Problem Set 0 – Advanced Macroeconomics

Prerequisites: dynamic optimization, business cycles models and consumption-saving problems

Thomas Bourany*
Einaudi Institute for Economics and Finance – EIEF

Due Date: Monday, 14th of September 2026

Description of this homework:

This "Homework 0" should make sure that all of you start the course with the same set of tools and the same common background. It is aimed at recalling the central building blocks of the standard DSGE models[1] of modern macroeconomics: the Neoclassical Growth model (Ramsey–Cass–Koopmans and its stochastic extension by Brock–Mirman), the Real Business Cycle (RBC) model, the New-Keynesian (NK) model, and the consumption-saving problem with idiosyncratic risk (Bewley–Huggett–Aiyagari), which is the workhorse of the heterogeneous-agent literature. These models are covered in standard first-year sequences, and they will be the starting point of this course.

This problem set is organized in three blocks. ***Part 0*** recalls the mathematical toolbox – constrained optimization and the Karush–Kuhn–Tucker (KKT) theorem – that we will use throughout the course. ***Parts 1 to 5*** cover the representative-agent business-cycle models in a deterministic environment: the firm problem and the determination of factor prices, the household problem solved first with the Lagrangian method and then with dynamic programming, the New-Keynesian model, and the interpretation of impulse response functions in the two models. ***Parts 6 to 8*** introduce risk and market structure: stochastic processes for income, the household problem under complete markets in the Arrow–Debreu sense, and the incomplete-market "self-insurance" problem solved recursively. This last block requires probability theory, stochastic processes and dynamic optimization (i.e. control theory) in discrete time, which we will use extensively in the course.

This work will not be graded. No need to work hundreds of hours on it: the main purpose is that all of you start with the same set of common knowledge, and that you can identify by yourself the tools you may want to review. For that reason, more explanations than usual are included in the questions themselves (to get you used to the notations and to the language). The number of points indicated at the beginning of each part is only a rough indication of the time and effort that each of them requires. For additional resources, see any standard textbook, e.g. D. Romer, *Advanced Macroeconomics*, J. Galí, *Monetary Policy, Inflation and the Business Cycle* (for Parts 4–5), or L. Ljungqvist & T. Sargent, *Recursive Macroeconomic Theory* (for Parts 6–8).

Rules for submission:

This problem set should be handed in on <u>Monday, 14th of September 2026</u> before the morning class. The exact date can still be discussed during the first class, if you feel the need to extend the deadline. Once the deadline is set, the solutions will be posted the day after, so no delay can be accepted. Problem sets should be handed in individually. However, it is strongly recommended to work in groups, especially for those of you who did not take the course corresponding courses in the past years. Clarity and conciseness will be appreciated: for most questions, a few lines of derivation and two sentences of economic intuition are enough.

* *thomas.bourany@eief.it*

[1] DSGE stands for *Dynamic Stochastic General Equilibrium*: the economic relations are dynamic (agents are forward-looking) and stochastic (the economy is subject to probabilistic "shocks" and agents form rational expectations). Moreover, the equilibrium is "general", as the economy clears between aggregate demand – from the households – and aggregate supply – from the firms. [For the sake of simplicity, the first parts are derived in a deterministic environment, and risk is introduced only from Part 6 onwards.]

0 Prerequisites (20 pts) Constrained optimization: Karush-Kuhn-Tucker Theorem

0.1 The static case

For an optimization problem, minimizing[2] a function $J(x)$ over a variable[3] x and subject to a set of M constraints $F_1(x) \leq 0 \ldots F_M(x) \leq 0$. Recall the (Karush-)Kuhn-Tucker (KKT) theorem that states necessary and sufficient conditions for optimality.

Solution:

We suppose that the objective function J and the constraints $F_1, \ldots, F_M$ are continuous functions on the set X and derivable on the set K defined by the constraints $K = \{x \in X, \text{ s.t. } F_i(x) \leq 0, \text{ for } 1 \leq i \leq M\}$.
Introducing the "Lagrangian" $\mathcal{L}$ function associated to this problem:

$$\mathcal{L}(x, \lambda_1, \ldots, \lambda_M) = J(x) + \sum_{i=1}^{M} \lambda_i \, F_i(x) \qquad \forall (x, \lambda_i) \in X \times \mathbb{R}_+ \quad \forall 1 \leq i \leq M$$

The optimality condition is to find the solution x^* that reach a saddle point of this Lagrangian function, under the condition that the constraints are "qualified"[4]:

Under all the previous hypothesis, the four following conditions are ***necessary*** for optimality. Formally, if x^* is a global minimum, then the four conditions are satisfied:

1. Stationarity:
$$J'(x^*) + \sum_{i=1}^{M} \lambda_i \, F_i'(x^*) = 0$$
(Equivalent to the "saddle point conditions" on the Lagrangian: $\frac{\partial \mathcal{L}}{\partial x}(x, \lambda) = 0$, $\frac{\partial \mathcal{L}}{\partial \lambda_i}(x, \lambda) = 0 \quad \forall\, 1 \leq i \leq M$)
2. Primal feasibility (simply, constraints should be satisfied): $F_i(x^*) \leq 0 \quad \text{for} \quad 1 \leq i \leq M$
3. Dual feasibility: $\lambda_i \geq 0 \qquad \forall 1 \leq i \leq M$
4. Complementarity: $\sum_{i=1}^{M} \lambda_i \, F_i(x^*) = 0$
(If the constraint is binding at optimum ($F(x^*) = 0$) then the Lagrange multiplier is strictly positive (it stands for the "shadow value" of relaxing the constraint) and conversely)

Under an additional assumption, that the objective function J and the constraints $F_1, \ldots, F_M$ are ***convex***, then these four conditions are also ***sufficient***.
Said differently, if x^* satisfy the four conditions, then x^* is global minimum.

$\cdots$———$\cdots$

[2]Note, this is a minimization problem, minimizing "$J(x)$" is the "same" as maximizing "$-J(x)$". Formally: $\min_x J(x) = -\max_x -J(x)$

[3]Here, for simplicity, we suppose that $x \in \mathbb{R}$ or even $x \in \mathbb{R}^n$. The KKT theorem can be generalized to the case where $x \in X$ where X is a reflexive Banach space (therefore, x can be many different objects, for example a function! (in the case where X is an infinite-dimensional space)).

[4]The constraints are "qualified" when $\forall 1 \leq i \leq M$, *the derivative of the constraint function* $F_i'(u^*)$ *should be negative (or equal to zero if* F_i *are affine). These conditions are sometimes called "Slater condition" in case of convex constraint functions, and "Mangasarian-Fromovitz constraint qualification" in the general case (where there are also equality constraint, which is not the case here). The main idea of qualification (very important for the proof of the "necessary condition" of KKT theorem) is that you can look in the neighborhood of the local minimum to find the optimality condition (after some "linearization" along the lines defined by gradients).*

0.2 The dynamic casein finite horizon

We extend the problem in a dynamic environment, often met in economic models in discrete time. Now the objective function is time-separable and the budget constraints and the variables are time-specific. Therefore, there are an infinite number of constraints[5].

$$\min_{\{x_t\}_{t=0}^{T}} \sum_{t=0}^{\infty} J_t(x_t)$$

$$\text{s.t.} \qquad F_t(x_t, x_{t+1}) \leq 0 \qquad \forall t \geq T$$

Adapt the previous KKT conditions to this new environment.

Solution:

Assuming the regularity conditions stated above (continuity, derivability and qualification of the constraints), and assuming the convexity of objective and constraints functions, we can redefine the Lagrangian and restate the four (necessary and sufficient) optimality conditions:

$$\mathcal{L}(x_t, \lambda_t) = \sum_{t=0}^{T} J_t(x_t) \; + \; \sum_{t=0}^{T} \lambda_t \; F_t(x_t, x_{t+1})$$

1. Stationarity:

$$J_t'(x_t^*) + \lambda_t \, \frac{\partial F_t(x_t^*, x_{t+1})}{\partial x_t} + \lambda_{t-1} \, \frac{\partial F_{t-1}(x_{t-1}, x_t^*)}{\partial x_t} = 0, \qquad \forall \, t \geq 0$$

   (Again equivalent to the "saddle point conditions" on the Lagrangian). Here, this condition should be satisfied at each point of time (Agents are forward-looking and the environment is deterministic here). Moreover, the fact that we end up with only $J_t'(x_t)$ is due to the time-separability of the objective function.

2. Primal feasibility: $F_t(x_t^*, x_{t+1}) \leq 0$ for $t \geq 0$
3. Dual feasibility: $\lambda_t \geq 0 \quad \forall t \geq 0$
4. Complementarity: $\sum_{t=0}^{\infty} \lambda_t \, F_t(x_t^*, x_{t+1}) = 0 \quad \forall t \geq 0$
   (When the constraint is binding at time t, the Lagrange multiplier is strictly positive and stands for the "shadow value" of relaxing the constraint, i.e. the marginal amount you will be able to minimize on the objective function if you could loosen the inequality constraint a little bit))

. . . ——— . . .

[5]Therefore, the Lagrange multipliers are time-specific: λ_t

1 The Firm problem (10 pts) Profit maximization, factor prices, and constant returns to scale

1) Suppose the representative firm should maximize its profit, in an fully-competitive environment, and adjusting its input factor (labor L_t and capital K_t). Profit are defined as output $Y_t = F_t(K_t, L_t)$ discounted for rental cost of capital ($R_t K_t$) and wages ($w_t L_t$). From this *static and unconstrained* problem, derive the First Order Condition, and show how marginal productivity of a factor equals its cost.

2) Suppose now that the technology shows constant return to scale, and in particular a Cobb-Douglas function: $F_t(K_t, L_t) = (K_t)^\alpha (Z_t L_t)^{(1-\alpha)}$.
a) Show how the Cobb-Douglas technology has constant return to scale.
b) Plug the optimality condition in the objective function (the profits), and show that the profits are zero for this representative firm.

