SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Discrete-time dynamic programming applied to economic theory

av

Markus Pettersson

2019 - No K25

Discrete-time dynamic programming applied to economic theory

Markus Pettersson

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Yishao Zhou

2019

Contents

1. Introduction

An important area within optimisation concerns decision making in the presence of time. In such dynamic contexts where decisions are made in stages, the outcome of a decision made today will typically affect not only the optimal decision but also the space of available decisions in subsequent stages. Therefore, decisions cannot be viewed in isolation since the desire for an optimal outcome today must be balanced against the desire for optimal outcomes in future stages. As an illustration, consider the following prototypical problem:

Example 1.1 (The stagecoach problem)**.** A stagecoach is travelling from A to J in Figure 1.1. An arrow from one node to another indicates a possible route to travel and its label indicates the length of that route. What is the shortest path from A to J? $\triangleleft$

In Example 1.1, choosing the shortest route at each stage yields the overall route $A \to B \to F \to I \to J$ of total length 13. This is not the shortest overall route however; for instance, the route $A \to D \to F \to I \to J$ of length 11 is shorter.

Dynamic programming presents a way of solving problems of this type. Introduced by Bellman (1952, 1953), dynamic programming simplifies the problem at hand by recursively breaking it down into a collection of simpler subproblems where each of those subproblems are solved once. The purpose of this thesis is to address this topic. Specifically, it aims to describe the underlying theory of dynamic programming and to formulate the type of problems it can solve.

Since its introduction in the 1950s, dynamic programming has become a standard tool in a variety of applied fields that rely heavily on optimisation. One such field is economics, where solving theoretical models in areas ranging from game theory to labour economics to macroeconomics typically boils down to some dynamic optimisation problem. Indeed, in their standard reference on modern macroeconomics, Ljungqvist and Sargent (2012, ch. 1) discuss "the imperialism of recursive methods" and point out that

> *Dynamic programming is now recognized as a powerful method for studying private agents' decisions and also the decisions of a government that wants to design an optimal policy in the face of constraints imposed on it by private agents' best responses to that government policy.*

A second aim of this paper is therefore to illustrate and analyse the application of dynamic programming within economic theory. We do so by using our theory to derive two cornerstone results in economics – the permanent income hypothesis and Tobin's q – as well as to practically solve two economic models – one labour economics search model and one macroeconomic growth model.

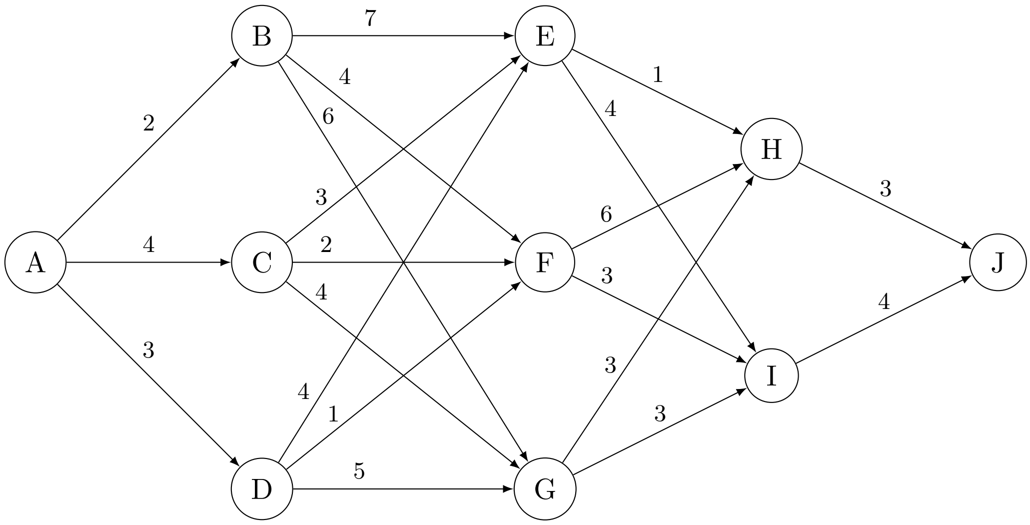

Figure 1.1. The stagecoach problem

In short, dynamic programming is used to solve models of dynamic optimisation, and such models in general require several important considerations. For instance:

(i) Should time be considered continuous or discrete?
(ii) Does the problem have a fixed end stage or does it proceed infinitely far into the future?
(iii) Is the problem deterministic or stochastic? That is, are there random variables that are out of the decision maker's control that should be considered?
(iv) In case of a stochastic problem, is the decision maker fully aware of the state he or she is in (or the location he or she is at in Example 1.1) at all times? In other words, is there perfect information with respect to the state or is some estimation of the state required?

In what follows, we will consider discrete-time problems with perfect state information only. We also include stochastic elements but keep it necessarily simple in order to avoid the use of measure theory and Markov chains, which go beyond the scope of this paper. Under these premises, we first consider a rather general theory for problems with finite horizon and then extend this theory into a subclass of problems with infinite time horizon. In particular, we consider stationary and discounted infinite-time problems with a bounded goal function. These choices reflect the fact that discrete-time and discounted models are by far the most common choice considered in economics.

The rest of the paper proceeds as follows. The theory of dynamic programming with finite time horizon is presented in Section 2. In Section 3 we extend this framework to problems with infinite time horizon. The applications of dynamic programming in economics are covered in Section 4 and Section 5 concludes.

2. Dynamic programming in finite time

In this section, we cover the basic theory behind dynamic programming in finite time. In particular, we formulate the type of problems for which we can apply dynamic programming to and present the algorithm used to solve them. In what follows below, Definitions 2.1 and 2.2 are generalised versions of deterministic counterparts in Voorneveld (2016, ch. 25), while Theorem 2.2 with corresponding proof is from Bertsekas (2005, ch. 1). Lemma 2.1 is formalised and proved independently.

2.1 Problem formulation

In general, suppose we wish to optimise the additively separable objective function

$$\sum_{t=0}^{T} g_t\big(x_t, u_t, z_t\big),$$

where x_t evolves according to the system

$$x_{t+1} = f_t\big(x_t, u_t, z_t\big), \qquad t = 0, \ldots, T-1.$$

Here, t indexes time and T is the horizon of the system. At each time t there is

(i) a state vector x_t that summarises where the system is at time t. It lies in a set X called the state space. The initial state x_0 is assumed to be given;
(ii) a control vector u_t which is the choice vector. It lies in a set $U(x_t)$ called the control space, which in turn depends on the realised state;
(iii) a vector of random variables z_t drawn from the set $\mathcal{Z}$ (which may be empty).

In order to avoid the use of measure theory and Markov chains, we make the following simplifying assumption throughout the thesis:

Assumption 2.1. The vector of random variables z_t is drawn from a finite set $\mathcal{Z} = \{z_1, \ldots, z_N\}$ and is independent across time, states, and controls. That is, for each period t and each feasible history $x^t = \{x_t, \ldots, x_0\}$, $u^t = \{u_t, \ldots, u_0\}$, $z^{t-1} = \{z_{t-1}, \ldots, z_0\}$,

$$\Pr\big[z_t = z_i \,\big|\, x^t, u^t, z^{t-1}\big] \;=\; \Pr\big[z_t = z_i\big] \;\equiv\; p_i \qquad \text{with} \qquad \sum_{i=1}^{N} p_i = 1,$$

where p_i denotes the constant unconditional probability of observing z_i. ◁

The presence of a stochastic element z_t makes the optimisation problem stochastic as we do not know *ex ante* the future realisations of z_t. This implies that we in fact optimise the expectation of the objective function with respect to the probability distribution of z_t, rather than optimising the objective function itself. We can therefore formally define this optimisation problem as follows:

Definition 2.1 (The finite-time dynamic programming problem)**.** A discrete-time dynamic programming problem with finite horizon is of the form

$$\begin{aligned} &\sup_{\{u_t\}_{t=0}^T} && \mathbb{E}\left[\sum_{t=0}^T g_t\big(x_t, u_t, z_t\big)\right] && \\ &\text{subject to} && u_t \in U(x_t), && t = 0, \dots, T, \\ & && z_t \in \mathcal{Z}, && t = 0, \dots, T, \\ & && x_{t+1} = f_t\big(x_t, u_t, z_t\big), && t = 0, \dots, T-1, \\ & && x_0 \text{ given}, && \end{aligned}$$

where Assumption 2.1 is satisfied and $\mathbb{E}$ is the expectations operator with respect to z_t, defined as $\mathbb{E}[z_t] = \sum_{i=1}^N p_i z_i$. ◁

Note that we could just as well have stated Definition 2.1 as a minimisation problem by defining the function $h_t(x_t, u_t, z_t) = -g_t(x_t, u_t, z_t)$. We define the feasible options of the dynamic programming problem as follows:

Definition 2.2. Consider the dynamic programming problem in Definition 2.1.

- For x_0 given, a pair $(x, u) = \{(x_0, u_0), \dots, (x_T, u_T)\}$ such that $u_t \in U(x_t)$ for all t and $x_{t+1} = f_t(x_t, u_t, z_t)$ for all $t + 1 > 0$ is said to be admissible or feasible. An admissible pair is optimal and $u = \{u_0, \dots, u_T\}$ is an optimal control if there is no other admissible pair that yields a higher expected value of the objective function.
- A policy is a sequence of functions $\pi = \{\pi_0, \dots, \pi_T\}$ where π_t maps states x_t into controls $u_t = \pi_t(x_t)$ in $U(x_t)$ for all x_t. Thus, a policy π yields an admissible pair (x, u) and such a policy is optimal if there is no other policy with a higher expected value of the objective function. ◁

Since a policy yields an admissible pair, choosing a policy is equivalent to choosing controls u_t; given $u_t = \pi_t(x_t)$, for any function g of x_t, u_t, and z_t we have

$$\sup_{\pi_t \in \Pi_t} g\big(x_t, \pi_t(x_t), z_t\big) = \sup_{u_t \in U(x_t)} g\big(x_t, u_t, z_t\big),$$

where Π_t is the set of all functions $\pi_t(x_t)$ such that $\pi_t(x_t) \in U(x_t)$ for all x_t.

