggering explosive dynamics. Although $s_t = 0$ in all periods the forecast errors may be non-zero. In particular, y_t may respond to non-fundamental sunspot shocks.

B.6 Probabilistic Voting

Suppose that two political candidates, a and b, compete for office in a democratic election. Voters are indexed by $i \in [0, I]$. Which candidate a voter supports depends both on the candidate's *policy platform* and on the voter's *ideological* attachment. Voter i supports candidate a if the voter's utility in the equilibrium implemented by a's proposed policy choice (and the continuation policies induced by policy functions of future decision makers) exceeds the utility resulting from b's proposal by more than a threshold value. This threshold value, which reflects the voter's ideological attachment to candidate b, is a random variable with two components: A voter specific, i.i.d. component, ξ^i; and an aggregate component, ξ. Accordingly, the outcome of the vote is *probabilistic* as well.

The voter-specific component is drawn from a uniform distribution with density ϕ^i, $\xi^i \sim U[-1/(2\phi^i), 1/(2\phi^i)]$. A positive ξ^i reflects a permanent ideological bias of voter i in favor of candidate b. The average bias across voters equals zero. The aggregate component is drawn from a uniform distribution with density φ, $\xi \sim U[-1/(2\varphi), 1/(2\varphi)]$. It represents an aggregate shock to ideological attachment which is realized after the candidates have proposed their policy platforms. The sum of the two components represents the total ideological bias of voter i in favor of candidate b in the current election.

Let $V^i(a)$ and $V^i(b)$ denote utility of voter i when the policy platform of candidate a or b, respectively, is implemented. Voter i supports candidate a iff

$$V^i(a) \geq V^i(b) + \xi + \xi^i.$$

Let $\Delta^i \equiv V^i(a) - V^i(b)$. Conditional on ξ, the probability that i supports candidate a equals[5]

$$\text{prob}(\xi^i \leq \Delta^i - \xi) = \frac{1}{2} + \phi^i \times (\Delta^i - \xi)$$

and candidate a's vote share conditional on ξ therefore equals

$$\frac{1}{2} + \frac{\int_i \phi^i \times (\Delta^i - \xi) di}{I}.$$

The unconditional probability that candidate a's vote share exceeds one half thus is given by

$$\text{prob}\left(\frac{\int_i \phi^i \times (\Delta^i - \xi) di}{I} \geq 0\right) = \text{prob}\left(\frac{\int_i \phi^i \Delta^i di}{\int_i \phi^i di} \geq \xi\right) = \frac{1}{2} + \varphi \frac{\int_i \phi^i \Delta^i di}{\int_i \phi^i di}.$$

[5]To keep the notation simple we assume that probabilities are interior.

The probability that candidate b wins the election equals one minus the former expression.

Conditional on b's platform, candidate a's vote share is a continuous function of a's platform (unlike in the median-voter setup); the optimal platform choice maximizes $\int_i \phi^i \Delta^i di$. Similarly, candidate b's optimal choice minimizes this expression (conditional on a's platform). In equilibrium, the policy platforms of the two candidates therefore coincide and maximize the weighted sum of utilities,

$$\int_i \phi^i V^i(a) di = \int_i \phi^i V^i(b) di.$$

Note that the equilibrium platform attaches larger weight to the policy preferences of voters with low variability of ideological attachment (a large value for ϕ^i); voters that care a lot about policy relative to the candidate's other characteristics thus have more political influence. Intuitively, voters whose ideological attachment is unlikely to be biased are more likely to alter their political support in response to small changes in the policy platform. In equilibrium, these groups of *swing voters* thus tilt policy in their favor. If all voters are equally responsive to changes in the policy platform, electoral competition implements the utilitarian optimum with respect to voters.

B.7 Bibliographic Notes

Gorman (1953) and Lewbel (1989) derive conditions under which the summed demand functions of heterogenous households can be represented as the demand of a fictional representative household, see Deaton and Muellbauer (1980, 6) for a discussion. Caselli and Ventura (2000) analyze several dimensions of heterogeneity in a growth model with generalized CIES preferences. Kirman (1992) offers a scathing critique of the "pseudo-microfoundations" (p. 125) of the representative household construct.

Stokey and Lucas (1989, Theorem 4.15) and Acemoglu (2009, Theorem 6.10) discuss the transversality condition in the infinite-horizon planner problem.

Kreps and Porteus (1978) and Epstein and Zin (1989) provide axiomatic foundations for recursive, non-expected utility specifications. Epstein and Zin (1989) and Weil (1989) analyze the asset pricing implications of the model discussed in section B.4, see also Bansal and Yaron (2004).

Sargent and Wallace (1973) discuss the forward solution and Blanchard and Kahn (1980) propose the solution strategy for linear rational expectations models, following Vaughan (1970).

The probabilistic voting model is due to Lindbeck and Weibull (1987).

Related Topics Mas-Colell et al. (1995, 4.B–D, 17.E) cover aggregation, including the Sonnenschein-Mantel-Debreu and Gorman aggregation theorems. Miao (2014, 2) reviews solution strategies for linear rational expectations models. Persson and Tabellini (2000, 3) cover the probabilistic voting model.