Solution:

1. This unconstrained problem takes the form:

$$\begin{aligned}\Pi_t(K_t, L_t) &= F_t(K_t, L_t) - R_t K_t - w_t L_t \\ &\max_{K_t, L_t} [Pi_t]\end{aligned}$$

The first-order condition are the following:

$$\begin{aligned}\frac{\partial \Pi_t(K_t, L_t)}{\partial K_t} &= \frac{\partial F(K_t, L_t)}{\partial K_t} - R_t = 0 \\ \frac{\partial \Pi_t(K_t, L_t)}{\partial L_t} &= \frac{\partial F(K_t, L_t)}{\partial L_t} - w_t = 0 \\ &\Rightarrow R_t = F_k(K_t, L_t) \\ &\Rightarrow w_t = F_\ell(K_t, L_t)\end{aligned}$$

2. Using a Cobb-Douglas function: $F(K_t, L_t) = (K_t)^\alpha (Z_t L_t)^{(1-\alpha)}$, it has constant return to scale:

$$F_t(\theta K_t, \theta L_t) = (\theta K_t)^\alpha (Z_t \theta L_t)^{(1-\alpha)} = \theta^{\alpha+1-\alpha} (K_t)^\alpha (Z_t L_t)^{(1-\alpha)} = \theta\, F_t(K_t, L_t)$$

But, careful!, it has diminishing returns in each variables:

$$\begin{aligned}\frac{\partial F_t(K_t, L_t)}{\partial K_t} &= \alpha \big(\tfrac{Z_t L_t}{K_t}\big)^{1-\alpha} && \text{decreasing in } K_t \\ \frac{\partial F_t(K_t, L_t)}{\partial L_t} &= (1-\alpha) Z_t^{1-\alpha} \left(\tfrac{L_t}{K_t}\right)^{-\alpha} && \text{decreasing in } L_t\end{aligned}$$

Using the FOC optimality conditions above, profits rewrites:

$$\begin{aligned}R_t &= \alpha \big(\tfrac{Z_t L_t}{K_t}\big)^{1-\alpha} \\ w_t &= (1-\alpha) \big(\tfrac{Z_t L_t}{K_t}\big)^{-\alpha} \\ \Pi_t(K_t, L_t) &= (K_t)^\alpha (Z_t L_t)^{(1-\alpha)} - K_t\, \alpha \big(\tfrac{Z_t L_t}{K_t}\big)^{1-\alpha} - L_t\, (1-\alpha) Z_t^{1-\alpha} \left(\tfrac{L_t}{K_t}\right)^{-\alpha} \\ &= (K_t)^\alpha (Z_t L_t)^{(1-\alpha)} - (\alpha + 1 - \alpha)(K_t)^\alpha (Z_t L_t)^{(1-\alpha)} \\ &= 0\end{aligned}$$

This is a consequence of the assumption that technology has constant returns to scale.

$\cdots$ ——— $\cdots$

2 The Household problem (15 pts) Lagrangian method

Now, let us consider the Household problem:

$$\max_{\{C_t\}_0^\infty, \{A_{t+1}\}_0^\infty, \{L_t\}_0^\infty} \sum_{t=0}^{\infty} \beta^t U(C_t, L_t)$$

$$\text{s.t.} \qquad C_t + A_{t+1} \leq (1+r_t)A_t + w_t L_t$$

A representative consumer receives income $w_t L_t$, which is time-varying (but that is deterministic, i.e. whose fluctuations can be perfectly anticipated in advance), consumes a good C_t, and save using an asset A. In period t, the consumer owns a quantity A_t of savings (when $A_t \leq 0$ this saving is in fact a debt) and decides how much to save for new period A_{t+1}.

Note that we could intuitively relabel this budget constraints as $C_t + Q_t \tilde{A}_{t+1} = \tilde{A}_t + w_t L_t$, with the one-period asset $\tilde{A}_t = (1+r_t)A_t$ is bought at price $Q_t = \frac{1}{1+r_{t+1}}$, and "returns" 1 next period. This convention will be used in later sections.

(a.) Apply the KKT theorem to this dynamic setting (i.e. use the first exercise/second case) to solve the Household problem. (Applying a theorem means checking explicitly the different hypothesis). The HH maximizes over its consumption path and labor (-supply) path.You should obtain (among other things) two[6] type of optimality conditions:
(i) an Euler equation showing relation between present and future marginal utility of consumption
(ii) a Labor-Consumption tradeoff, with a ratio between marginal utility of consumption vs. marginal disutility of labor.
(b.) Apply these conditions to the special functional form of utility: separable between labor and consumption, CRRA in consumption and with a constant Frisch-elasticity: $U(C_t, L_t) = \frac{C_t^{1-\sigma}}{1-\sigma} - \frac{L_t^{1+\varphi}}{1+\varphi}$, where σ is the risk aversion coefficient – which is also the *inverse* of intertemporal elasticity of substitution (IES), and φ is the Inverse of the Frish elasticity of labor supply.
- Explain these relations with economic intuitions.
- How the consumption path changes with r_t? How do consumption and saving react when r_t changes, in the case where IES is low ($\sigma > 1$) vs. when it is high ($\sigma < 1$).
- Lastly, when expressing $Q_t = \frac{1}{1+r_{t+1}}$ as function of the other variables, how is the asset priced (what is Q_t), if $C_{t+1} \gg C_t$.

———

Solution:
We make the hypothesis that the utility function is continuous and derivable (expect when consumption is null, which is unlikely to be a optimum) and concave in consumption and labor. The constraint rewrites: $F_t(c, \ell, a_t, a_{t+1}) = C_t + A_{t+1} - (1+r_t)A_t - w_t L_t \leq 0$. Note that A_{t+1} appears twice: once in F_t and once in F_{t+1} (recall the budget is a constraint at all time t). Therefore, the optimality of the path $(C_t^*, L_t^*)_{t=0}^\infty$ is equivalent to the following conditions:

[6] Here, as there are 3 variables, there should be 3 types of optimality conditions, but here you should get rid of the Lagrange multipliers to determine explicitly the consumption path

1. Stationarity (C_t, A_{t+1} and L_t):

$$
\begin{array}{lll}
[C_t] & \beta^t \dfrac{\partial U(C_t^*, L_t^*)}{\partial C_t} & = \lambda_t\,(1), \qquad \forall\, t \geq 0 \\
[A_{t+1}] & \{\dfrac{\partial U(C_t^*, L_t^*)}{\partial A_{t+1}} = 0\} & 0 = \lambda_t\,(1) + \lambda_{t+1}\{-(1+r_{t+1})\} \\
\text{(combining 1 \& 2)} & \dfrac{\partial U(C_t^*, L_t^*)}{\partial C_t} & = \beta \dfrac{\partial U(C_{t+1}^*, L_{t+1}^*)}{\partial C_{t+1}}(1+r_{t+1}) \\
[L_t] & \beta^t \dfrac{\partial U(C_t^*, L_t^*)}{\partial L_t} & = -\lambda_t\, w_t \\
\text{(combining 1 \& 3)} & \dfrac{\partial U(C_t^*, L_t^*)}{\partial L_t} & = -\dfrac{\partial U(C_t^*, L_t^*)}{\partial C_t}\, w_t
\end{array}
$$

Rewriting the previous equations more compactly:

$$
\begin{array}{rl}
U_c(C_{t+1})\beta(1+r_{t+1}) = U_c(C_t) & [\textbf{\textit{Euler equation}}] \\
-\dfrac{U_\ell(C_t, L_t)}{U_c(C_t, L_t)} = w_t & [\textbf{\textit{Labor consumption tradeoff}}]
\end{array}
$$

2. Primal feasibility: $C_t + A_{t+1} \leq (1+r_t)A_t + w_t L_t \quad \text{for} \quad t \geq 0$
3. Dual feasibility: $\lambda_t \geq 0 \qquad \forall t \geq 0$
4. Complementarity: $\sum_{t=0}^{\infty} \lambda_t \left(C_t + A_{t+1} - (1+r_t)A_t + w_t L_t\right) = 0 \quad \forall t \geq 0$
   (the more the budget constraint binds, the more the Lagrange multiplier increases)

With the given functional form, the two relations rewrite:

$$
\beta(1+r_{t+1}) = \Big(\frac{C_{t+1}}{C_t}\Big)^{\sigma} \qquad\qquad C_t^{\sigma}\, L_t^{\varphi} = w_t
$$

or, equivalently, expressing the intertemporal condition in terms of the price $Q_t = \frac{1}{1+r_{t+1}}$ of the one-period asset:

$$
Q_t \equiv \frac{1}{1+r_{t+1}} = \beta\Big(\frac{C_t}{C_{t+1}}\Big)^{\sigma}
$$
$$
\text{or} \quad \frac{C_{t+1}}{C_t} = \Big(\frac{\beta}{Q_t}\Big)^{-1/\sigma}
$$

(b.) Let us provide intuitions.

The Euler-equation[7] represent an intertemporal condition. Everything else held constant, today's consumption C_t increases if future consumption C_{t+1} increases, if the future interest rate r_{t+1} decreases – equivalently, if the asset price Q_t rises – or if the agent is less patient (smaller β).

Obviously, holding everything else constant is somehow counterintuitive: the Euler equation characterizes a "consumption smoothing" behavior, balancing substitution effects vs. income effects. When asset price Q_t falls / interest rate r_{t+1} rises, the substitution effect implies that saving is more profitable and the agents save more for future consumption. When asset price Q_t falls / interest rate r_{t+1} rises, the income effect implies that you will get the same (ex-post) saving amount A_{t+1} at a lower cost (i.e. a smaller initial amount), therefore agents keep more money for present consumption and gross saving decreases. When $\sigma < 1$, i.e. with high intertemporal elasticity of substitution (IES), the substitution effect is stronger (consistent with the name :p). When $\sigma > 1$, i.e. low IES, the income effects dominates.

This equation is also a pricing equation Q_t is the value agents give to the asset that allow them to transfer wealth from t to $t+1$. Hence if $C_{t+1} \geq \beta^{1/\sigma} C_t$ the value of the asset today is worth less

[7] Which has not much to do with the swiss mathematician Euler ... only by the fact that Euler worked on optimality conditions in variational problems

than 1 (i.e. interest rate is positive). The asset price drops (interest rises) further the larger the gap between future and present consumption, and the lower the IES is (or higher σ is). Again, low IES implies higher value for transferring money.

The labor consumption tradeoff is a intra-temporal condition: everything else held constant, if labor increases (e.g. due to an exogenous rise in labor demand), then the consumption shall drop. In particular, this relation explains why the RBC model can difficultly generate positive/large government spending multipliers (mainly because higher government driven aggregate demand does not have any feedback effect on consumption behavior). Concerning real wages, the leisure-consumption trade-off can also be seen as a balance between substitution effect (if wage rise, and if leisure and consumption are substitute, people will exchange leisure against "cheaper" consumption) and income effect (need to work less to get the same consumption basket). RBC do not have the best foundations for labor supply, and that is why many recent papers include "search and matching" features in macro models.

In the previous Euler equation, when risk is introduced, the RHS is often referred as the stochastic discount factors (more on that in Sections 7 and 8) or pricing kernel in asset-pricing finance.

...——— ...

3 The Household problem, (15 pts) Dynamic programming

Now, let us consider the Household problem:

$$\max_{\{C_t\}_0^\infty, \{L_t\}_0^\infty} \sum_{t=0}^{\infty} \beta^t U(C_t, L_t)$$
$$\text{s.t.} \qquad C_t + A_{t+1} \leq (1+r_t)A_t + w_t L_t$$

This is exactly the same problem as in the previous section, but we now solve it ***recursively***, using the "Dynamic programming" method. Again, the HH maximizes over the consumption path and labor (supply) path, and you should obtain the very same two types of optimality conditions:
(a) an Euler equation and (b) a Labor-Consumption tradeoff. The point of the exercise is the method, not the result: the two approaches must agree, and the recursive one is the only one that will still work once we introduce risk (in Section 8).