2.2 Bellman's principle of optimality

Solving the optimisation problem in Definition 2.1 relies on the *principle of optimality* which Bellman (1953) states as follows:

> *An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.*

In essence, the optimality principle states that if $u^* = \{u_0^*, \ldots, u_T^*\}$ is an optimal control for the dynamic programming problem and x_s at time s occurs with positive probability under this policy, then the sequence of controls $\{u_s^*, \ldots, u_T^*\}$ must be optimal in the subproblem of optimising the remaining objective function

$$\mathbb{E}\left[\sum_{t=s}^{T} g_t\big(x_t, u_t, z_t\big)\right],$$

that starts in period s and state x_s. We formulate this principle with the lemma below.

Lemma 2.1 (The principle of optimality)**.** *If the admissible pair (x^*, u^*) solves the dynamic programming problem, then $\{(x_t^*, u_t^*)\}_{t=s}^T$ solves the subproblem starting in period s with initial state x_s^*.*

Proof. We use induction over s. The result holds by assumption for $s = 0$. Now assume it holds for some $s \geq 0$. By linearity of the expectations operator, we can write

$$\mathbb{E}\left[\sum_{t=s}^{T} g_t(x_t, u_t, z_t)\right] = \mathbb{E}\Big[g_s(x_s, u_s, z_s)\Big] + \mathbb{E}\left[\sum_{t=s+1}^{T} g_t(x_t, u_t, z_t)\right].$$

Since $\{x_t^*, u_t^*\}_{t=s}^T$ is optimal, we have by definition that for any feasible pair $\{\hat{x}_t, \hat{u}_t\}_{t=s+1}^T$,

$$\mathbb{E}\Big[g_s(x_s^*, u_s^*, z_s)\Big] + \mathbb{E}\left[\sum_{t=s+1}^{T} g_t(x_t^*, u_t^*, z_t)\right] \geq \mathbb{E}\Big[g_s(x_s^*, u_s^*, z_s)\Big] + \mathbb{E}\left[\sum_{t=s+1}^{T} g_t(\hat{x}_t, \hat{u}_t, z_t)\right].$$

Simplifying this expression we get

$$\mathbb{E}\left[\sum_{t=s+1}^{T} g_t(x_t^*, u_t^*, z_t)\right] \geq \mathbb{E}\left[\sum_{t=s+1}^{T} g_t(\hat{x}_t, \hat{u}_t, z_t)\right],$$

so $\{(x_t^*, u_t^*)\}_{t=s+1}^T$ is also optimal for the subproblem starting in period $s+1$ with initial state x_{s+1}^*. □

This result can be illustrated with the following simplistic travel analogy: if the fastest route from Stockholm to Copenhagen goes through Malmö, then the Malmö-to-Copenhagen part of that route is also the fastest route from Malmö to Copenhagen.

2.3 The dynamic programming algorithm

The principle of optimality suggests that we can solve the dynamic programming problem as follows. First, find the optimal control for the subproblem involving the last period only. By the principle of optimality we know that the optimal last-period control u_T^* must also be optimal in period T in the subproblem involving the last two periods. Thus, we can substitute backwards and use this control to solve for the optimal controls of the subproblem involving the last two periods. Continuing backwards in this manner enables us to sequentially solve for the optimal controls of the full problem. More formally, for a given time $s \in \{0, \ldots, T\}$ and state x_s, define the *optimal value function* as

$$J_s^*(x_s) \;=\; \sup_{\{u_t\}_{t=s}^T} \mathbb{E}\left[\sum_{t=s}^{T} g_t\big(x_t, u_t, z_t\big)\right], \tag{2.1}$$

where $(x_t, u_t, z_t) \in X \times U(x_t) \times \mathcal{Z}$ for all $t \in \{s, \ldots, T\}$. For the final-period problem, where $s = T$, we need not worry about future consequences from the choice of control. Thus, $J_T^*(x_T)$ must necessarily satisfy

$$J_T^*(x_T) \;=\; \sup_{u_T \in U(x_T)} \mathbb{E}\Big[g_T\big(x_T, u_T, z_T\big)\Big].$$

Now, in the subproblem involving the last two periods, we know that choosing a control u_{T-1} yields an instantaneous payoff $g_{T-1}(x_{T-1}, u_{T-1}, z_{T-1})$ and a next-period state $x_T = f_{T-1}(x_{T-1}, u_{T-1}, z_{T-1})$. Since we know the optimal value of the subproblem involving only the final period, we therefore know that

$$J_T^*\big(x_T\big) \;=\; J_T^*\big(f_{T-1}(x_{T-1}, u_{T-1}, z_{T-1})\big).$$

Clearly, the best thing to do is to optimise the sum of these two expressions:

$$J_{T-1}^*(x_{T-1}) \;=\; \sup_{u_{T-1}} \mathbb{E}\Big[g_{T-1}(x_{T-1}, u_{T-1}, z_{T-1}) + J_T^*\big(f_{T-1}(x_{T-1}, u_{T-1}, z_{T-1})\big)\Big].$$

Continuing backwards we then get for each time $s \in \{0, \ldots, T-1\}$ and state x_s that

$$J_s^*(x_s) \;=\; \sup_{u_s \in U(x_s)} \mathbb{E}\Big[g_s(x_s, u_s, z_s) + J_{s+1}^*\big(f_s(x_s, u_s, z_s)\big)\Big].$$

This is exactly the dynamic programming algorithm which we state and prove below.

Theorem 2.2 (The dynamic programming algorithm)**.** *The optimal value function satisfies*

$$J_T(x_T) \;=\; \sup_{u_T \in U(x_T)} \mathbb{E}\Big[g_T\big(x_T, u_T, z_T\big)\Big], \tag{2.2}$$

$$J_s\,(x_s) \;=\; \sup_{u_s \in U(x_s)} \mathbb{E}\Big[g_s(x_s, u_s, z_s) + J_{s+1}\big(f_s(x_s, u_s, z_s)\big)\Big], \quad s = 0, \ldots, T-1, \tag{2.3}$$

where the expectation is taken with respect to the probability distribution of z_s. For x_0 given, it follows that $J_0(x_0)$ is the optimal value of the dynamic programming problem. Moreover, if $u_s^* = \pi_s^*(x_s)$ solves the right-hand sides of Equations (2.2) and (2.3) for each x_s and s, the policy $\pi^* = \{\pi_0^*, \ldots, \pi_T^*\}$ is optimal and $u^* = \{u_0^*, \ldots, u_T^*\}$ is an optimal control.

Proof. We use induction over s to show that for all $s \in \{0, \dots, T\}$, the optimal value function $J_s^*(x_s)$ defined by Equation (2.1) is identical to $J_s(x_s)$ given by the dynamic programming algorithm in Equations (2.2) and (2.3). For the base case, let $s = T$. Then by Equation (2.1), we have for all x_T that

$$J_T^*(x_T) = \sup_{\{u_t\}_{t=T}^T} \mathbb{E}\left[\sum_{t=T}^{T} g_t\big(x_t, u_t, z_t\big)\right] = \sup_{u_T} \mathbb{E}\Big[g_T\big(x_T, u_T, z_T\big)\Big],$$

which is indeed identical to Equation (2.2). So the base case holds. For the induction step, suppose Equations (2.2) and (2.3) hold for $s+1, \dots, T$ for some $s < T$. Then for all x_s,

$$\begin{aligned}
J_s^*(x_s) &= \sup_{\{u_t\}_{t=s}^T} \mathbb{E}\left[g_t\big(x_s, u_s, z_s\big) + \sum_{t=s+1}^{T} g_t\big(x_t, u_t, z_t\big)\right] && \text{(by (2.1))}\\
&= \sup_{u_s} \mathbb{E}\left[g_s\big(x_s, u_s, z_s\big) + \sup_{\{u_t\}_{t=s+1}^T} \mathbb{E}\left\{\sum_{t=s+1}^{T} g_t\big(x_t, u_t, z_t\big)\right\}\right] && \text{(by Lemma 2.1)}\\
&= \sup_{u_s} \mathbb{E}\Big[g_s\big(x_s, u_s, z_s\big) + J_{s+1}^*\big(f_s(x_s, u_s, z_s)\big)\Big] && \text{(by (2.1))}\\
&= \sup_{u_s} \mathbb{E}\Big[g_s\big(x_s, u_s, z_s\big) + J_{s+1}\big(f_s(x_s, u_s, z_s)\big)\Big] && \text{(induction hypothesis)}\\
&= J_s(x_s) && \text{(by (2.3))}.
\end{aligned}$$

It follows that the induction step also holds and this proves the theorem. □

We finish this section by demonstrating the dynamic programming algorithm by applying it on the stagecoach problem in Example 1.1.

Example 1.1 (Cont.)**.** In Figure 1.1, we wish to find the shortest path from A to J. This is equivalent to maximising the additive inverses of the path lengths, so we can apply the dynamic programming algorithm. Moreover, there is no stochastic element involved so $\mathcal{Z} = \varnothing$ and we do not have to consider expected values. In stage t, the state is the current location and the control is the route chosen. Then we immediately have that $J_3^*(H) = 3$ and $J_3^*(I) = 4$ since there is only one feasible control in these states; $u_3 = J$. Working backwards we see that

$$\begin{aligned}
J_2^*(E) &= \min\big\{1 + J_3^*(H),\ 4 + J_3^*(I)\big\} = \min\{4,\ 8\} = 4,\\
J_2^*(F) &= \min\big\{6 + J_3^*(H),\ 3 + J_3^*(I)\big\} = \min\{9,\ 7\} = 7,\\
J_2^*(G) &= \min\big\{3 + J_3^*(H),\ 3 + J_3^*(I)\big\} = \min\{6,\ 7\} = 6,
\end{aligned}$$

with corresponding optimal controls $u_2 = H$, $u_2 = I$, and $u_2 = H$, respectively. Going back one more period we have

$$\begin{aligned}
J_1^*(B) &= \min\big\{7 + J_2^*(E),\ 4 + J_2^*(F),\ 6 + J_2^*(G)\big\} = \min\{11,\ 11,\ 12\} = 11,\\
J_1^*(C) &= \min\big\{3 + J_2^*(E),\ 2 + J_2^*(F),\ 4 + J_2^*(G)\big\} = \min\{7,\ 9,\ 10\} = 7,\\
J_1^*(D) &= \min\big\{4 + J_2^*(E),\ 1 + J_2^*(F),\ 5 + J_2^*(G)\big\} = \min\{8,\ 8,\ 11\} = 8,
\end{aligned}$$