The idea here is to solve[8] the "Bellman" equation: the utility function can be rewritten as an iterative "Value function":

$$V_t(A_t) = \max_{\{C_t\}_t, \{A_{t+1}\}_t, \{L_t\}_t} \Big[\sum_{j=t}^{\infty} \beta^{j-t} U(C_j, L_j) \Big]$$
$$\Leftrightarrow V_t(A_t) = \max_{C_t, A_{t+1}, L_t} \Big[U(C_t, L_t) + \beta V_{t+1}(A_{t+1}) \Big]$$

To solve this problem, you need to derive (i) the KKT FOC-conditions for this new (iterative) maximization problem[9], and (ii) a new "Envelope condition". This second condition shows the derivative of the value function (w.r.t. A_t) and, through the use of the chain-rule, is linked to the "Envelope Theorem". In this theorem[10], when a function is maximized over a variable (V_t is max over A_{t+1} here) and hence optimal in this sense, its derivative w.r.t. this same variable A_{t+1} is null.
Combining the different equations and manipulating the terms, you should obtain: (a) the Euler equation and (b) the Labor-consumption trade-off (exactly as before!).

Apply this conditions to the special functional form of utility: separable between labor and consumption, CRRA in consumption and with a constant Frisch-elasticity: $U(C_t, L_t) = \frac{C_t^{1-\sigma}}{1-\sigma} - \frac{L_t^{1+\varphi}}{1+\varphi}$

[8] During the course, we will discuss the general method to solve such problems (and prove that the sequential problem and its recursive formulation are equivalent).
[9] Note that we are back to the first "static" version of the KKT theorem (exo 0.1)
[10] We will come back to this theorem during the course as well

Solution:

As a preliminary exercise, let us recall the **Envelope Theorem**. This theorem arises when one want to differentiate a function that is expressed as an optimization problem. The most important case in economics (micro and macro) is when a value function depends on a "state-variable" x but is the optimum over a "control" variable y:

$$v(x) = \max_{y \in K} f(x, y) \qquad \forall x \in \mathbb{R}^d$$

We develop the multivariate case here, where: K is a compact set (eventually the set resulting from a set of constraints) $K \subset \mathbb{R}^m$ where m is the dimension of the control (y) and d the dimension of the state x. We assume that, for every value of $y \in K$, $f(., y)$ is differentiable, and $\nabla_x f(x, y)$ is continuous.

We assume, after using the KKT theorem above if the case is appropriate, that this maximization problem has an unique optimum. This optimal y^* depends on the state variable: $y^*(x)$, implicitly derived from the KKT optimality conditions (and that can be analytically expressed using the ... implicit theorem!). In macro, the result of the maximization problem is often called the *policy function*: $y^*(x) = \pi(x)$.

Given all these assumptions, the value function v is of class C^1 and we can explicitly obtain its derivative:
$$\nabla v(x) = \nabla_x f(x, y^*(x)), \qquad \forall x \in \mathbb{R}^d$$

Important remark: the main thing here is that the derivative v is the same as the derivative of f w.r.t. its first variable, and $\nabla_y f(x, y^*(x)) = 0$. Why? Because the function $f(., y)$ is already optimal w.r.t. y and implied by the (KKT-) FOC conditions. Therefore, the chain rule implies ($d = m = 1$): $v'(x) = f'_x(x, y(x)) + f'_y(x, y(x))\, y'(x) = f'_x(x, y(x))$

Alright? Ok, let's come back to our economic problem and our **Bellman equation**:

$$V_t(A_t) = \max_{C_t, A_{t+1}, L_t} \Big[U(C_t, L_t) + \beta V_{t+1}(A_{t+1})\Big]$$

Here, the value function depends on the state A_t, i.e. your saving, (in heterogeneous agents models, it will be the wealth), while it is optimal w.r.t. the controls C_t, A_{t+1}, L_t.

Therefore, we need to derive the **FOC conditions** of the problem at time t to obtain the policy functions ($C_t = \pi(A_t)$ and $L_t = \tilde{\pi}(A_t)$ both determined using the ... KKT theorem). Rewrite the inequality constraint and Lagrangian:

$F_t(c, \ell, a_{t+1}) = C_t + A_{t+1} - (1+r_t)A_t - w_t L_t \leq 0 \qquad \mathcal{L}_t(c, \ell, a_{t+1}, \lambda) = U(c, \ell) + \beta V_{t+1}(A_{t+1}) - \lambda_t\, F_t(c, \ell, a_{t+1})$

The FOC yields the following equations (note that we won't rewrite the primal and dual feasibility and complementarity conditions that are exactly the same as above) deriving w.r.t. C_t, A_{t+1} and L_t:

$$
\begin{array}{lll}
[C_t] & & \dfrac{\partial U(C_t^*, L_t^*)}{\partial C_t} = \lambda_t \ (1) \\
[A_{t+1}] & \{\dfrac{\partial U(c,\ell)}{\partial A_{t+1}} = 0\} & \beta \dfrac{dV_{t+1}(A_{t+1}^*)}{dA_{t+1}} = \lambda_t \ (1) \\
[L_t] & & \dfrac{\partial U(C_t^*, L_t^*)}{\partial L_t} = -\lambda_t \ w_t \\
\text{(combining 1 \& 3)} & & U_\ell(c^*, \ell^*) = -U_c(c^*, \ell^*) \ w_t \\
\text{(combining 1 \& 2)} & & U_c(c^*, \ell^*) = \beta V_{t+1}'(A_{t+1}^*)
\end{array}
$$

Here, the fourth equation is the ***Labor-consumption tradeoff*** (intratemporal FOC)! Now, to derive the **envelope condition**, we need to obtain the derivative of V_t (w.r.t. A_t) (why? In the fifth equation we still have the term $V_{t+1}'(A_{t+1})$ unknown). Coming back to the Bellman equation, let us obtain $V'(A_t)$. Here, we can "simply" use the Envelope theorem: we need to derive the value function w.r.t. the state variable, while it is already maximized over the three controls. ***But***, there is a trick, in the sense that, when the budget constraint is binding, one of the control can be expressed as a function of the other (state and control) variables:

$$C_t = \phi(A_t, A_{t+1}, L_t) = -A_{t+1} + (1 + r_t)A_t + w_t L_t$$

This implies that, after using the chain-rule, the envelope condition is expressed as:

$$V_t'(A_t) = \frac{\partial U(C_t, L_t)}{\partial C_t} \frac{\partial \phi(A_t, A_{t+1}, L_t)}{\partial A_t} + \underbrace{\Big[\frac{\partial U(C_t, L_t)}{\partial C_t} \frac{\partial \phi(A_t, A_{t+1}, L_t)}{\partial A_{t+1}} + \beta \frac{dV_{t+1}(A_{t+1})}{dA_{t+1}}\Big]}_{=0 \ \ \text{(FOC on } A_{t+1})} \frac{dA_{t+1}}{dA_t} + \underbrace{\Big[\frac{\partial U(C_t, L_t)}{\partial C_t} \frac{\partial \phi(A_t, A_{t+}}{\partial L_t}}_{=0 \ \ \text{(FOC}}$$

And ***here*** we can use the Envelope theorem: the two brackets are exactly the first-order conditions with respect to the controls A_{t+1} and L_t, hence they are zero at the optimum, as $\beta V_{t+1}' = U_c$. Therefore:

$$
\begin{aligned}
V_t'(A_t) &= \frac{\partial U(C_t, L_t)}{\partial C_t} \frac{\partial \phi(A_t, A_{t+1}, L_t)}{\partial A_t} \\
&= U_c(C_t, L_t)\,(1 + r_t)
\end{aligned}
$$

This condition holds at all time, and we can iterate ($t \to t+1$), and obtain:

$$V_{t+1}'(A_{t+1}) = U_c(C_{t+1}, L_{t+1})\,(1 + r_{t+1})$$

and replacing in the fifth equation above :

$$U_c(C_t, L_t) = \beta V_{t+1}'(A_{t+1}) = \beta U_c(C_{t+1}, L_{t+1})\,(1 + r_{t+1})$$

We got the ***Euler equation*** (intertemporal FOC)!

Again, we can recompute the two equations with a specific functional form:

$$\Big(\frac{C_{t+1}}{C_t}\Big)^\sigma = \beta(1 + r_{t+1}) \qquad\qquad C_t^\sigma \, L_t^\varphi = w_t$$

With the dynamic programming methods, that seems a lot of work for two simple (and already known) equations ... but the good news is that the method is completely general: we will re-use it in Section 8, where risk makes the sequential Lagrangian approach much less convenient.

$\cdots$ ——— $\cdots$

4 The New-Keynesian model, (10 pts) Nominal rigidities, IS-curve, Phillips curve and monetary policy

1. [Optional question] To solve the New Keynesian model, one of the step is to solve the following Household problem:

$$\max_{\{C_t\}_0^\infty, \{B_{t+1}\}_0^\infty, \{L_t\}_0^\infty} \sum_{t=0}^{\infty} \beta^t U(C_t, L_t)$$
$$\text{s.t.} \qquad C_t P_t + Q_t B_{t+1} \leq B_t + W_t L_t + T_t$$

This problem is really similar to the one of the RBC problem (again everything is deterministic for the sake of simplicity). Using the given functional form $U(C_t, L_t) = \frac{C_t^{1-\sigma}}{1-\sigma} - \frac{L_t^{1+\varphi}}{1+\varphi}$, derive the Euler-equation and labor-consumption tradeoff, using the method of your choice (the Lagrangian method or the Dynamic programming method)[11].

Briefly explain what are the changes in the Euler equation and Labor-consumption trade-off, for instance when inflation changes the price level.

2. ***[Compulsory question]*** Recall the log-linearized formulas[12] for (a) the dynamic IS-equation (linking the present and future output gaps and interest rate), (b) the New-Keynesian Phillips Curve (linking present and future inflation rates and aggregate demand) and (c) the interest rate rule[13] (showing how evolves the monetary policy). Describe these relations and their economic intuitions.

For the monetary policy rule, briefly explain the Taylor principle.

———

Solution:

1. In the New Keynesian models, the inter- and intra-temporal First Order Conditions rewrites:

$$Q_t = \beta\Big(\frac{P_t}{P_{t+1}}\Big)\Big(\frac{C_t}{C_{t+1}}\Big)^{\sigma} \qquad\qquad C_t^{\sigma} L_t^{\varphi} = \frac{W_t}{P_t}$$

Now, in the Euler equation, a wedge appears due to price level distortion: when prices are expected to increase in $t+1$ (again here there should be an expectation "$\mathbb{E}_t$" sign), then – everything else held constant – consumption (C_t) should increases: consumption smoothing is weighted given price level changes. Moreover, if wages decreases but if prices take time to adjust, then real wage fall and Household decrease their consumption and/or labor supply.

[11] This exercise is advised for those who feel the need to practice

[12] Simply states the three equations, no need to derive it! The main textbook for this type of model is: J. Gali - Monetary Policy, Inflation and the Business cycle

[13] There are different types of monetary policy rule, choose one and explain it

2. Noting that $i_t = -log(Q_t)$, the Euler equation described above yields, through its log-linearized version, the ***Dynamic IS equation***[14]:

$$\hat{c_t} = \mathbb{E}_t(\hat{c}_{t+1}) - \frac{1}{\sigma}(i_t - \mathbb{E}_t(\pi_{t+1}) - \rho) \qquad \text{Euler equation}$$

$$\widetilde{y_t} = \mathbb{E}_t(\widetilde{y}_{t+1}) - \frac{1}{\sigma}(i_t - \mathbb{E}_t(\pi_{t+1}) - r_t^n) \qquad \text{IS equation}$$

$$\text{with } r_t^n = \rho + \sigma\psi_{yz}^n\,\mathbb{E}_t(\Delta z_{t+1}) \qquad \text{Natural int. rate}$$

Behind the IS equation lies the usual mechanisms underlying the standard Euler equations. However, the impact of monetary policy is clearer (it is a monetary model after all): when real interest rate rises above its natural level, this has a negative impact on consumption (as agents save instead of consuming), and output gap widens. Again, due to consumption smoothing motive, future consumption impacts linearly today's consumption: expectations of agents over future policy have therefore a critical impact (and this situation even creates what we call the "forward guidance" puzzle: in this type of model, future policy decision have oversized effects).

The ***New-Keynesian Phillips Curve*** is derived from the definition of marginal cost (directly coming from the labor-consumption tradeoff) and the price setting of firms:

$$\hat{w}_t - p_t = \sigma\hat{c}_t + \varphi\hat{\ell}_t \qquad \text{labor-cons. tradeoff}$$

$$mc_t = (\hat{w}_t - p_t) - mpl_t \qquad \text{effective marginal cost}$$

$$\widehat{mc}_t = \left(\sigma + \tfrac{\varphi+\alpha}{1-\alpha}\right)(y_t - y_t^n) = \psi_{mc}\,\widetilde{y}_t \qquad \text{"marginal cost" gap}$$

$$\pi_t = \beta\,\mathbb{E}_t(\pi_{t+1}) + \lambda\widehat{mc}_t \qquad \text{firms price setting}$$

$$\pi_t = \beta\,\mathbb{E}_t(\pi_{t+1}) + \kappa\,\widetilde{y}_t \qquad \text{New Keynesian Philipps Curve}$$

Behind the NKPC, the firms pricing decision have an impact on inflation: when aggregate demand increases, raising output above its flexible price level, it increases the labor disutility (households dislike working more than steady state level) and, through workers-firms bargaining, wages increases and so do marginal costs. However, only a fraction $1-\theta$ of firms can reset their price in a given period (Calvo pricing) – the coefficient λ of the NKPC is not that fraction, but the slope $\lambda = \frac{(1-\theta)(1-\beta\theta)}{\theta}\Theta$, which decreases with the degree of price stickiness θ. Moreover, when firms expect future inflation (and future rise in demand/marg. cost), they know that some of them will not be able to update their prices in the future (when the shocks arise) and prefer rising their prices now: this creates an expectation-driven "inflationary spiral".

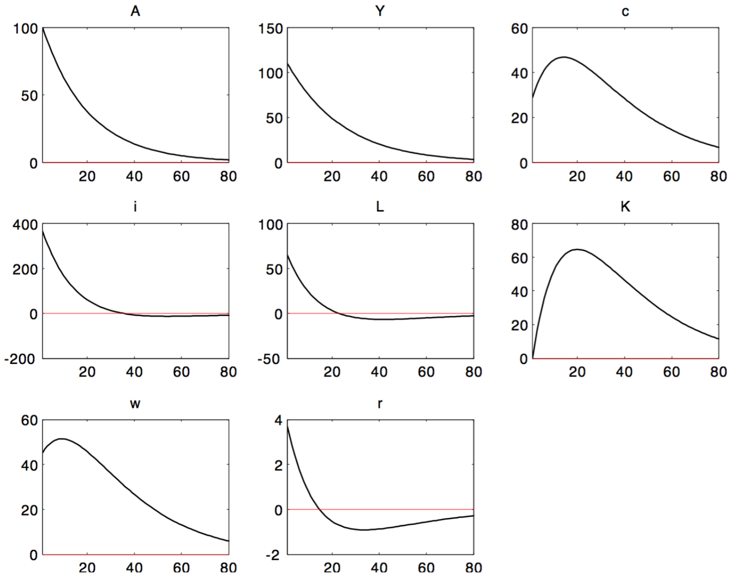
... ——— ...

[14]This time I provide the stochastic version, with expectation signs

5 The dynamics of the two models, (12 pts) Impulse response functions (IRF) to technology shock

This is the response of the two standard models to a Technology shock.[15]

Figure 1: *Baseline RBC model*



In the first RBC model, the variables are the following: Z=TFP (labelled "A"), Y=Output, C=consumption, i=Investment, L=Labor, K=Capital, w=Wages, r=interest rate. The relations structuring the models are the Euler equations, the Labor-consumption trade-off, the definition of output (technology: Cobb-Douglas, and $Y = C + i$), the evolution of capital ($K_{t+1} = (1 - \delta)K_t + i$), the factor prices, wages and interest rate, as marginal productivity of factors, and an AR(1) evolution of TFP.

In the second NK model, the variables are described (with z=TFP, labelled "a"), and the relations structuring the models are the IS equation (or Euler equation), the NK-Phillips Curve, the monetary policy rule for the nominal interest rate, the definitions of: the natural rate of output and natural rate of interest (rates that would prevail in the flexible price economy), the output gap (effective output − natural output), the real rate as interest rate discounted for inflation, and employment.

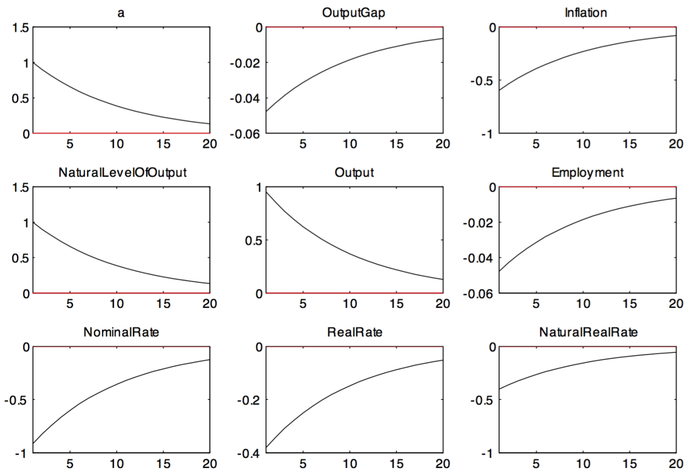
Describe the economic mechanisms underlying the two sets of Impulse Response function (IRF) following a technology shock, first in the Real-Business Cycle model, and second in the New-Keynesian model.

Solution:

In the RBC model, when technology momentarily improves above trend, the marginal product of both capital and labor increases, which leads to higher wages and higher interest rates. Agents respond

[15] *Notation in the figures.* The plots were produced in Dynare/Matlab, with notations differing slightly from the ones used previously. In the *RBC* figure: "A" is the TFP Z_t, "Y" is output Y_t, "c" is consumption C_t, "i" is investment I_t, "L" is labor L_t, "K" is capital K_t, "w" is the real wage w_t and "r" is the interest rate r_t. In the *NK* figure: "a" is the (log) TFP z_t, "NaturalLevelOfOutput" is y_t^n, "Output" is y_t, "OutputGap" is the output gap $\widetilde{y}_t = y_t - y_t^n$, "Inflation" is π_t, "Employment" is $\hat{\ell}_t$, "NominalRate" is i_t, "RealRate" is the ex-ante real rate $i_t - \mathbb{E}_t(\pi_{t+1})$ and "NaturalRealRate" is r_t^n. In both figures, all variables are plotted as deviations from their steady-state value, and the horizontal axis is the number of periods after the shock.

Figure 2: *Baseline NK model*



to the higher wage rate by increasing their labor supply. Output has to increase. Since agents want to have a smooth consumption pattern, aggregate consumption initially increases (proportionally) by less than output and, hence, investment increases (proportionally) by more than output. Capital slowly accumulates over time. The accumulation of capital lowers its marginal product and leads to an interest rate below its steady state level. Once technology is almost back to trend, there remains a wealth effect due to the higher capital stock. Agents therefore consume more while working slightly less than in steady state.

In the NK model, the favorable TFP shock increases the natural level of output (that would prevail under flexible prices), raises the natural interest rate (which is simply the discount rate adjusted for the evolution of technology) decreases firms marginal costs. However, prices are sticky and not all firms can adjust their prices (i.e. some firms would like to decrease them but can not because of Calvo-style staggered price setting). This relative price distortion reduces the output of those firms, the household consumption (due to monopolistic competition and "taste for diversity"), and therefore employment: output gap widens. Notice that the improvement in technology is partly accommodated by the monetary policy, easing nominal interest rate. Nevertheless, this expansionary policy is not sufficient to close the negative output gap, yielding deflation.

...——— ...

6 Stochastic income processes, (30 pts) Random walk, Markov chain, Martingale, AR(1), e tutti quanti

The first half of this problem set was written in *general equilibrium* and in a *deterministic* environment: households and firms interacted, and Y_t denoted aggregate output. In this second half, we zoom in on a single household in *partial equilibrium* – prices and the income process are taken as given – and we introduce *risk*. Accordingly, the notation shifts: from now on Y_t denotes the household's exogenous *income* ($\sim$ labor income $w_t L_t$, now random, and labor supply is no longer chosen by agents). This is the notation often used in the consumption-saving literature, the sections 7 and 8, and in the heterogeneous-agent literature more broadly.

We now introduce a simple shock for income $Y_t = Y(s_t)$. It fluctuate randomly given a stochastic process, where s_t is a "state of the world" (or alea) at date t that belongs to the information set $\mathcal{I}_t$ (what is known to be possible to happen at date t)[16]. More specifically, a sequence of state-of-the-world $\{s_0, s_1, \dots, s_{t-1}, s_t\} = s^t$ defines a "history" s^t of shocks until t (this is the economist way of writing things compactly).

Putting some structure on this process $Y(s^t)$ (making it depends on past), we consider an Autoregressive – AR(1) – process for income:

$$\begin{cases} Y_t(s^t) & = \rho_Y Y_{t-1}(s^{t-1}) + \varepsilon(s_t) \\ Y_{t_0} & = y \end{cases}$$

where $\rho_Y \in \mathbb{R}$ is a parameter and ε is a i.i.d. random variable (i.e. a function of the state of the world), that simply[17] takes three values: $+1$ with proba p, -1 with proba q and 0 with proba $1 - p - q$. The following questions (b-f) are independent from each other (but most of them need the answer of (a.).)

(a.) Compute the conditional expectation $\mathbb{E}(Y_t(s^t)|\mathcal{I}_{t-1})$, as a function of the parameters and note that this expectation is a function of a variable depending on s^{t-1}.
- Using the same logic, write the value $\mathbb{P}(Y_t(s^t) = z|\mathcal{I}_{t-1})$ (Hint: you may need to use indicator functions $\mathbb{1}\{W = w\}$ for a random variables W and values w to be determined).