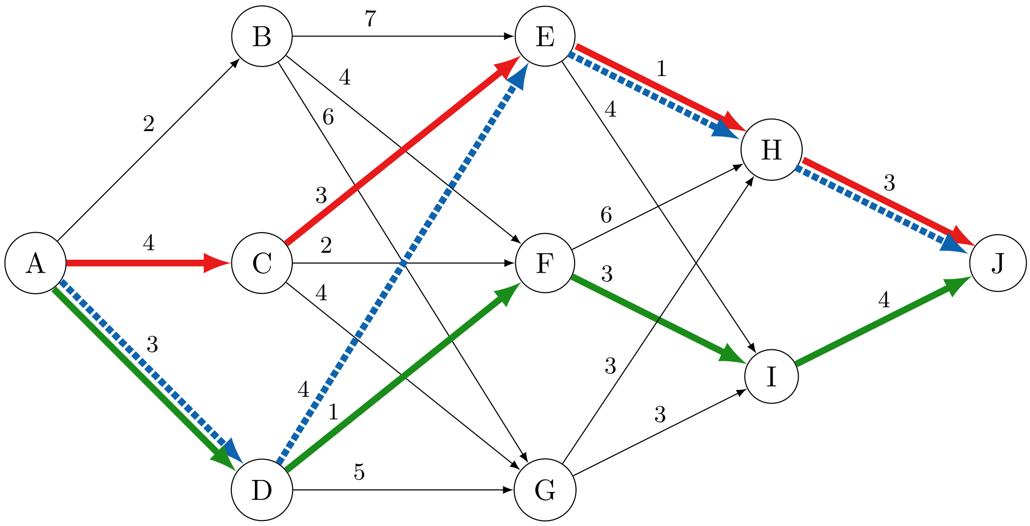

FIGURE 2.1. Solution to the stagecoach problem

with corresponding optimal controls $u_1 \in \{E, F\}$, $u_1 = E$, and $u_1 \in \{E, F\}$, respectively. It follows that the length of the shortest path is

$$J_0^*(A) = \min\left\{2 + J_1^*(B),\ 4 + J_1^*(C),\ 3 + J_1^*(D)\right\} = \min\{13,\ 11,\ 11\} = 11$$

and the routes that achieve this are

$$\begin{aligned} &A \to C \to E \to H \to J, \\ &A \to D \to E \to H \to J, \\ &A \to D \to F \to I \ \ \to J. \end{aligned}$$

These routes are illustrated in Figure 2.1. ◁

3. Extension to infinite time

We now generalise the results of the previous section to allow for an infinite time horizon. That is, we let $T \to \infty$. Doing so introduces a number of complications. For instance, the goal function now becomes a series, so we need to ensure that it is well-defined. Moreover, the dynamic programming algorithm in Section 2 works backward from some finite end period. In infinite time, no such period exists. Below we show how to account for these issues for a subclass of infinite-time problems, that of discounted, stationary problems with bounded goal function, and outline how to solve them. Unless stated otherwise, the theorems in this section are from Bertsekas (2007, ch. 1) while the proofs and definitions are generalised versions of deterministic counterparts in Voorneveld (2016, ch. 27–28).

3.1 Problem formulation

Bertsekas (2007, p. 2) identifies four principal classes of infinite-time dynamic programs: (i) stochastic shortest path problems; (ii) stationary discounted problems with bounded objective functions; (iii) problems with unbounded objective functions; and (iv) problems that optimise the average of the per-stage objective functions. Since the primary concern in this thesis is to investigate dynamic programming within the context of economic theory, we restrict our attention here to the second class of problems, that of bounded discounted problems, as it is by far the most common type in economics.

Definition 3.1 (The discounted infinite-time dynamic programming problem)**.** A stationary and discounted discrete-time dynamic programming problem with infinite horizon and discount factor $\beta \in (0, 1)$ is of the form

$$\begin{aligned}
&\sup_{\{u_t\}_{t=0}^{\infty}} && \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t g\big(x_t, u_t, z_t\big)\right] && \\
&\text{subject to} && u_t \in U(x_t), && t = 0, 1, 2, \dots, \\
& && z_t \in \mathcal{Z}, && t = 0, 1, 2, \dots, \\
& && x_{t+1} = f\big(x_t, u_t, z_t\big), && t = 0, 1, 2, \dots, \\
& && x_0 \text{ given},
\end{aligned}$$

where Assumption 2.1 is satisfied and $\mathbb{E}$ is defined as in Definition 2.1. The problem is stationary since neither g nor f depend on time. ◁

Admissible pairs, controls, and policies are defined analogously to Definition 2.2. To guarantee that the objective function is summable, we assume the following throughout:

Assumption 3.1. For some real scalar M, the function g satisfies $|g(x,u,z)| \leq M$ for all $(x,u,z) \in X \times U(x) \times \mathcal{Z}$. ◁

This makes the objective function well-defined; if (x,u) is an admissible pair, then

$$\mathbb{E}\left[\sum_{t=0}^{T} \beta^t \Big| g\big(x_t,u_t,z_t\big)\Big|\right] \leq \sum_{t=0}^{T} \beta^t M = \frac{1-\beta^{T+1}}{1-\beta} M \to \frac{M}{1-\beta} \quad \text{as} \quad T \to \infty,$$

making the left-hand side summable. Note also that the objective function in Definition 3.1 in fact should be $\lim_{T\to\infty} \mathbb{E}\Big[\sum_{t=0}^{T} \beta^t g\big(x_t,u_t,z_t\big)\Big]$, which is not in general equal to our formulation. However, Bertsekas (2007, pp. 3–4) points out that under Assumptions 2.1 and 3.1, these are indeed equal and thus allow for the formulation above. We prove this in Appendix A.

3.2 The Bellman equation and optimality

Solving the infinite-time dynamic programming problem still rests on the principle of optimality. In our case, this principle is completely analogous to the finite-horizon case; just let $g_t(\cdot) = \beta^t g(\cdot)$ and $T \to \infty$ in Lemma 2.1 and the corresponding proof to get that if (x^*, u^*) is optimal in the infinite-time problem, then $\{(x_t^*, u_t^*)\}_{t=s}^{\infty}$ is optimal in the infinite-time subproblem starting in period s with initial state x_s^*. Thus, in what follows we refer to Lemma 2.1 also as the infinite-horizon version of the optimality principle. Now, in order to solve the infinite-time problem, we start off as in the finite-horizon case and define the optimal value function:

Definition 3.2 (The value function)**.** For a given state $x \in X$, we define the optimal value function as

$$J^*(x) = \sup_{\{u_t\}_{t=0}^{\infty}} \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t g\big(x_t,u_t,z_t\big)\right], \tag{3.1}$$

where $x_t \in X$, $u_t \in U(x_t)$, and $z_t \in \mathcal{Z}$ for all t. ◁

This definition does not specify a time period for x, so the value function applies to any subproblem starting in an arbitrary time period s with initial state x_s. From this definition it is clear that the optimal value of such subproblem is $\beta^s J^*(x_s)$ and thus that $J^*(x_0)$ is the optimal value of the full dynamic program. Solving the original problem is therefore equivalent to finding $J^*(x_0)$. This is in general not a straightforward task, but the following theorem provides us with a helpful tool in this regard.

Theorem 3.1 (The Bellman equation)**.** *For all $x \in X$, the value function satisfies the Bellman equation*

$$J(x) = \sup_{u \in U(x)} \mathbb{E}\Big[g(x,u,z) + \beta J\big(f(x,u,z)\big)\Big], \tag{3.2}$$

where u is the control and z is the realisation of the stochastic element in the time period of state x.

Proof. For all $x \in X$,

$$
\begin{aligned}
J^*(x) &= \sup_{\{u_t\}_{t=0}^{\infty}} \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t g\big(x_t, u_t, z_t\big)\right] && \text{(by (3.1))} \\
&= \sup_{u} \mathbb{E}\left[g(x,u,z) + \sup_{\{u_t\}_{t=1}^{\infty}} \mathbb{E}\left\{\sum_{t=1}^{\infty} \beta^t g(x_t, u_t, z_t)\right\}\right] && \text{(by Lemma 2.1)} \\
&= \sup_{u} \mathbb{E}\left[g(x,u,z) + \beta \sup_{\{u_{t+1}\}_{t=0}^{\infty}} \mathbb{E}\left\{\sum_{t=0}^{\infty} \beta^t g(x_{t+1}, u_{t+1}, z_{t+1})\right\}\right] && \text{(relabelling)} \\
&= \sup_{u} \mathbb{E}\Big[g(x,u,z) + \beta J^*\big(f(x,u,z)\big)\Big] && \text{(by (3.1)).} \quad \square
\end{aligned}
$$

Theorem 3.1 essentially states that the value function $J^*(x)$ is a fixed point[1] to the mapping $F : X \to X$ defined by

$$
FJ(x) \;=\; \sup_{u \in U(x)} \mathbb{E}\Big[g(x,u,z) + \beta J\big(f(x,u,z)\big)\Big]. \tag{3.3}
$$

Also note the similarity between the Bellman equation (3.2) and the dynamic programming algorithm in Section 2. The difference here is that since we do not iterate backwards from some end period, the value function on the right-hand side is typically unknown. Moreover, the Bellman equation is only a necessary condition for the value function. There may, however, be other functions that satisfy the Bellman equation. The first of these two concerns can be handled via the following result:

Theorem 3.2. *For any bounded function $J : X \to \mathbb{R}$, the value function satisfies*

$$
J^*(x) \;=\; \lim_{N \to \infty} F^N J(x) \qquad \text{for all } x \in X,
$$

where $F : X \to X$ is defined by Equation (3.3) and F^N denotes the composition of F with itself N times. By convention, $F^0 J(x) \equiv J(x)$.