Solution:
The conditional expectation of $Y_t(s^t)$ given the information $\mathcal{I}_{t-1}$ is :

$$\begin{aligned} \mathbb{E}(Y_t(s^t)|\mathcal{I}_{t-1}) &= \mathbb{E}(\rho_Y Y_{t-1}(s^{t-1}) + \varepsilon(s_t)|\mathcal{I}_{t-1}) \\ &= \rho_Y Y_{t-1}(s^{t-1}) + \mathbb{E}(\varepsilon(s_t)|\mathcal{I}_{t-1}) \\ &= \rho_Y Y_{t-1}(s^{t-1}) + \mathbb{E}(\varepsilon(s_t)) \\ &= \rho_Y Y_{t-1}(s^{t-1}) + p - q \end{aligned}$$

We pass from the first to second line because $Y_{t-1}(s^{t-1})$ is known at time $t-1$ (we already have information about its value at $\mathcal{I}_{t-1}$, i.e. $Y_{t-1}(s^{t-1})$ is measurable w.r.t. to $\mathcal{I}_{t-1}$ in measure theory language).
We get the third line because $\varepsilon(s_t)$ is independent of past shocks: information at $t+1$ doesn't help predict the expectation of $\varepsilon(s_t)$, i.e. $\varepsilon(s_t)$ is not measurable w.r.t. $\mathcal{I}_{t-1}$.

[16] In measure theory language, this information set is the $\sigma-$algebra generated by the stochastic process until t (and $\{\mathcal{I}_t\}_t$ is a filtration: $\mathcal{I}_{t-1} \subset \mathcal{I}_t \subset \mathcal{I}_{t+1}$), i.e. it implies that $Y(s_t)$ is measurable with respect to this $\sigma-$algebra: it doesn't include values that are "outside" this set (if measure theory is too abstract for you, don't worry and just ignore this last sentence/footnote!)

[17] If you find that too easy, feel free to consider that as a random variable $\varepsilon(s_t) \sim \mathcal{N}(\mu_Y, \sigma_Y)$

Last line is the definition of the law of $\varepsilon(s_t)$ (it is $\mathbb{E}(\varepsilon(s_t)) = \mu_Y$ if you assume $\varepsilon(s_t) \sim \mathcal{N}(\mu_Y, \sigma_Y)$). We see that this conditional expectation $\mathbb{E}(Y_t(s^t)|\mathcal{I}_{t-1})$ is a *function* of period s^{t-1} information $\mathcal{I}_{t-1}$.

Similarly, the conditional probability can also be computed with the same logic. We note that because $\mathbb{P}(Z = z|\mathcal{F}) = \mathbb{E}[\mathbb{1}\{Z = z\}|\mathcal{F}]$, the approach follows :

$$\begin{aligned}
\mathbb{P}(Y_t(s^t) = z|\mathcal{I}_{t-1}) &= \mathbb{P}(\rho_Y Y_{t-1}(s^{t-1}) + \varepsilon(s_t) = z|\mathcal{I}_{t-1}) \\
&= \mathbb{P}\big(\varepsilon(s_t) = 1 \ \&\ \rho_Y Y_{t-1}(s^{t-1}) = z - 1\big|\mathcal{I}_{t-1}\big) + \mathbb{P}\big(\varepsilon(s_t) = 0 \ \&\ \rho_Y Y_{t-1}(s^{t-1}) = z\big|\mathcal{I}_{t-1}\big) \\
&\qquad + \mathbb{P}\big(\varepsilon(s_t) = -1 \ \&\ \rho_Y Y_{t-1}(s^{t-1}) = z + 1\big|\mathcal{I}_{t-1}\big) \\
&= \mathbb{P}(\varepsilon(s_t) = 1) \times \mathbb{P}\big(\rho_Y Y_{t-1}(s^{t-1}) = z - 1\big|\mathcal{I}_{t-1}\big) + \mathbb{P}(\varepsilon(s_t) = 0) \times \mathbb{P}\big(\rho_Y Y_{t-1}(s^{t-1}) = z\big|\mathcal{I}_{t-1}\big) \\
&\qquad + \mathbb{P}(\varepsilon(s_t) = -1) \times \mathbb{P}\big(\rho_Y Y_{t-1}(s^{t-1}) = z + 1\big|\mathcal{I}_{t-1}\big) \\
&= p\mathbb{1}\{Y_{t-1}(s^{t-1}) = \tfrac{z-1}{\rho_Y}\} + (1 - p - q)\mathbb{1}\{Y_{t-1}(s^{t-1}) = \tfrac{z}{\rho_Y}\} + q\mathbb{1}\{Y_{t-1}(s^{t-1}) = \tfrac{z+1}{\rho_Y}\}
\end{aligned}$$

We get the third line – i.e. we could separate the probability sign because the events are disjoint[18] To get the four line, we use the last that the events inside the proba sign are independents[19] (as ε is i.i.d. and hence independent of Y_{t-1}). The last line is obtained because again information on $Y_{t-1}(s^{t-1})$ is included in $\mathcal{I}_{t-1}$ and we don't need to take expectation (because we know its value at $t - 1$). Again, we realize that this conditional probability is a function of $Y_{t-1}(s^{t-1})$.

$\cdots$———$\cdots$

(b.) Show what are the conditions on the parameters (ρ_Y, p and q, etc.[20]), for Y_t to be a martingale (or a sub-martingale, or a super-martingale).

———

Solution:

Let us first verify that the sequence has finite first moment : $\mathbb{E}\big[|Y_t|\big] = \mathbb{E}\big[|Y_{t-1} + \varepsilon|\big] \leq \mathbb{E}\big[|Y_{t-1}|\big] + \overbrace{\mathbb{E}\big[|\varepsilon|\big]}^{\leq 1}$ (triangle inequality). Hence if $\mathbb{E}\big[|Y_{t-1}|\big] < \infty$ then $\mathbb{E}\big[|Y_t|\big] < \infty$.
We can thus say (by recursion) that if $\mathbb{E}\big[|Y_{t_0}|\big] = \mathbb{E}\big[|y|\big] = |y| < \infty$, then $\forall\, t \geq 1$, $\mathbb{E}\big[|Y_t|\big]$ has a finite first moment.

Now let us check the second condition. We computed the conditional expectation in question a):

$$\mathbb{E}(Y_t(s^t)|\mathcal{I}_{t-1}) = \rho_Y Y_{t-1}(s^{t-1}) + p - q$$

It is obvious that Y_t is a martingale when $\rho_Y = 1$ and $p = q$, s.t. $\mathbb{E}(Y_t|\mathcal{I}_{t-1}) = Y_{t-1}$. This is called a random walk (more on this in question d. and g.). Still with $\rho_Y = 1$, the process is a submartingale (it increases on average) if $p \geq q$, and a supermartingale (it decreases on average) if $p \leq q$. ***Careful*** when $\rho_Y \neq 1$: the supermartingale condition $\mathbb{E}(Y_t|\mathcal{I}_{t-1}) \leq Y_{t-1}$ reads $(p - q) \leq (1 - \rho_Y)Y_{t-1}$, which depends on the sign and the size of Y_{t-1} and hence cannot hold for every realization. A mean-reverting AR(1) is therefore neither a super- nor a sub-martingale globally. In these two cases, on average the process falls or grows forever, but it doesn't prevent the process to have finite moments: $\mathbb{E}\big[|Y_t|\big] < \infty, \forall\, t$ isn't the same as $\mathbb{E}\big[\sup_{t\geq 1} |Y_t|\big] < \infty$.

$\cdots$———$\cdots$

(c.) Does this stochastic process respect the Markov property? Can we say that it is a Markov chain (i.e. Markov process)? What is the set of states? (Hint: this needs not be finite). Write the transition matrix as function of the parameters, for $\rho_Y = 1$.

[18] The events A and B are disjoint when $\mathbb{P}(A \cap B) = 0$ and we used $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B)$.

[19] The events A and B are independents, then $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B)$.

[20] This includes μ_Y if you consider a Normal law for ε

Solution:

We check the memoryless property, namely that

$$\mathbb{P}\big(Y_t \in A \mid \mathcal{I}_{t-1}\big) := \mathbb{P}\big(Y_t \in A \mid Y_{t-1}, Y_{t-2}, \dots\big) = \mathbb{P}\big(Y_t \in A \mid \sigma(Y_{t-1})\big)$$

We compute the conditional probability of this RHS in question a., and we saw that it is a function that only depends on $Y_{t-1}(s^{t-1})$, and doesn't require the knowledge (memory) of past events beside this value Y_{t-1}. Hence, $\{Y_t\}_t$ is a sequence of random variables respecting the Markov property, hence it is a Markov process, or discrete-time Markov chain.

Let us assume that $y \in \mathbb{N}$ and $\rho_Y = 1$. As a consequence, the list of state is discrete (but infinite), i.e. $\mathbb{N}$, because ε is a discrete random variable. If $y \notin \mathbb{N}$ or $\rho_Y \neq 1$, then the state spaces would be respectively $\mathbb{N} + \{y\}$ and $\mathbb{R}$. If we assume that $\varepsilon \sim \mathcal{N}(\mu_Y, \sigma_Y)$, the process would still be Markovian (because ε is i.i.d. over time) but the state space would be $\mathbb{R}$.

The transition matrix is $\{P_{i,j}\}_{i,j\in\mathbb{N}}$ with :

$$P_{i,j} = \mathbb{P}(Y_t = j \mid Y_{t-1} = i/\rho_Y) = \begin{cases} q & \text{if } j = i-1 \\ (1-p-q) & \text{if } j = i \\ p & \text{if } j = i+1 \end{cases}$$

If $\varepsilon \sim \mathcal{N}(\mu_Y, \sigma_Y)$, then $\frac{Y_t - \rho_Y Y_{t-1} - \mu_Y}{\sigma_Y} \sim \mathcal{N}(0,1)$, and hence the Markov kernel would be

$$p(dx, y_{t-1}) = \phi\Big(\frac{x - \rho_Y y_{t-1} - \mu_Y}{\sigma_Y}\Big)dx$$

with $\phi(\cdot)$ the density of the standard normal distribution. This is a hint for question f.

$\cdots$——— $\cdots$

(d.) Assume that $\rho_Y = 1$. This process is called a random walk. Suppose that there are many agents with the same income process and $y \in \mathbb{N}$. What is the law of motion of the distribution π_t over the income "states"?

- Give explicitly the distribution of income in date $t = 1$ and $t = 2$, when $y = 2$
- Do you expect this process to have a stationary distribution of over income states? (if yes, which one? if not, why not?)

Solution:

Let us denote $\pi_t = [\dots \pi_{i,t} \dots]$, an infinite row vector, and $P = \{P_{i,j}\}_{i,j\in\mathbb{N}}$ the transition matrix. The law of motion is given by the equation:

$$\pi_t = \pi_{t-1}P$$

More explicitly, we have :

$$\pi_{i,t} = \pi_{i-1,t-1}\, p + \pi_{i,t-1}(1-p-q) + \pi_{i+1,t-1}\, q \qquad \forall i$$

When $y = 2$, the distribution is concentrated at 2 : $p_0 = [0, \dots, 0, 1, 0, \dots]$.

At time $t = 1$, the distribution is $\pi_1 = [0\dots 0, q, 1-p-q, p, 0, \dots]$, when the three values are respectively at column 1, 2 and 3. Similarly $\pi_2 = [\dots p_{i,2}, \pi_{i+1,2}, \dots]$ is :

$$\pi_{i,2} = \begin{cases} 0 & \text{if } i < 0 \\ q^2 & \text{if } i = 0 \\ q(1-p-q) + (1-p-q)q & \text{if } i = 1 \\ qp + (1-p-q)^2 + pq & \text{if } i = 2 \\ 2p(1-p-q) & \text{if } i = 3 \\ p^2 & \text{if } i = 4 \\ 0 & \text{if } i > 4 \end{cases}$$

Where, for example the mass at $i = 2$ is the sum of the mass of people moving down and then up, the people moving up and then down, and the people staying in place for two periods.

We can realize that the distribution spreads over time. If $q > 0$ and/or $p > 0$ The stationary distribution converges to $\pi_\infty = [0, \dots, 0, \dots 0]$

$\cdots$ ——— $\cdots$

(e.) Assume that $\rho_Y < 1$. Write the expression of $Y(s^t)$ as a function of past shocks and parameters.
- In the VAR language, rewrite this with lag operators. Moreover, if the first date was $t_0 \to -\infty$, can you reexpress this process as a MA(∞) process? As a consequence, can you rewrite this lag polynomial in a more compact way? (Hint: try to "inverting" an operator "as if" it was a standard division and use the rule for geometric series).
- In this setting, (or in the case where $t_0 = 0$ and $y = 0$) can you compute the variance of $\mathbb{V}\text{ar}(Y_t)$ and the covariance $\mathbb{C}\text{ov}(Y_t, Y_{t-1})$? (Hint: recall that $\mathbb{V}\text{ar}(Z_t) = \mathbb{E}(Z_t^2) - \mathbb{E}(Z_t)^2$, for random variables Z_t that admit a finite second moment).
- Using the answer for the last subquestion, can you propose a method of moments (GMM) to estimate p, q and ρ_Y.

———

Solution:

(i) Using a recursion formula, we have:

$$Y_t = \rho_Y Y_{t-1} + \varepsilon_t = \rho_Y^2 Y_{t-2} + \rho_Y \varepsilon_{t-1} + \varepsilon_t = \cdots = \rho_Y^{t-t_0} Y_{t_0} + \sum_{k=0}^{t-t_0-1} \rho_Y^k \varepsilon_{t-k} \qquad (*)$$

(ii) Using Lag operators, such that $Y_{t-1} = LY_t$, and $Y_{t-k} = L^k Y_t$ we have:

$$Y_t = \rho_Y L Y_t + \varepsilon_t = \rho_Y^2 L^2 Y_t + \rho_Y L \varepsilon_t + \varepsilon_t = \cdots = \rho_Y^{t-t_0} L^{t-t-0} Y_t + \sum_{k=0}^{t-t_0-1} \rho_Y^k L^k \varepsilon_t$$

If $t_0 \to -\infty$, and because $\rho_Y < 1$ (s.t. we have $\rho_Y^{t-t_0} \to 0$), the power series converges and the initial term Y_{t_0} fades away. Hence we have:

$$Y_t = \sum_{k=0}^{t-t_0-1} \rho_Y^k L^k \varepsilon_t$$

Moreover, heuristically, because we know that the sum of the power series is : $\sum_{k=0}^{T} \alpha^k e = \frac{(1-\alpha^{T+1})e}{1-\alpha}$, we can intuitively obtain:

$$\begin{aligned} Y_t &= \lim_{t_0 \to -\infty} \sum_{k=0}^{t-t_0-1} \rho_Y^k L^k \varepsilon_t \\ &= \lim_{t_0 \to -\infty} (1 - \rho_Y L)^{-1} (1 - \rho_Y^{t-t_0-1+1}) \varepsilon_t \\ &= (1 - \rho_Y L)^{-1} \varepsilon_t \end{aligned}$$

Another way to see it :

$$Y_t = \rho_Y Y_{t-1} + \varepsilon_t \qquad \Rightarrow \qquad Y_t - \rho_Y Y_{t-1} = (1-\rho_Y L)Y_t = \varepsilon_t$$
$$\Rightarrow \qquad Y_t = (1-\rho_Y L)^{-1}\varepsilon_t$$

(iii) To compute the variances, we use the equation $(*)$ and the fact that $\mathbb{V}\mathrm{ar}(Z) = \mathbb{E}(Z^2) - \mathbb{E}(Z)^2$. First, see that because either the first term fades away (when $t_0 = -\infty$) or is null when $y = 0$, we have the second term:

$$\begin{aligned}
\mathbb{E}[Y_t] &= \mathbb{E}\Big[\sum_{k=0}^{t-t_0-1} \rho_Y^k \varepsilon_{t-k}\Big] \\
&= \sum_{k=0}^{t-t_0-1} \rho_Y^k \mathbb{E}[\varepsilon_{t-k}] = \sum_{k=0}^{t-t_0-1} \rho_Y^k (p-q) \\
&= \frac{1-\rho_Y^{t-t_0}}{1-\rho_Y}(p-q)
\end{aligned}$$

Now, for the second moment, we know that $\mathbb{E}[\varepsilon_t^2] = p+q$ and $\mathbb{E}[\varepsilon_t] = p-q$. It is easier to work directly with the *centered* shocks: write $\varepsilon_t = \mu_Y + \eta_t$ with $\mu_Y = \mathbb{E}[\varepsilon_t] = p-q$ and $\mathbb{E}[\eta_t] = 0$, $\mathbb{V}\mathrm{ar}[\eta_t] = \mathbb{V}\mathrm{ar}[\varepsilon_t] = \sigma_Y^2 = (p+q) - (p-q)^2$. ***Careful***: independence gives $\mathbb{E}[\varepsilon_{t-k}\varepsilon_{t-k'}] = \mathbb{E}[\varepsilon_{t-k}]\mathbb{E}[\varepsilon_{t-k'}] = \mu_Y^2$ for $k \neq k'$, which is *not* zero unless $p = q$ – this is why we center the shocks first. Using the linearity of the variance operator on independent terms:

$$\begin{aligned}
\mathbb{V}\mathrm{ar}[Y_t] &= \mathbb{V}\mathrm{ar}\Big[\sum_{k=0}^{t-t_0-1} \rho_Y^k \varepsilon_{t-k}\Big] = \sum_{k=0}^{t-t_0-1} \rho_Y^{2k}\mathbb{V}\mathrm{ar}[\varepsilon_{t-k}] \\
&= \frac{1-\rho_Y^{2(t-t_0)}}{1-\rho_Y^2}\sigma_Y^2 \quad \xrightarrow[t_0\to-\infty]{} \quad \frac{\sigma_Y^2}{1-\rho_Y^2} = \frac{(p+q)-(p-q)^2}{1-\rho_Y^2}
\end{aligned}$$

Similarly, for the covariance, only the terms with the same shock survive (by changing the indices $k'' = k'+1$, the shock η_{t-k} appears in Y_{t-1} with one power of ρ_Y less):

$$\begin{aligned}
\mathbb{C}\mathrm{ov}[Y_t, Y_{t-1}] &= \mathbb{C}\mathrm{ov}\Big[\sum_{k=0}^{t-t_0-1} \rho_Y^k \eta_{t-k}\,,\ \sum_{k'=0}^{t-t_0-2} \rho_Y^{k'} \eta_{t-1-k'}\Big] \\
&= \sum_{k''=1}^{t-t_0-1} \rho_Y^{k''}\rho_Y^{k''-1}\sigma_Y^2 = \rho_Y \frac{1-\rho_Y^{2(t-t_0-1)}}{1-\rho_Y^2}\sigma_Y^2 \quad \xrightarrow[t_0\to-\infty]{} \quad \rho_Y\frac{\sigma_Y^2}{1-\rho_Y^2}
\end{aligned}$$

so that $\mathbb{C}\mathrm{ov}[Y_t, Y_{t-1}] = \rho_Y \mathbb{V}\mathrm{ar}[Y_t]$: the first-order autocorrelation of an AR(1) is simply ρ_Y. These expressions hold verbatim in the gaussian case $\varepsilon \sim \mathcal{N}(\mu_Y, \sigma_Y)$, with σ_Y^2 now the variance of the normal law.

(iv) We have the following moments conditions:

$$\mathbb{E}[Y_t] = \frac{\mu_Y}{1-\rho_Y} = \frac{p-q}{1-\rho_Y} \qquad \mathbb{V}\mathrm{ar}[Y_t] = \frac{\sigma_Y^2}{1-\rho_Y^2} = \frac{(p+q)-(p-q)^2}{1-\rho_Y^2}$$
$$\mathbb{C}\mathrm{ov}[Y_t, Y_{t-1}] = \rho_Y\frac{\sigma_Y^2}{1-\rho_Y^2} = \rho_Y\frac{(p+q)-(p-q)^2}{1-\rho_Y^2}$$

Let us call $\overline{Y}$, $var(Y_t)$ and $cov(Y_t, Y_{t-1})$ the empirical counterparts. An estimator given by the generalized method of moment should search for the parameters p, q and ρ_Y such that the theoretical moments match the empirical ones. Note that here the system can even be inverted by hand: $\rho_Y = cov(Y_t, Y_{t-1})/var(Y_t)$, then $\sigma_Y^2 = (1-\rho_Y^2)var(Y_t)$ and $\mu_Y = (1-\rho_Y)\overline{Y}$, and finally p and q from

$\mu_Y = p - q$ and $\sigma_Y^2 = (p+q) - (p-q)^2$. More generally, one minimizes the quadratic loss function:

$$\min_{p,q,\rho_Y} \alpha_1 \Big[\frac{p-q}{1-\rho_Y} - \overline{Y}\Big]^2 + \alpha_2 \Big[\frac{(p+q)-(p-q)^2}{1-\rho_Y^2} - var(Y_t)\Big]^2 + \alpha_3 \Big[\rho_Y \frac{(p+q)-(p-q)^2}{1-\rho_Y^2} - cov(Y_t, Y_{t-1})\Big]^2$$

$$\min_{p,q,\rho_Y} (m(p,q,\rho_Y) - \mathcal{M}_N)' W (m(p,q,\rho_Y) - \mathcal{M}_N)$$

with $m(p,q,\rho_Y)$ is the vector of theoretical moments and $\mathcal{M}_N$ is the empirical counterpart for a sample of N observations. The second line is more general with a weighting matrix to balance the different terms in functions of their variances.

Here we have 3 moments and 3 parameters, so we are exactly identified. Adding more moments (like $\mathbb{C}\text{ov}(Y_t, Y_{t-k}), k \geq 2$), i.e. increasing the degree of overidentification, can lead to small sample bias (there is a tradeoff sample-size/number of moments conditions), and you will see more on that in your econometrics courses.

...——— ...

(f.) Assume $\varepsilon \sim \mathcal{N}(\mu_Y, \sigma_Y)$ for this question. Suppose you had data on $Y_1, Y_2, \dots, Y_T$, and you would like to estimate the parameters ρ_Y, μ_Y and σ_Y using Maximum Likelihood estimation (MLE). Write the (log-)likelihood, i.e. the objective of this maximization problem.