Proof (derived independently). We use induction to first show that

$$
F^N J(x_0) \;=\; \sup_{\{u_t\}_{t=0}^{N-1}} \mathbb{E}\left[\sum_{t=0}^{N-1} \beta^t g(x_t, u_t, z_t) + \beta^N J(x_N)\right].
$$

The base case $N = 0$ holds trivially as $F^0 J(x_0) \equiv J(x_0)$. For the induction step, suppose the result holds for some $N - 1 \geq 0$. Then by Equation (3.3),

$$
F^N J(x_0) \;=\; F\Big(F^{N-1} J(x_0)\Big) \;=\; \sup_{u_0} \mathbb{E}\Big[g(x_0, u_0, z_0) + \beta F^{N-1} J\big(f(x_0, u_0, z_0)\big)\Big]
$$

[1] A point $x \in X$ is called a fixed point under a function $f : X \to X$ if $x = f(x)$.

and invoking the induction hypothesis yields

$$
\begin{aligned}
F^N J(x_0) & \\
&= \sup_{u_0} \mathbb{E}\left[g(x_0,u_0,z_0) + \beta \sup_{\{u_{t+1}\}_{t=0}^{N-2}} \mathbb{E}\left\{ \sum_{t=0}^{N-2} \beta^t g(x_{t+1},u_{t+1},z_{t+1}) + \beta^{N-1} J(x_N) \right\}\right] \\
&= \sup_{\{u_t\}_{t=0}^{N-1}} \mathbb{E}\left[\sum_{t=0}^{N-1} \beta^t g(x_t,u_t,z_t) + \beta^N J(x_N) \right].
\end{aligned}
$$

So the result also holds for N which proves the result. Now, since J is bounded by assumption and $\beta \in (0,1)$, we necessarily have that $\beta^N J(x_N) \to 0$ as $N \to \infty$. It follows that

$$
F^N J(x_0) \;\to\; \sup_{\{u_t\}_{t=0}^{\infty}} \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t g\big(x_t,u_t,z_t\big) \right] = J^*(x_0) \qquad \text{as} \qquad N \to \infty.
$$

Thus, $J^*(x) = \lim_{N\to\infty} F^N J(x)$ as required. □

This allows us to at least approximate $J^*(x)$. Moreover, an immediate consequence of Theorem 3.2 is that the Bellman equation is indeed a sufficient condition within the class of bounded functions, which takes care of our second concern.

Theorem 3.3. *$J^*(x)$ is the unique bounded solution to the Bellman equation.*

Proof (Bertsekas, 2007, p. 12). Suppose J is a bounded function that satisfies the Bellman equation. Then $J(x) = FJ(x)$ and it follows that $J(x) = \lim_{N\to\infty} F^N J(x)$. By Theorem 3.2, we have $J(x) = J^*(x)$. □

Having shown that there is just one bounded solution to the Bellman equation, we can state the necessary and sufficient condition for optimality, which in turn shows that we can solve the dynamic programming problem by solving the Bellman equation.

Theorem 3.4 (Necessary and sufficient condition for optimality)**.** *The policy $\pi^* = \{\pi_t^*\}_{t=0}^{\infty}$ is optimal and $u^* = \{u_t^*\}_{t=0}^{\infty}$ is an optimal control to the infinite-time dynamic programming problem if and only if u^* solves the Bellman equation for the value function J^*.*

Proof. Suppose (x^*, u^*) is an admissible pair that solves the Bellman equation (3.2) for the value function J^*:

$$
J^*(x_t^*) \;=\; \mathbb{E}\Big[g(x_t^*,u_t^*,z_t) + \beta J^*\big(f(x_t^*,u_t^*,z_t)\big)\Big] \qquad \text{for all } t = 0,1,2,\dots .
$$

By recursion over t, we then have

$$\begin{aligned} J^*(x_0^*) &= \mathbb{E}\left[g(x_0^*, u_0^*, z_0) + \beta\Big(g(x_1^*, u_1^*, z_1) + \beta J^*\big(f(x_1^*, u_1^*, z_1)\big)\Big)\right] \\ &= \dots \\ &= \mathbb{E}\left[\sum_{t=0}^{T} \beta^t g(x_t^*, u_t^*, z_t) + \beta^{T+1} J^*\big(f(x_T^*, u_T^*, z_T)\big)\right]. \end{aligned}$$

Since J^* is bounded by Assumption 3.1, $\beta^{T+1} J^*(x_{T+1}^*) \to 0$ as $T \to \infty$. It follows that

$$J^*(x_0^*) \;\to\; \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t g\big(x_t^*, u_t^*, z_t\big)\right] \qquad \text{as} \qquad T \to \infty,$$

so u^* solves the dynamic programming problem. Conversely, suppose that the admissible pair (x^*, u^*) solves the dynamic programming problem. We then know by Lemma 2.1 that $\{u_{t+s}^*\}_{s=0}^{\infty}$ solves the subproblem starting in period t with initial state x_t^*. Using Lemma 2.1 twice yields

$$\begin{aligned} J^*(x_t^*) &= \mathbb{E}\left[g(x_t^*, u_t^*, z_t) + \beta \sum_{s=0}^{\infty} \beta^s g(x_{t+s+1}^*, u_{t+s+1}^*, z_{t+s+1})\right] \\ &= \mathbb{E}\Big[g(x_t^*, u_t^*, z_t) + \beta J^*\big(f(x_t^*, u_t^*, z_t)\big)\Big]. \end{aligned}$$

Since the value function satisfies the Bellman equation by Theorem 3.1, it follows that u^* solves the Bellman equation. □

3.3 Solving the Bellman equation

The necessary and sufficient condition for optimality presents us with a method to solve the dynamic programming problem: by solving the Bellman equation. In principle this is easy. However, Theorem 3.4 only applies *if* we already know the value function J^*. As previously pointed out, this is generally not the case, and the key concern when solving these problems is thus to find the value function, We therefore finish off this section by discussing the methods used in this regard. In particular, Ljungqvist and Sargent (2012, ch. 3.1.1) list three main types of computational methods to find the value function:

(I). Guess and verify. The first method involves guessing (a bounded) J^* and verifying that it is indeed a solution to the Bellman equation (3.2). This method relies on Theorem 3.3; if we find a bounded function that satisfies the Bellman equation, then it has to be the value function. However, as this method depends on luck in making a good guess, it is of limited practical use.

(II). Value function iteration. A second method, value function iteration, applies the mapping F defined by Equation (3.3) to construct a sequence of value functions and corresponding controls by iteration as follows:

(i) Start with some bounded function $J_0 : X \to \mathbb{R}$, for instance the zero function $J_0(x) = 0$ for all $x \in X$.
(ii) At each iteration n, calculate $J_{n+1}(x) = FJ_n(x)$. That is, let

$$J_{n+1}(x) \;=\; \sup_{u \in U(x)} \mathbb{E}\Big[g(x,u,z) + \beta J_n\big(f(x,u,z)\big)\Big].$$

By Theorem 3.2, we know that $F^n J_0(x) \to J^*(x)$ as $n \to \infty$ for any bounded J_0, so the constructed sequence of value functions is guaranteed to converge to J^*.

(III). Policy function iteration. The last method, policy function iteration, is similar to value function iteration and uses the same mapping F, but it iterates over feasible policies instead of the value function. It consists of the following steps:

(i) Start with some feasible policy $\pi_0 : X \to U$ such that $\pi_0(x) \in U(x)$ for all $x \in X$.
(ii) Policy evaluation: at the start of each iteration n, calculate the value J_n of the chosen policy $\pi_n(x)$:

$$J_n(x) \;=\; \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t g\big(x_t, \pi_n(x_t), z_t\big)\right] \qquad \text{with } x_{t+1} = f\big(x_t, \pi_n(x_t), z_t\big) \text{ and } x_0 = x.$$

(iii) Policy improvement: generate a new policy $\pi_{n+1}(x) \in U(x)$ that satisfies the mapping F for J_n. That is, let

$$\pi_{n+1}(x) \;=\; \operatorname*{arg\,max}_{\pi(x) \in U(x)} \mathbb{E}\Big[g\big(x, \pi(x), z\big) + \beta J_n\Big(f\big(x, \pi(x), z\big)\Big)\Big].$$

This algorithm generates better and better policies: $J_{n+1} \geq J_n$ for all n. In the limit, $J_n \to J^*$ and the corresponding policy is optimal. To see this, we follow Voorneveld (2016, p. 149) and define the function $F_n : X \to X$ as

$$F_n J(x) \;=\; \mathbb{E}\Big[g\big(x, \pi_n(x), z\big) + \beta J\Big(f\big(x, \pi_n(x), z\big)\Big)\Big].$$

This is a monotonic function; suppose $J(x) \geq V(x)$, then

$$\begin{aligned} F_n J(x) \;&=\; \mathbb{E}\Big[g\big(x, \pi_n(x), z\big) + \beta J\Big(f\big(x, \pi_n(x), z\big)\Big)\Big] \\ &\geq\; \mathbb{E}\Big[g\big(x, \pi_n(x), z\big) + \beta V\Big(f\big(x, \pi_n(x), z\big)\Big)\Big] \;=\; F_n V(x). \end{aligned}$$

After the policy evaluation and policy improvement we then have, respectively,

$$\begin{aligned} F_n J_n(x) \;&=\; \mathbb{E}\Big[g\big(x, \pi_n(x), z\big) + \beta J_n\Big(f\big(x, \pi_n(x), z\big)\Big)\Big] \;=\; J_n(x), \\ F_{n+1} J_n(x) \;&=\; \mathbb{E}\Big[g\big(x, \pi_n(x), z\big) + \beta J_n\Big(f\big(x, \pi_{n+1}(x), z\big)\Big)\Big]. \end{aligned}$$

Since $\pi_{n+1}(x)$ is chosen optimally, we necessarily have that $F_{n+1}J_n(x) \geq F_nJ_n(x)$. Monotonicity of F_{n+1} implies $F_{n+1}^kJ_n(x) \geq F_nJ_n(x)$ for all k, where F_{n+1}^k is the composition of F_{n+1} with itself k times. By recursion over k, we can write F_{n+1}^k as

$$F_{n+1}^k J_n(x) = \mathbb{E}\left[\sum_{t=0}^{k-1} g\big(x_t, \pi_{n+1}(x_t), z_t\big) + \beta^k J_n(x_k)\right],$$

and we therefore have $F_{n+1}^kJ_n(x) \to J_{n+1}(x)$ in the limit. This shows that $J_{n+1} \geq J_n$. By Assumption 3.1, J_n is bounded, and in particular, it is bounded from above by J^* due to Definition 3.2. The algorithm therefore constructs a sequence $\{J_n\}_{n=0}^{\infty}$ that is increasing and bounded from above: it must converge. In the limit, $\pi_n(x) = \pi_{n+1}(x)$ and so

$$\begin{aligned}
J_n(x) &= \mathbb{E}\left[g\big(x, \pi_n(x), z\big) + \beta J_n\Big(f\big(x, \pi_n(x), z\big)\Big)\right] \\
&= \mathbb{E}\left[g\big(x, \pi_{n+1}(x), z\big) + \beta J_n\Big(f\big(x, \pi_{n+1}(x), z\big)\Big)\right] \\
&= \sup_{\pi(x)\in U(x)} \mathbb{E}\left[g\big(x, \pi(x), z\big) + \beta J_n\Big(f\big(x, \pi(x), z\big)\Big)\right].
\end{aligned}$$

Thus, J_n is a bounded function that satisfies the Bellman equation. By Theorem 3.3 it follows that $J_n = J^*$.