———

Solution:

The vector of parameter is $\theta = (\rho_Y, \mu_Y, \sigma_Y)$

$$\begin{aligned}
\ell(\theta) &= \log \prod_{t=1}^{T} \phi\Big(\frac{y_t - \rho_Y y_{t-1} - \mu_Y}{\sigma_Y}\Big) \\
&= \log \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi\sigma_Y^2}} \exp\Big(-\frac{(y_t - \rho_Y y_{t-1} - \mu_Y)^2}{2\sigma_Y^2}\Big) \\
&= -\frac{T}{2}\log(2\pi) - T\log(\sigma_Y) - \sum_{t=1}^{T} \frac{(y_t - \rho_Y y_{t-1} - \mu_Y)^2}{2\sigma_Y^2} \\
&= -\frac{T}{2}\log(2\pi) - T\log(\sigma_Y) - \frac{T}{2\sigma_Y^2}\mu_Y^2 - \frac{\mu_Y\rho_Y}{\sigma_Y^2}\sum_{t=1}^{T} y_{t-1} + \frac{\mu_Y}{\sigma_Y^2}\sum_{t=1}^{T} y_t - \frac{\rho_Y^2}{2\sigma_Y^2}\sum_{t=1}^{T} y_{t-1}^2 + \frac{\rho_Y}{\sigma_Y^2}\sum_{t=1}^{T} y_{t-1}y_t - \frac{1}{2\sigma_Y^2}\sum_{t=1}^{T} y_t^2
\end{aligned}$$

One could take the first order conditions w.r.t to μ_Y, σ_Y ρ_Y, to obtain three equations that allow to determine the MLE for these parameters. However, for more complicated problems it is frequent to optimize this log-likelihood numerically!

...——— ...

(g.) Suppose ε is again the original three states $\{+1, 0, -1\}$ random variable, with probability p, q such that $\mathbb{E}(\varepsilon) = 0$ and $\mathbb{V}\text{ar}(\varepsilon) = 1$ (what are these values of p and q?. Assume that the process Y_t is both a martingale and Markov chain (i.e. we have the properties for questions (b.) or (c.)).
- What do you expect this process to look like when the time period becomes infinitesimally small, i.e. $\Delta t = t_{i+1} - t_i \to 0$. Can you give the name of this stochastic process?
- More formally, define the rescaled process

$$W^{(n)}(t) = \frac{Y_{\lfloor nt \rfloor}}{\sqrt{n}}, \qquad t \in [0,1]$$

Can you determine if $W^{(n)}(1)$ converge in distribution toward a specific random variable? If yes, which one? What is the name of this theorem?
- (Advanced) Can one show this convergence for every t, and hence determine the law of $W(t)$ at the limit for n? What is the name of this theorem?

Solution:

When the time steps converge to zero, we obtain a Brownian motion (or Wiener process), the most simple stochastic process which is both a Markov process, a Martingale and continuous. The more general Levy processes (Jump and diffusion processes) are also Markov process and can be martingales, but may not be continuous.

Formally, $W^{(n)}(1) = \frac{Y_{\lfloor n \rfloor}}{\sqrt{n}} = \frac{\sum_{k=1}^{n} \varepsilon_k}{\sqrt{n}}$, where ε_k has finite variance and i.i.d. Hence by the central limit theorem, $W^{(n)}(1)$ converges in distribution towards a standard gaussian random variable $\mathcal{N}(0,1)$.

Now, for every $t \in [0,1]$, $W^{(n)}(t)$ converges to the Brownian motion W_t, which has a law $\mathcal{N}(0,t)$, where the standard deviation in square root in time $\sigma(W_t) = \sqrt{t}$. This convergence in distribution happens in the (infinite dimensional) Skorokhod space $\mathcal{D}([0,1])$ (say, the space of continuous processes on $[0,1]$). This is called the Donsker theorem. This is also the foundation for the Kolmogorov-Smirnov test for non-parametric estimation (for estimation the cdf $F(x)$ of any random variables X on space $\mathbb{X}$, which is also a function in $\mathcal{D}(\mathbb{X})$).

Because the Brownian motion is the continuous time limit of any random walk, it explains its ubiquity in continuous-time models in economics and finance.

$\cdots$——— $\cdots$

7 The Household problem in Arrow-Debreu economies, (10 pts) State-dependent consumption & portfolio choice

Using a stochastic process for income, we assume that the market is complete in the Arrow Debreu (AD) sense: for each history s^t and each state of the world s_{t+1}, there exists an asset $A(s_{t+1}|s^t)$ that yield one unit of consumption in state s_{t+1}, that can be purchased at date s^t and at price $Q_t(s_{t+1}|s^t)$ (for compact notation: $A(s_{t+1}|s^t) \equiv A(s^{t+1})$ and $Q_t(s_{t+1}|s^t) = Q_t(s^{t+1})$). Now, given this complete market structure, the household/consumer maximizes its expected utility, given the probability $\pi(s^{t+1})$ of history s^{t+1}.

$$\max_{\{C_t(s^t)\}_0^\infty, \{A_{t+1}(s^{t+1})\}_0^\infty} \sum_{t=0}^{\infty} \sum_{s^t \subset \mathcal{I}_t} \beta^t \pi(s^t) U\big(C_t(s^t)\big)$$

$$\text{s.t.} \qquad C_t(s^t) + \sum_{s_{t+1}} Q_t(s_{t+1}|s^t) A(s_{t+1}|s^t) \leq A(s^t) + Y_t(s^t)$$

Similarly, using the Lagrangian optimization to derive the Euler equation, and specify the stochastic discount factor.

Solution:
Note that this time, the Lagrange multiplier is state-dependent $\lambda(s^t)$. The Lagrangian :

$$\mathcal{L}\big(C_t(s^t), \{A_{t+1}(s_{t+1}|s^t)\}_{s_{t+1} \subset \mathcal{I}_{t+1}}, \lambda_t(s^t)\big) = \sum_{t=0}^{\infty} \sum_{s^t \subset \mathcal{I}_t} \Big\{\beta^t \pi(s^t) U\big(C_t(s^t)\big) - \lambda_t(s^t) \Big[C_t(s^t) + \sum_{s_{t+1}} Q_t(s_{t+1}|s^t) A_{t+1}(s_{t+1}|s^t) - A(s^t) - Y_t(s^t)\Big]\Big\}$$

1. Stationarity (C_t, A_{t+1} and L_t):

$$[C_t] \qquad \beta^t \pi(s^t) U'(C_t^*(s^t)) = \lambda_t(s^t) \qquad \forall\, t \geq 0$$

$$[A_{t+1}(s_{t+1}|s^t)] \qquad [\frac{\partial U(C_t^*)}{\partial A_{t+1}} = 0] \qquad 0 = \lambda_t(s^t)\, Q_t(s_{t+1}|s^t) + \lambda_{t+1}(s^{t+1})(-1)$$

$$\text{(combining 1 \& 2)} \qquad Q_t(s_{t+1}|s^t) \pi(s^t) U'(C_t^*(s^t)) = \beta \pi(s^{t+1}) U'(C_{t+1}^*(s^{t+1}))$$

   The other optimality conditions (the ones that economist usually forget) are the same, but this time for every states-of-the-world:

2. Primal feasibility: $C_t(s^t) + \sum_{s_{t+1}} Q_t(s_{t+1}|s^t) A(s_{t+1}|s^t) \leq A(s^t) + Y_t(s^t)$ for $t \geq 0, \forall\ s^t \subset \mathcal{I}_t$
3. Dual feasibility: $\lambda_t(s^t) \geq 0 \quad \forall t \geq 0 \forall\ s^t \subset \mathcal{I}_t$
4. Complementarity: $\sum_{t=0}^{\infty} \lambda_t(s^t) \Big(C_t(s^t) + \sum_{s_{t+1}} Q_t(s_{t+1}|s^t) A(s_{t+1}|s^t) - A(s^t) - Y_t(s^t)\Big) = 0 \quad \forall t \geq 0, \forall\, s^t$
   (the more the budget constraint binds, the more the Lagrange multiplier increases)

We see that the Lagrange multiplier is proportional to the marginal utility of consumption and the probability of the event/state of the world s^t. Moreover, $\pi(s^{t+1}) = \pi(s^t \cup s_{t+1})$, and so $\frac{\pi(s^{t+1})}{\pi(s^t)} = \frac{\pi(s^t \cup s_{t+1})}{\pi(s^t)} = \pi(s_{t+1}|s^t)$. Rewriting the previous equations more compactly:

$$Q_t(s_{t+1}|s^t) = \beta \frac{\pi(s^{t+1})\, U'(C_{t+1}(s^{t+1}))}{\pi(s^t)\, U'(C_t(s^t))} = \pi(s_{t+1}|s^t) \beta \frac{U'(C_{t+1})}{U'(C_t)} \qquad [\textbf{\textit{Euler equation}}]$$

In asset pricing (and dynamic/stochastic macro) the stochastic discount factor (s.d.f.) $m(s_{t+1}|s^t)$ is defined such that:

$$Q_t = \mathbb{E}[m(\widetilde{s}_{t+1}) R_t(\widetilde{s}_{t+1}) | \mathcal{I}_t]$$

Where Q_t is the asset price and $R_t(\widetilde{s}_{t+1}) = \mathbb{1}\{\widetilde{s}_{t+1} = s_{t+1}\}$ is its return – which is one for state of the world s_{t+1}, and zero otherwise, in our case here – with $\widetilde{s}_{t+1}$ the (dummy) variables on which we integrate in the expectation. Hence the s.d.f is again the ratio of marginal utilities, this time degenerated at the event s^{t+1}, i.e.

$$m(\widetilde{s}_{t+1}|s^t) = \beta\frac{U'(C_{t+1}(\widetilde{s}_{t+1}))}{U'(C_t(s^t))}$$

which makes sense, because in our (rather irrealistic) Arrow Debreu economy, there is one asset (and asset price) for each state of the world s_{t+1}, and hence :

$$Q_t(s_{t+1}|s^t) = \mathbb{E}[m(\widetilde{s}_{t+1})\mathbb{1}\{\widetilde{s}_{t+1} = s_{t+1}\}|\mathcal{I}_t] = \pi(s_{t+1}|s^t)m(s_{t+1})$$

In the next exercise, we will remove this strong assumption of complete markets. ***Keep this formula in mind***: with complete-markets, there is a full menu of traded payoffs, shown by the indicator $\mathbb{1}\{\widetilde{s}_{t+1} = s_{t+1}\}$, i.e. *one asset per state*. Therefore, the household equates $\beta U'(C_{t+1})/U'(C_t)$ to a price *state by state*: the marginal rate of substitution is pinned down separately in every s_{t+1}, and the household can therefore choose a consumption profile that no longer depends on which idiosyncratic shock it draws – this is *perfect risk sharing*, and it is the reason complete markets are the natural benchmark.