4. Applications in economics

Having laid out the necessary framework for dynamic programming in finite and infinite time, we now motivate the use of the theory by applying it to a number of problems concerning economic theory. Below we cover four examples. In the first two, we use dynamic programming simply as a tool to derive two classic results in economics: Milton Friedman's permanent income hypothesis and James Tobin's q-theory of investment.[1] In the last two examples we turn more practical and use dynamic programming to find explicit solutions to two standard problems: McCall's (1970) job search model, where an unemployed worker must decide on the optimal timing to accept a job offer, and Brock and Mirman's (1972) stochastic growth model, where a benevolent social planner must decide on the welfare-maximising allocation of consumption, investment, and labour supply.

4.1 The permanent income hypothesis

The permanent income hypothesis concerns the issue of dividing consumption between the present and the future for a household facing an uncertain income stream. The hypothesis was first discussed by Fisher (1930) and Friedman (1957), while Hall (1978) formalised it mathematically. According to the hypothesis, households smooth consumption across time by forming expectations of their total lifetime income (or permanent income) and then setting current consumption as an appropriate fraction of that income. We illustrate this hypothesis with the use of dynamic programming below.

Consider an infinitely-lived household that gets utility from consumption c according to a function $u(c)$ which is strictly increasing, continuously differentiable, and strictly concave. We naturally restrict consumption to be non-negative: $c \geq 0$. In each period t, the household earns a random wage w_t which is independent over time and drawn from the set $\mathcal{W} = \{w_1, \ldots, w_N\}$. The household wishes to maximise the expectation of its discounted lifetime utility given by

$$\mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t u(c_t) \right],$$

where $\beta \in (0,1)$ is a subjective discount factor and $\mathbb{E}_0$ denotes the expectation over w_t, conditioned on the information available in time 0. The expectation is necessary here because future values of consumption are stochastic as they depend on the wage realisations w_t. Lastly, the household also has the opportunity to lend and borrow assets

[1] Both of which received the Nobel Memorial Prize in Economic Sciences partly due to these results.

a_t freely at some constant interest rate r. For all t, this yields the per-period budget constraint

$$a_{t+1} + c_t \;=\; (1+r)a_t + w_t. \tag{4.1}$$

Starting in period t and using recursion forward, we can rewrite Equation (4.1) as

$$\sum_{s=0}^{\infty}\left(\frac{1}{1+r}\right)^s c_{t+s} \;=\; \sum_{s=0}^{\infty}\left(\frac{1}{1+r}\right)^s w_{t+s} + a_t,$$

and since $c_t \geq 0$ and $w_t \geq \min \mathcal{W}$ for all t, we therefore have the borrowing constraint

$$a_t \;\geq\; -\sum_{s=0}^{\infty}\left(\frac{1}{1+r}\right)^s \min \mathcal{W} \;\equiv\; -b.$$

It follows from Equation (4.1) that $a_{t+1} \in \big[-b,\, (1+r)a_t + w_t\big]$ and $c_t \in \big[0,\, (1+r)a_t + w_t + b\big]$. Hence, we are optimising a continuous function over a non-empty, compact, and convex set, so a maximum exists by Weierstrass' maximum theorem. By strict concavity of $u(\cdot)$, this maximum is unique. For a_0 and w_0 given, we can therefore write the household optimisation problem as

$$\begin{aligned}
&\max_{\{c_t\}_{t=0}^{\infty}} && \mathbb{E}_0\left[\sum_{t=0}^{\infty}\beta^t u(c_t)\right] \\
&\text{subject to} && c_t \in \big[0,\, (1+r)a_t + w_t + b\big], && t = 0,1,2,\ldots, \\
& && w_t \in \mathcal{W}, && t = 0,1,2,\ldots, \\
& && a_{t+1} = (1+r)a_t + w_t - c_t, && t = 0,1,2,\ldots, \\
& && a_0 \text{ and } w_0 \text{ given.}
\end{aligned}$$

Conversely, by Equation (4.1) we can write $c_t = (1+r)a_t + w_t - a_{t+1}$, substitute for c_t in $u(c_t)$, and maximise over a_{t+1} instead of c_t. It turns out that this approach is easier here. We thus write the Bellman equation for this problem as

$$J^*(a_t, w_t) \;=\; \max_{a_{t+1}\in[-b,(1+r)a_t+w_t]} \mathbb{E}_t\Big[u\Big((1+r)a_t + w_t - a_{t+1}\Big) + \beta J^*(a_{t+1}, w_{t+1})\Big].$$

Differentiating the right-hand side of the Bellman equation, the first-order condition for an internal solution of the maximisation problem gives

$$-u'\Big((1+r)a_t + w_t - a_{t+1}\Big) + \beta\mathbb{E}_t\left[\frac{\partial J^*(a_{t+1}, w_{t+1})}{\partial a_{t+1}}\right] \;=\; 0,$$

where u' denotes the derivative of u. Now, by definition we have

$$J^*(a_t, w_t) \;=\; \max_{\{a_{t+s+1}\}_{s=0}^{\infty}} \mathbb{E}_t\left[\sum_{s=0}^{\infty}\beta^s u\Big((1+r)a_{t+s} + w_{t+s} - a_{t+s+1}\Big)\right],$$

so

$$\frac{\partial J^*(a_t, w_t)}{\partial a_t} \;=\; (1+r)u'\Big((1+r)a_t + w_t - a_{t+1}\Big).$$

Substituting this into the first-order condition and rearranging terms, we get the necessary optimality condition

$$u'\Big((1+r)a_t + w_t - a_{t+1}\Big) \;=\; \beta(1+r)\mathbb{E}_t\Big[u'\Big((1+r)a_{t+1} + w_{t+1} - a_{t+2}\Big)\Big]. \qquad (4.2)$$

We can solve Equation (4.2) using policy function iteration: conjecture the policy $a_{t+2} = \pi_0(a_{t+1})$ such that for a_t and w_t given, Equation (4.2) becomes a function in a_{t+1} only. Solving for a_{t+1} yields a solution to the right-hand side of the Bellman equation, so this solution gives a policy update $a_{t+1} = \pi_1(a_t)$. Iterating until convergence yields the optimal policy. Moreover, by the budget constraint (4.1), we can write Equation (4.2) as

$$u'(c_t) \;=\; \beta(1+r)\mathbb{E}_t\Big[u'(c_{t+1})\Big]. \qquad (4.3)$$

This equation, called the consumption Euler equation, is *the* cornerstone of modern macroeconomics[2] and captures precisely the permanent income hypothesis: consumption today is not only based on current income but on the expected income and consumption in future years, and households smooth consumption across time accordingly.

4.2 Optimal investment and Tobin's q

The "q-theory of investment" is a canonical model of firm investment first introduced by Tobin (1969), where q is the ratio of the average market value of a firm's capital to its replacement cost. Tobin argues that investment is positively related to this q: if q is greater than one, capital has more value within the firm than outside it, so the firm should invest in more capital, and vice versa if q is less than one. Hayashi (1980) later established that investment in fact should relate to a marginal q. That is, q is the ratio of the market value of *new additional* capital to its replacement cost. It turns out that we can derive Tobin's marginal q using dynamic programming.

Consider a firm which uses capital K to earn revenue according to some function $\pi(K)$ which is strictly increasing, continuously differentiable, strictly concave, and bounded from above. We naturally restrict capital to be non-negative: $K \geq 0$. In each period, there is a random demand shock z which is independent over time, drawn from the set $\mathcal{Z} = \{z_1, \ldots, z_N\}$, and adjusts revenues π linearly. The firm can choose to invest some amount I in new capital each period at a constant price p. In addition to the direct cost pI of investment, investment is also subject to adjustment costs captured by the function $\Phi(I)$, where Φ is strictly increasing, continuously differentiable, and strictly convex. We additionally assume that $\Phi(0) = \Phi'(0) = 0$, where Φ' denotes the derivative of Φ. The firm seeks to maximise the expectation of discounted total profits given by

$$\mathbb{E}_0\left[\sum_{t=0}^{\infty}\left(\frac{1}{1+r}\right)^t \Big(z_t\pi(K_t) - pI_t - \Phi(I_t)\Big)\right],$$

where r is a constant interest rate and $\mathbb{E}_0$ denotes the expectation over z_t, conditioned on the information available in time 0. Lastly, capital is assumed to depreciate at a rate

[2] Indeed, Ljungqvist and Sargent (2012) call it "the common ancestor" of all macroeconomics.

$\delta \in (0,1)$ such that $K_{t+1} = (1-\delta)K_t + I_t$. By non-negativity of K_t, we then necessarily have $I_t \in [-(1-\delta)K_t, +\infty)$. For K_0 and z_0 given, this yields the optimisation problem

$$\begin{aligned} \max_{\{I_t\}_{t=0}^{\infty}} \quad & \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \left(\frac{1}{1+r} \right)^t \Big(z_t \pi(K_t) - pI_t - \Phi(I_t) \Big) \right] \\ \text{subject to} \quad & I_t \in \big[-(1-\delta)K_t,\ +\infty \big), && t = 0,1,2,\ldots, \\ & z_t \in \mathcal{Z}, && t = 0,1,2,\ldots, \\ & K_{t+1} = (1-\delta)K_t + I_t, && t = 0,1,2,\ldots, \\ & K_0 \text{ and } z_0 \text{ given.} \end{aligned}$$

As in the case of the permanent income hypothesis, we can write $I_t = K_{t+1} - (1-\delta)K_t$, substitute for I_t in the profit function, and optimise over K_{t+1} instead of I_t. Per-period profits are then $z_t\pi(K_t) - p\big(K_{t+1} - (1-\delta)K_t\big) - \Phi\big(K_{t+1} - (1-\delta)K_t\big)$. Moreover, by boundedness of π we necessarily have that $\overline{K} \equiv \arg\max_{K \geq 0} \big(\pi(K) - p\delta K\big)$ is finite. For any $K_t \in [0, \overline{K}]$, it can never be optimal to choose $K_{t+1} > \overline{K}$ since by reducing capital to $K_{t+1} = \overline{K}$, the firm attains a higher expected revenue tomorrow at a lower investment cost today. Thus, without loss of generality we can assume that $K_{t+1} \in [0, \overline{K}]$. Again, we are maximising a continuous and strictly concave function over a non-empty, compact, and convex set, so a unique maximum exists. It follows that the Bellman equation is

$$\begin{aligned} J^*(K_t, z_t) \;=\; \max_{K_{t+1} \in [0, \overline{K}]} \; \mathbb{E}_t \bigg[& z_t \pi(K_t) - p\big(K_{t+1} - (1-\delta)K_t\big) - \Phi\Big(K_{t+1} - (1-\delta)K_t\Big) \\ & + \frac{1}{1+r} J^*(K_{t+1}, z_{t+1}) \bigg]. \end{aligned}$$

The solution procedure is completely analogous to the case of the permanent income hypothesis. The first-order condition for an internal solution of the right-hand side is

$$-p - \Phi'\big(K_{t+1} - (1-\delta)K_t\big) + \frac{1}{1+r} \mathbb{E}_t \left[\frac{\partial J^*(K_{t+1}, z_{t+1})}{\partial K_{t+1}} \right] \;=\; 0.$$

By the definition of the value function,

$$\frac{\partial J^*(K_t, z_t)}{\partial K_t} \;=\; z_t \pi'(K_t) + (1-\delta)\Big(p + \Phi'\big(K_{t+1} - (1-\delta)K_t\big)\Big),$$

where π' denotes the derivative of π. Substituting this into the first-order condition, rearranging terms, and using $I_t = K_{t+1} - (1-\delta)K_t$, we get the optimality condition

$$p + \Phi'\big(I_t\big) \;=\; \frac{1}{1+r}\, \mathbb{E}_t \Big[z_{t+1} \pi'(K_{t+1}) + (1-\delta)\Big(p + \Phi'\big(I_{t+1}\big)\Big) \Big]. \tag{4.4}$$

Again, this can be solved using policy function iteration to find the optimal investment. Moreover, if we define

$$q_t \;\equiv\; \frac{1}{p} \left(\frac{1}{1+r}\, \mathbb{E}_t \left[\frac{\partial J^*(K_{t+1}, z_{t+1})}{\partial K_{t+1}} \right] \right)$$

and substitute Equation (4.4) into the right-hand side of $\frac{\partial J^*(K_{t+1},z_{t+1})}{\partial K_{t+1}}$, we get recursively

$$q_t = \frac{1}{p}\left(\frac{1}{1+r}\mathbb{E}_t\left[\sum_{s=0}^{\infty}\left(\frac{1-\delta}{1+r}\right)^s z_{t+s+1}\pi'(K_{t+s+1})\right]\right). \tag{4.5}$$

Since p is the price of investment, it is also the marginal replacement cost of capital. Thus, Equation (4.5) states that q_t is the expected increase in future (discounted) profits that results from increasing the capital stock by one unit, relative to the replacement cost of that unit of capital. In other words, we have derived Tobin's q. Moreover, by definition of q_t and by the first-order condition of the Bellman equation, $p+\Phi'(I_t) = pq_t$. By strict convexity of Φ with $\Phi'(0) = 0$, there exists an inverse function Φ'^{-1} which is strictly increasing with $\Phi'^{-1}(0) = 0$. It follows that $I_t = \Phi'^{-1}\big[p(q_t - 1)\big]$ with positive investment if $q_t > 1$ and negative investment if $q_t < 1$, in line with Tobin's prediction.

4.3 Job search and unemployment

Another type of problem that is well-suited for dynamic programming techniques is optimal stopping problems. The labour economics literature has taken advantage of this since theoretical models of unemployment and job search are typically of this kind. For instance, consider the following modified version of McCall's (1970) job search model.

An individual enters the labour market as unemployed at time $t = 0$ and retires after period T. Each period, the individual receives a job offer with a randomly drawn wage w_t that is independent over time and drawn from the set $\mathcal{W} = \{w_1, \dots, w_N\}$. The individual has the option to either reject or accept the offer. If the offer is rejected, he or she receives some unemployment compensation c in this period and waits until the next period to draw a new offer from $\mathcal{W}$. If the offer is accepted, he or she receives this wage in every subsequent period until retirement. Thus, if y_t is the income received in period t, then $y_t = c$ if the offer is rejected and $y_t = w_t$ if it is accepted. The individual wishes to maximise the expected sum of discounted income $\mathbb{E}_0 \sum_{t=0}^{T} \beta^t y_t$.

Suppose we let the choice of accepting or rejecting the offer be denoted by a binary variable u_t equal to one if the offer is accepted and zero otherwise. Then all variables lie in finite sets so we know a solution to the maximisation problem exists and for w_0 given we can state this problem as

$$\max_{\{u_t\}_{t=0}^{T}} \mathbb{E}_0\left[\sum_{t=0}^{T}\beta^t y_t\right] \quad \text{subject to} \quad u_t \in \{0,1\},\ w_t \in \mathcal{W} \text{ with } w_0 \text{ given, and}$$

$$y_t = \begin{cases} w_s & \text{if } u_s = 1 \text{ and } u_0 = \ldots = u_{s-1} = 0, \\ c & \text{otherwise,} \end{cases}$$

where $t \in \{0, \dots, T\}$, $s \in \{0, \dots, t\}$, and $\mathbb{E}_0$ denotes the expectation over w_t, conditioned on the information available in time 0. Clearly, this is of the same form as Definition 2.1, so we can apply the dynamic programming algorithm. Since a worker who accepts an offer receives that wage in every subsequent period, the value from accepting an offer w_t

in period t is

$$\sum_{s=0}^{T-t} \beta^s w_t \;=\; w_t \cdot \frac{\beta^{T-t+1}-1}{\beta-1},$$

while the value from rejecting is $c + \mathbb{E}_t\big[J^*_{t+1}(w_{t+1})\big]$. Thus, by Theorem 2.2,

$$J^*_T(w_T) \;=\; \max_{u_T} \Big\{ u_T w_T + (1-u_T)c \Big\} \;=\; \max\big\{ w_T,\, c \big\},$$

and for all $t \in \{0, \dots, T-1\}$,

$$\begin{aligned} J^*_t(w_t) \;&=\; \max_{u_t} \left\{ u_t w_t \cdot \frac{\beta^{T-t+1}-1}{\beta-1} + (1-u_t)\Big(c + \mathbb{E}_t\Big[J^*_{t+1}(w_{t+1})\Big]\Big) \right\} \\ &=\; \max \left\{ w_t \cdot \frac{\beta^{T-t+1}-1}{\beta-1},\; c + \mathbb{E}_t\Big[J^*_{t+1}(w_{t+1})\Big] \right\}. \end{aligned}$$

Clearly, the optimal control is to accept the offer if the first term in the maximum operator is greater than the second term. Denote the *reservation wage* (that is, the wage at which the worker is indifferent between accepting and rejecting the offer) by $\overline{w}_t$. We then necessarily have $\overline{w}_T = c$ and, if q_i denotes the probability that w_i is drawn from $\mathcal{W}$ (naturally with $\sum_i q_i = 1$),

$$\overline{w}_t \cdot \frac{\beta^{T-t+1}-1}{\beta-1} \;=\; c + \sum_{i=1}^{N} q_i J^*_{t+1}(w_i) \;=\; c + \frac{\beta^{T-t}-1}{\beta-1} \left(\sum_{w_i \le \overline{w}_{t+1}} q_i \overline{w}_{t+1} + \sum_{w_i > \overline{w}_{t+1}} q_i w_i \right)$$

for $t \in \{0, \dots, T-1\}$. It follows that the individual's reservation wage is given recursively by

$$\overline{w}_t \;=\; \begin{cases} c & \text{if } t = T, \\ \dfrac{\beta-1}{\beta^{T-t+1}-1} \left[c + \dfrac{\beta^{T-t}-1}{\beta-1} \left(\displaystyle\sum_{w_i \le \overline{w}_{t+1}} q_i \overline{w}_{t+1} + \sum_{w_i > \overline{w}_{t+1}} q_i w_i \right) \right] & \text{if } t < T, \end{cases} \tag{4.6}$$

and the optimal control is

$$u^*_t \;=\; \begin{cases} 1 & \text{if } \overline{w}_t < w_t, \\ 0 & \text{if } \overline{w}_t \ge w_t. \end{cases}$$

The formula for the reservation wage provides an easy way to solve numerical examples. Suppose for instance that $T = 20$, $N = 10$, $w_i = i$, $q_i = \frac{1}{N}$, $c = 3$, and $\beta = 0.99$. Equation (4.6) then gives a decreasing reservation wage over time which is plotted in Figure 4.1.

4.4 Optimal economic growth

In general, there is no closed-form solution to the value function, and in applications we rather use approximate computational methods to solve a problem. In rare instances,

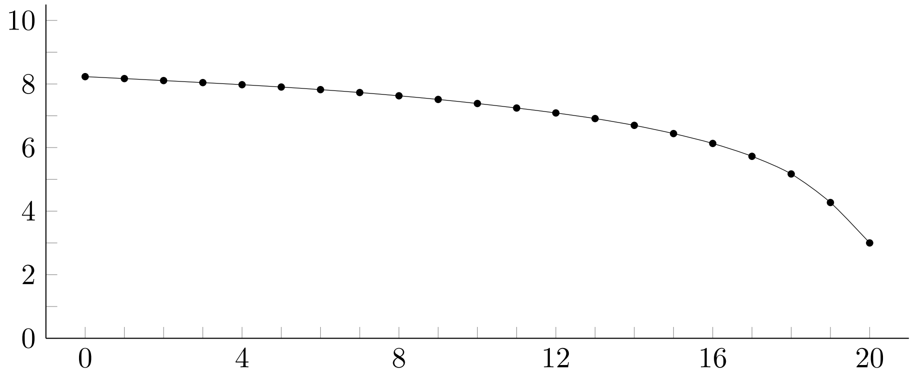

FIGURE 4.1. The reservation wage for the numerical example of the job search model

however, it is possible to solve for the value function analytically. This is the case for the stochastic model of economic growth by Brock and Mirman (1972). Their model concerns infinitely-lived households that only care about consumption, just as in the case of the permanent income hypothesis. Below we extend their model to include a labour supply choice and show that there is still an analytical solution for the value function.