$\cdots$ ——— $\cdots$

8 The Household problem in incomplete markets (20 pts)

Consider the same problem as in the previous exercise. This time, the market is incomplete : there does not exist assets for each states of the world s_{t+1}. The only asset is a "risk-free bond"

$$\max_{\{C_t(s^t)\}_0^\infty, \{A_{t+1}(s^{t+1})\}_{s_{t+1}}\}_0^\infty} \sum_{t=0}^{\infty} \sum_{s^t \subset \mathcal{I}_t} \pi(s^t)\beta^t U\big(C_t(s^t)\big)$$

$$\text{s.t.} \qquad C_t(s^t) + Q_t(s^t)A_{t+1}(s^t) \leq A_t(s^t) + Y_t(s^t)$$

To alleviate notations, you can remove the notations s^t where it is clear that the variables is *not* measurable w.r.t. future shocks, information $s_{t+1} \subset \mathcal{I}_{t+1}$ but $\not\subset \mathcal{I}_t$

For that, we will use a "Dynamic programming" procedure. Again, the HH maximizes over the consumption path and we will obtain a Euler equation.

The idea here is to solve[21] the "Bellman" equation: the utility function can be rewritten as an iterative "Value function":

$$\begin{aligned}
V_t(A_t, Y_t) &= \max_{\{C_t\}_t, \{A_{t+1}\}_t} \mathbb{E}\Big[\sum_{j=t}^{\infty} \beta^{j-t} U(C_j) \big| \mathcal{I}_t\Big] \\
&= \max_{\{C_t\}_t, \{A_{t+1}\}_t} \sum_{j=t}^{\infty} \sum_{s^j \subset \mathcal{I}_j} \pi(s^j|s^t)\beta^{j-t} U(C_j) \\
\Leftrightarrow \qquad V_t(A_t, Y_t) &= \max_{C_t, A_{t+1}} \Big[U(C_t) + \beta \mathbb{E}\big(V_{t+1}(A_{t+1}, Y_{t+1}) \big| \mathcal{I}_t\big)\Big] \\
&= \max_{C_t, A_{t+1}} \Big[U(C_t) + \beta \sum_{s_{t+1} \subset \mathcal{I}_{t+1}} \pi(s_{t+1}|s^t)\, V_{t+1}(A_{t+1}, Y_{t+1}(s_{t+1}))\Big]
\end{aligned}$$

To solve this problem, you need to derive
(i) the KKT FOC-conditions for this new (iterative) maximization problem[22], and
(ii) a new "Envelope condition". This second condition shows the derivative of the value function (w.r.t. A_t) and, through the use of the chain-rule, is linked to the "Envelope Theorem". In this theorem, when a function is maximized over a variable (V_t is max over A_{t+1} here) and hence optimal in this sense, its derivative w.r.t. this same variable A_{t+1} is null.

Combining the different equations and manipulating the terms, show the Euler equation and show the similarities and differences with the Arrow-Debreu/complete market economy.

Solution:

(0.) The **Envelope Theorem** is the same tool as in Section 3: for a value function $v(x) = \max_{y \in K} f(x, y)$ whose optimum is interior and unique, $\nabla v(x) = \nabla_x f(x, y^*(x))$ – the derivative of v is the derivative of f with respect to its *first* (state) variable only, because $\nabla_y f(x, y^*(x)) = 0$ by the first-order conditions on the controls. We refer to the solution of Section 3 for the statement and the assumptions, and go directly to the economic problem here.

Let us come back to our **Bellman equation**:

$$V_t(A_t, Y_t) = \max_{C_t, A_{t+1}} U(C_t) + \beta \mathbb{E}\Big[V_{t+1}(A_{t+1}, Y_{t+1}(s_{t+1}) \Big| \mathcal{I}_t\Big]$$

Here, the value function depends on the state A_t, i.e. saving, (in most macro models, it will be the wealth in one form or another) and on the stochastic Y_t, i.e. income (most of the time an exogenous

[21] During the lecture, we discussed the general method to solve such problem (and the proof that show that the sequential problem and its recursive formulation are equivalent.

[22] Note that we are back to the first "static" version of the KKT theorem (exo 0.1)

economic shock), while it is optimal w.r.t. the controls C_t, A_{t+1}.

(1.) First, we need to derive the **FOC conditions** of the problem at time t to obtain the policy functions ($C_t = \mathcal{C}(A_t, Y_t)$ and $A_{t+1} = \mathcal{A}(A_t, Y_t)$ and determined using the ... KKT theorem). Rewrite the inequality constraint and Lagrangian: everything is state dependent, all the variables should have the s^t, but we only specify Y_t and $\lambda_t(s^t)$ to make clear where the shock comes from and affect the rest of the system:

$$F_t(C_t, A_{t+1}|s^t) = C_t + Q_t A_{t+1} - A_t - Y_t(s^t) \leq 0$$
$$\mathcal{L}_t(C_t, A_{t+1}, \lambda_t|s^t) = U(C_t) + \beta\mathbb{E}\Big[V_{t+1}(A_{t+1}, Y_{t+1})\Big|\mathcal{I}_t\Big] - \lambda_t\, F_t(C_t, A_{t+1}|s^t)$$

The FOC yields the following equations (note that we won't rewrite the primal and dual feasibility and complementarity conditions that are exactly the same as above) deriving w.r.t. C_t and A_{t+1}. Note that again, the Lagrange multiplier is state-dependent $\lambda(s^t)$. The Lagrangian :

$$[C_t] \qquad U'(C_t^*) = \lambda_t(s^t)$$
$$[A_{t+1}] \qquad \sum_{s_{t+1}} \pi(s_{t+1}|s^t)\beta \frac{\partial V_{t+1}(A_{t+1}^*, Y_{t+1}(s_{t+1}))}{\partial A_{t+1}} = \lambda_t(s^t) Q_t \qquad \{\frac{\partial U(C_t)}{\partial A_{t+1}} = 0\}$$
$$\text{(combining 1 \& 2)} \qquad Q_t U'(C_t) = \beta \sum_{s_{t+1}} \pi(s_{t+1}|s^t) \partial_A V_{t+1}(A_{t+1}^*, Y_{t+1}(s^{t+1}))$$

(2.) Now, to derive the **envelope condition**, we need to obtain the derivative of V_t (w.r.t. A_t) (why? In the last equation we still have the term $\partial_A V_{t+1}(A_{t+1}, Y_{t+1})$ unknown). Coming back to the Bellman equation, let us obtain $\partial_A V(A_t, Y_t)$. Here, we can "simply" use the Envelope theorem: we need to derive the value function w.r.t. the state variable, while it is already maximized over the three controls. ***But***, there is a trick, in the sense that, when the budget constraint is binding, one of the control can be expressed as a function of the other (state and control) variables:

$$C_t = \phi(A_t, A_{t+1}, Y_t) = -Q_t A_{t+1} + A_t + Y_t$$

This implies that, after using the chain-rule, the envelope condition is expressed as:

$$\partial_A V(A_t, Y_t) = U'(C_t)\, \frac{\partial \phi(A_t, A_{t+1}, Y_t)}{\partial A_t} + \underbrace{\left[U'(C_t) \frac{\partial \phi(A_t, A_{t+1}, Y_t)}{\partial A_{t+1}} + \beta\, \mathbb{E}_t \frac{\partial V_{t+1}(A_{t+1}, Y_{t+1})}{\partial A_{t+1}}\right]}_{=0 \;\; \text{(FOC on } A_{t+1})} \frac{dA_{t+1}}{dA_t}$$

And ***here*** we can use the Envelope theorem: the whole bracket is exactly the first-order condition with respect to the control A_{t+1}, hence it is zero at the optimum, as $\beta\mathbb{E}_t[\partial_A V_{t+1}] = Q_t U'(C_t)$. Therefore:

$$\begin{aligned} \partial_A V(A_t, Y_t(s_t)) &= U'(C_t)\, \frac{\partial \phi(A_t, A_{t+1}, L_t)}{\partial A_t} \\ &= U'(C_t) \times 1 \end{aligned}$$

In this very simple case, the marginal value of wealth (i.e. saving) equals the marginal utility of present consumption. Surprisingly this intuition holds in many macro models.

(3.) Since, this condition holds at all time, for all s^t, and we can iterate ($t \to t+1$), and obtain:

$$\partial_A V_{t+1}(A_{t+1}, Y_{t+1}(s_{t+1})) = U_c(C_{t+1})$$

and replacing in the last FOC equation above :

$$Q_t U_c(C_t) = \beta \mathbb{E}\big[\partial_A V_{t+1}(A_{t+1}, Y_{t+1}(s_{t+1})\big|\mathcal{I}_t\big] = \beta \mathbb{E}\big[U'(C_{t+1}(s_{t+1}))\big|\mathcal{I}_t\big]$$

We got the ***Euler equation*** (intertemporal FOC)! This implies the asset pricing equation:

$$Q_t = \beta \mathbb{E}\Big[\frac{U'(C_{t+1})}{U'(C_t)}\Big|\mathcal{I}_t\Big]$$

Hence, the stochastic discount factor is again the ratio of marginal utilities.

$$Q_t = \mathbb{E}[m(\widetilde{s}_{t+1})R_t(\widetilde{s}_{t+1})|\mathcal{I}_t] \qquad\qquad m(\widetilde{s}_{t+1}|s^t) = \beta\frac{U'(C_{t+1}(\widetilde{s}_{t+1}))}{U'(C_t(s^t))}$$

With again Q_t asset price, R_t asset return – one here – and $\widetilde{s}_{t+1}$ the (dummy) variable of integration. Hence the s.d.f is indeed *stochastic* (that the standard case!), as it depends on s^{t+1}. This represents how households "anticipate" future shocks – the foundation of rational expectations!

Complete vs. incomplete markets – the point of Sections 7 and 8.

Compare the two economies: in the Arrow-Debreu economy, with complete markets, there is one Euler equation *per state of the world* – the household can trade a claim contingent on each s_{t+1}, equalize its marginal utility across states, and thereby insure its idiosyncratic income risk completely: consumption ends up depending only on the aggregate state, not on the household's own draw of Y_t (*perfect risk sharing*). With a single riskless bond, only *one* Euler equation holds, and it holds only *in expectation*: $Q_t U'(C_t) = \beta\mathbb{E}_t[U'(C_{t+1})]$. The household can no longer move resources across states, only across time, so consumption inherits the fluctuations of income and can only be smoothed by saving and dissaving – this is *self-insurance*, and it is why the buffer of assets A_t becomes a state variable of the problem. This contrast – perfect insurance vs. self-insurance – is the starting point of the Bewley-Huggett-Aiyagari literature on heterogeneous agents.

With the dynamic programming methods, that seems a lot of work for one simple (and almost already known) equation ... but, the good news is that this method is general and can be applied to all kind of problems (especially when associated with the KKT theorem for inequality constraints). In the macro sequences, you will use the Dynamic programming approach to study macro models with incomplete markets, credit constraints, nominal rigidities, search & matching frictions, heterogeneous agents, etc. There will be two main methods: use the Bellman algorithm numerically (solving with an approximation of the value function in order to obtain the policy function) or using make clever assumptions to be able to derive the model analytically.