Consider a closed economy without a government that consists of a household sector and a production sector. The production sector produces output through an aggregated Cobb-Douglas production function[3] given by

$$Y_t = z_t K_t^{\alpha} L_t^{1-\alpha},$$

where K_t is the capital stock, z_t is the technology level, and $\alpha \in (0, 1)$ is the capital share of national income. The technology level is randomly drawn from the set $\mathcal{Z} = \{z_1, \ldots, z_N\}$ and is independent over time. The number of households is normalised to one, and households only care about consumption and leisure time. We naturally restrict consumption to be non-negative and normalise the time endowment to one. The average household maximises the expectation of its discounted lifetime utility function given by

$$\mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \Big(\ln C_t + \psi \ln \big(1 - L_t\big) \Big) \right],$$

where C_t denotes consumption, L_t denotes labour supply, $\mathbb{E}_0$ denotes the expectation over z_t conditioned on the information available in time 0, $\beta \in (0, 1)$ is a subjective discount factor, and ψ is a parameter denoting the relative weight put on leisure. Clearly, $C_t \geq 0$ and $L_t \in [0, 1]$. With neither a government nor access to international markets, the standard national income identity requires that output be divided between consumption and investment I_t, so $Y_t = C_t + I_t$. Capital is assumed to depreciate fully after one period, so the next-period capital stock is simply the amount invested today: $K_{t+1} = I_t$. We therefore have that $Y_t = z_t K_t^{\alpha} L_t^{1-\alpha} = C_t + K_{t+1}$, which motivates the presence of the expectations operator in the utility function. We wish to find the optimal allocation of consumption, investment, and labour supply from the perspective of the households.

[3]Introduced by Cobb and Douglas (1928), this is the standard production function used in most macroeconomic models as it aligns well with observed long-run data patterns.

For K_0 and z_0 given, this is equivalent to solving

$$\begin{aligned}
&\max_{\{C_t,L_t\}_{t=0}^{\infty}} && \mathbb{E}_0\left[\sum_{t=0}^{\infty}\beta^t\Big(\ln C_t+\psi\ln\big(1-L_t\big)\Big)\right] \\
&\text{subject to} && C_t\in\big[0,\ z_tK_t^{\alpha}L_t^{1-\alpha}\big], && t=0,1,2,\ldots, \\
& && L_t\in[0,1], && t=0,1,2,\ldots, \\
& && z_t\in\mathcal{Z}, && t=0,1,2,\ldots, \\
& && K_{t+1}=z_tK_t^{\alpha}L_t^{1-\alpha}-C_t, && t=0,1,2,\ldots, \\
& && K_0 \text{ and } z_0 \text{ given.}
\end{aligned}$$

Again we are optimising a continuous and strictly concave function over a non-empty, compact, and convex set, so a unique solution exists. As in the previous examples, we use that $C_t = z_tK_t^{\alpha}L_t^{1-\alpha}-K_{t+1}$, substitute for C_t in the objective function, and optimise over K_{t+1} instead of C_t. The Bellman equation is then

$$J^*(K_t,z_t) \;=\; \max_{\substack{K_{t+1}\in[0,Y_t]\\ L_t\in[0,1]}} \mathbb{E}_t\Big[\ln\big(z_tK_t^{\alpha}L_t^{1-\alpha}-K_{t+1}\big)+\psi\ln\big(1-L_t\big)+\beta J^*\big(K_{t+1},z_{t+1}\big)\Big].$$

We can apply value function iteration to this equation. Starting with $J_0=0$, we show in Appendix B that

$$J^*(K_t,z_t) \;=\; X+Y\ln K_t+Z\ln z_t,$$

where

$$\begin{aligned}
X &= \frac{1}{1-\beta}\bigg[\psi\ln\big(1-L\big)+\frac{\alpha\beta}{1-\alpha\beta}\ln\Big(\alpha\beta L^{1-\alpha}\Big)+\ln\Big((1-\alpha\beta)L^{1-\alpha}\Big) \\
&\qquad +\frac{\beta}{1-\alpha\beta}\mathbb{E}_t\Big[\ln z_{t+1}\Big]\bigg], \\
Y &= \frac{\alpha}{1-\alpha\beta}, \\
Z &= \frac{1}{1-\alpha\beta}, \\
L &= \frac{1-\alpha}{\psi(1-\alpha\beta)+(1-\alpha)}.
\end{aligned}$$

It follows that $\frac{\partial J^*(K_{t+1},z_{t+1})}{\partial K_{t+1}}=\frac{Y}{K_{t+1}}$, so the first-order conditions for an internal solution of the right-hand side of the Bellman equation are

$$K_{t+1}: \qquad -\frac{1}{z_tK_t^{\alpha}L_t^{1-\alpha}-K_{t+1}}+\frac{\beta Y}{K_{t+1}} \;=\; 0, \tag{4.7}$$

$$L_t: \qquad \frac{(1-\alpha)z_tK_t^{\alpha}L_t^{-\alpha}}{z_tK_t^{\alpha}L_t^{1-\alpha}-K_{t+1}}-\frac{\psi}{1-L_t} \;=\; 0. \tag{4.8}$$

By Equation (4.7), $K_{t+1} = \frac{\beta Y}{1+\beta Y} z_t K_t^{\alpha} L_t^{1-\alpha}$. Plugging this into Equation (4.8), solving for L_t, and using the definition of Y, we get

$$L_t = \frac{(1-\alpha)(1+\beta Y)}{\psi + (1-\alpha)(1+\beta Y)} = \frac{1-\alpha}{\psi(1-\alpha) + (1-\alpha)} = L.$$

Given L and Y, the optimal policies are therefore

$$\begin{aligned} L_t &= L, \\ I_t &= K_{t+1} = \frac{\beta Y}{1+\beta Y} z_t K_t^{\alpha} L_t^{1-\alpha} = \alpha\beta z_t K_t^{\alpha} L^{1-\alpha}, \\ C_t &= z_t K_t^{\alpha} L_t^{1-\alpha} - K_{t+1} = (1-\alpha\beta) z_t K_t^{\alpha} L^{1-\alpha}. \end{aligned}$$

Thus, in each period t, the optimal labour supply is constant while optimal consumption and investment are constant fractions of aggregate output.

5. Concluding summary

Concerning solution methods to discrete-time dynamic optimisation problems, this paper outlines the theory of dynamic programming. In addition to covering a rather general theory for problems with a finite number of stages, we have also considered a special case of infinite-time problems, namely problems with discounted and stationary objective functions. We saw that finite-time problems can be solved using backwards recursion, where the subproblem starting in the last stage is solved first and then used to solve subsequent subproblems starting in earlier stages, and that infinite-time problems can be solved using the Bellman equation.

A second aim of this thesis was to explore how dynamic programming can be applied to problems concerning economic theory. To this end, we used dynamic programming to derive two cornerstones of modern economics: the permanent income hypothesis and Tobin's q. We also used dynamic programming to explicitly solve two economic models (one regarding unemployment and one regarding macroeconomic growth) in order to show how the theory can be applied in practice.

From the two aims of this thesis – presenting the basic theory and applying it to economics – we have also illustrated one of the main advantages of dynamic programming. Optimisation problems in economic models typically fall within nonlinear programming, where problems are solved using Lagrange multipliers and the Karush-Kuhn-Tucker conditions. These conditions form equation systems where the entire sequence of optimal controls and Lagrange multipliers are solved simultaneously for all possible states. In principle, this is a viable strategy if the dimensionality of the problem is small, but becomes infeasible as we let the number of dimensions approach infinity. Dynamic programming on the other hand circumvents this issue by reducing the problem into a collection of easily solved subproblems.

Lastly, it is important to emphasise that the material covered here disregards a number of challenging technical issues. First, we do not consider more general forms of infinite-time problems, such as those with undiscounted or non-stationary objective functions, which are typically harder to solve. However, this restriction was motivated by the fact that most infinite-time problems in economics are discounted and stationary. Secondly, and more importantly, we have kept the treatment of stochastic elements to a minimum by assuming that they are independent over time and drawn from finite sets. This excludes many types of problems that occur in economics, where random variables often follow Markov processes or are drawn from continuous distributions. Extending the theory to account for this is possible but requires more sophisticated methods (such as measure theory), and thus goes beyond the scope of this paper.

References

Bellman, R. E. (1952). 'On the theory of dynamic programming'. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 38(8), pp. 716–719.

Bellman, R. E. (1953). *An introduction to the theory of dynamic programming* (RAND Corporation report No. R-245). Santa Monica, CA: RAND Corporation.

Bertsekas, D. P. (2005). *Dynamic programming and optimal control, vol. 1* (3rd ed.). Belmont, MA: Athena Scientific.

Bertsekas, D. P. (2007). *Dynamic programming and optimal control, vol. 2* (3rd ed.). Belmont, MA: Athena Scientific.

Brock, W. A. and Mirman, L. J. (1972). 'Optimal economic growth and uncertainty: the discounted case'. *Journal of Economic Theory*, vol. 4(3), pp. 479–513.

Cobb, C. W. and Douglas, P. H. (1928). 'A theory of production'. *American Economic Review*, vol. 18(1), pp. 139–165.

Fisher, I. (1930). *The theory of interest, as determined by impatience to spend income and opportunity to invest it*. New York, NY: Macmillan.

Friedman, M. (1957). *A theory of the consumption function*. Princeton, NJ: Princeton University Press.

Hall, R. E. (1978). 'Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence'. *Journal of Political Economy*, vol. 86(6), pp. 971–987.

Hayashi, F. (1980). 'Tobin's marginal and average q: a neoclassical interpretation'. *Econometrica*, vol. 50(1), pp. 213–224.

Ljungqvist, L. and Sargent, T. J. (2012). *Recursive macroeconomic theory* (3rd ed.). Cambridge, MA: The MIT Press.

McCall, J. J. (1970). 'Economics of information and job search'. *Quarterly Journal of Economics*, vol. 84(1), pp. 113–126.

Tobin, J. (1969). 'A general equilibrium approach to monetary theory'. *Journal of Money, Credit and Banking*, vol. 1(1), pp. 15–29.

Voorneveld, M. (2016). *A short course in static and discrete-time dynamic optimization*. Stockholm: Stockholm School of Economics. (Lecture notes, unpublished manuscript).

Appendix A

Proof of interchangeability between limit and expectation

We wish to show that under Assumptions 2.1 and 3.1,

$$\lim_{T\to\infty} \mathbb{E}\left[\sum_{t=0}^{T} \beta^t g\big(x_t, u_t, z_t\big)\right] = \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t g\big(x_t, u_t, z_t\big)\right].$$

Proof (derived independently). By linearity of the expectations operator,

$$\mathbb{E}\left[\sum_{t=0}^{T} \beta^t g\big(x_t, u_t, z_t\big)\right] = \sum_{t=0}^{T} \beta^t \,\mathbb{E}\Big[g(x_t, u_t, z_t)\Big],$$

and by Assumption 2.1 and the definition of $\mathbb{E}$ in Definition 2.1,

$$\mathbb{E}\Big[g(x_t, u_t, z_t)\Big] = \sum_{i=1}^{N} p_i\, g(x_t, u_t, z_i) \qquad \text{for all} \qquad t = 0, 1, 2, \ldots .$$

It follows that

$$\begin{aligned}\mathbb{E}\left[\sum_{t=0}^{T} \beta^t g\big(x_t, u_t, z_t\big)\right] &= \sum_{t=0}^{T} \beta^t \left(\sum_{i=1}^{N} p_i\, g(x_t, u_t, z_i)\right) \\ &= \sum_{i=1}^{N} p_i \left(\sum_{t=0}^{T} \beta^t g(x_t, u_t, z_i)\right). \end{aligned} \tag{A.1}$$

By Assumption 3.1, we know that

$$\lim_{T\to\infty} \sum_{t=0}^{T} \beta^t g(x_t, u_t, z_i) = \sum_{t=0}^{\infty} \beta^t g(x_t, u_t, z_i)$$

for a given z_i is well-defined. Thus, taking the limit of both sides of Equation (A.1) we get

$$
\begin{aligned}
\lim_{T\to\infty} \mathbb{E}\left[\sum_{t=0}^{T} \beta^t g\big(x_t, u_t, z_t\big)\right] &= \lim_{T\to\infty} \sum_{i=1}^{N} p_i \left(\sum_{t=0}^{T} \beta^t g(x_t, u_t, z_i)\right) \\
&= \sum_{i=1}^{N} p_i \left(\lim_{T\to\infty} \sum_{t=0}^{T} \beta^t g(x_t, u_t, z_i)\right) \\
&= \sum_{i=1}^{N} p_i \left(\sum_{t=0}^{\infty} \beta^t g(x_t, u_t, z_i)\right) \\
&= \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t g\big(x_t, u_t, z_t\big)\right],
\end{aligned}
$$

where the last equality follows from the definition of $\mathbb{E}$ in Definition 2.1. This proves the claim. □

Appendix B

Derivation of the optimal-growth value function

In Section 4.4 we wish to find the value function satisfying the Bellman equation

$$J^*(K_t, z_t) = \max_{\substack{K_{t+1}\in[0,Y_t] \\ L_t\in[0,1]}} \mathbb{E}_t\Big[\ln\big(z_t K_t^\alpha L_t^{1-\alpha} - K_{t+1}\big) + \psi \ln\big(1 - L_t\big) + \beta J^*\big(K_{t+1}, z_{t+1}\big)\Big].$$

Let $\{J_n\}_{n=0}^\infty$ be the sequence of value functions obtained from value function iteration, starting with $J_0 = 0$. We first show, using induction, that all these functions have the format

$$J_n(K_t, z_t) = X_n + Y_n \ln K_t + Z_n \ln z_t,$$

where X_n, Y_n, Z_n are some real scalars. The base case trivially holds with $X_0 = Y_0 = Z_0 = 0$. For the induction step, suppose the result is true for some $n \geq 0$. Then by the value function iteration algorithm,

$$\begin{aligned} J_{n+1}(K_t, z_t) = \max_{\substack{K_{t+1}\in[0,Y_t] \\ L_t\in[0,1]}} \mathbb{E}_t\Big[&\ln\big(z_t K_t^\alpha L_t^{1-\alpha} - K_{t+1}\big) + \psi \ln\big(1 - L_t\big) \\ &+ \beta\Big(X_n + Y_n \ln K_{t+1} + Z_n \ln z_{t+1}\Big)\Big]. \end{aligned}$$

The first-order conditions for an internal solution of the right-hand side are

$$K_{t+1}: \quad -\frac{1}{z_t K_t^\alpha L_t^{1-\alpha} - K_{t+1}} + \frac{\beta Y_n}{K_{t+1}} = 0, \tag{B.1}$$

$$L_t: \quad \frac{(1-\alpha) z_t K_t^\alpha L_t^{-\alpha}}{z_t K_t^\alpha L_t^{1-\alpha} - K_{t+1}} - \frac{\psi}{1 - L_t} = 0. \tag{B.2}$$

By Equation (B.1), $K_{t+1} = \frac{\beta Y_n}{1+\beta Y_n} z_t K_t^\alpha L_t^{1-\alpha}$. Plugging this into Equation (B.2) and solving for L_t we get

$$L_t = \frac{(1-\alpha)(1+\beta Y_n)}{\psi + (1-\alpha)(1+\beta Y_n)} \equiv L_n,$$

where L_n is a constant. Thus, $K_{t+1} = \frac{\beta Y_n}{1+\beta Y_n} z_t K_t^\alpha L_n^{1-\alpha}$. Using these results in the Bellman equation, we obtain

$$\begin{aligned}
J_{n+1}(K_t, z_t) &= \ln\left[\frac{1}{1+\beta Y_n} z_t K_t^\alpha L_n^{1-\alpha}\right] + \psi \ln\big(1 - L_n\big) + \beta X_n \\
&\quad + \beta Y_n \ln\left[\frac{\beta Y_n}{1+\beta Y_n} z_t K_t^\alpha L_n^{1-\alpha}\right] + \beta Z_n \mathbb{E}_t\Big[\ln z_{t+1}\Big] \\
&= \beta X_n + \psi \ln\big(1 - L_n\big) + \beta Y_n \ln\left[\frac{\beta Y_n L_n^{1-\alpha}}{1+\beta Y_n}\right] + \ln\left[\frac{L_n^{1-\alpha}}{1+\beta Y_n}\right] \\
&\quad + \beta Z_n \mathbb{E}_t\Big[\ln z_{t+1}\Big] + \alpha(1+\beta Y_n)\ln K_t + (1+\beta Y_n)\ln z_t.
\end{aligned}$$

It follows that $J_{n+1}(K_t, z_t) = X_{n+1} + Y_{n+1}\ln K_t + Z_{n+1}\ln z_t$ with

$$\begin{aligned}
X_{n+1} &= \beta X_n + \psi \ln\big(1 - L_n\big) + \beta Y_n \ln\left[\frac{\beta Y_n L_n^{1-\alpha}}{1+\beta Y_n}\right] + \ln\left[\frac{L_n^{1-\alpha}}{1+\beta Y_n}\right] \\
&\quad + \beta Z_n \mathbb{E}_t\Big[\ln z_{t+1}\Big], && \text{(B.3)} \\
Y_{n+1} &= \alpha(1+\beta Y_n), && \text{(B.4)} \\
Z_{n+1} &= 1 + \beta Y_n. && \text{(B.5)}
\end{aligned}$$

Given that L_n is constant, the right-hand sides of Equations (B.3) to (B.5) are constant, so the result holds also for $n+1$. This proves the claim. We then know by Theorem 3.2 that

$$J^*(K_t, z_t) = \lim_{n\to\infty} J_n(K_t, z_t) = \lim_{n\to\infty} X_n + \Big(\lim_{n\to\infty} Y_n\Big)\ln K_t + \Big(\lim_{n\to\infty} Z_n\Big)\ln z_t.$$

Iterating backwards over n in Equation (B.4), we obtain

$$Y_n = \alpha(1+\beta Y_{n-1}) = \alpha\Big(1 + \alpha\beta\big(1 + \beta Y_{n-2}\big)\Big) = \ldots = \alpha\sum_{i=0}^{n}(\alpha\beta)^i + (\alpha\beta)^n Y_0.$$

Since $\alpha, \beta \in (0,1)$ and $Y_0 = 0$,

$$Y \equiv \lim_{n\to\infty} Y_n = \alpha \lim_{n\to\infty}\sum_{i=0}^{n}(\alpha\beta)^i = \frac{\alpha}{1-\alpha\beta}.$$

By Equation (B.5),

$$Z \equiv \lim_{n\to\infty} Z_n = \lim_{n\to\infty}\frac{Y_n}{\alpha} = \frac{Y}{\alpha} = \frac{1}{1-\alpha\beta}.$$

We also have that the optimal labour supply in the limit is

$$L \equiv \lim_{n\to\infty} L_n = \frac{(1-\alpha)(1+\beta Y)}{\psi + (1-\alpha)(1+\beta Y)} = \frac{1-\alpha}{\psi(1-\alpha\beta) + (1-\alpha)}.$$

Defining $X \equiv \lim_{n\to\infty} X_n$ and using Y, Z, and L, the limit of Equation (B.3) becomes

$$\begin{aligned} X &= \lim_{n\to\infty} X_{n+1} \\ &= \lim_{n\to\infty} \Bigg[\beta X_n + \psi \ln\big(1 - L_n\big) + \beta Y_n \ln\left[\frac{\beta Y_n L_n^{1-\alpha}}{1+\beta Y_n}\right] + \ln\left[\frac{L_n^{1-\alpha}}{1+\beta Y_n}\right] \\ &\qquad + \beta Z_n \mathbb{E}_t\Big[\ln z_{t+1}\Big]\Bigg] \\ &= \beta X + \psi \ln\big(1-L\big) + \frac{\alpha\beta}{1-\alpha\beta}\ln\Big(\alpha\beta L^{1-\alpha}\Big) + \ln\Big((1-\alpha\beta)L^{1-\alpha}\Big) \\ &\quad + \frac{\beta}{1-\alpha\beta}\mathbb{E}_t\Big[\ln z_{t+1}\Big], \end{aligned}$$

so

$$\begin{aligned} X &= \frac{1}{1-\beta}\Bigg[\psi \ln\big(1-L\big) + \frac{\alpha\beta}{1-\alpha\beta}\ln\Big(\alpha\beta L^{1-\alpha}\Big) + \ln\Big((1-\alpha\beta)L^{1-\alpha}\Big) \\ &\qquad + \frac{\beta}{1-\alpha\beta}\mathbb{E}_t\Big[\ln z_{t+1}\Big]\Bigg]. \end{aligned}$$

Hence, we have found the value function: in sum,

$$J^*(K_t, z_t) \;=\; X + Y \ln K_t + Z \ln z_t,$$

where

$$\begin{aligned} X &= \frac{1}{1-\beta}\Bigg[\psi \ln\big(1-L\big) + \frac{\alpha\beta}{1-\alpha\beta}\ln\Big(\alpha\beta L^{1-\alpha}\Big) + \ln\Big((1-\alpha\beta)L^{1-\alpha}\Big) \\ &\qquad + \frac{\beta}{1-\alpha\beta}\mathbb{E}_t\Big[\ln z_{t+1}\Big]\Bigg], \\ Y &= \frac{\alpha}{1-\alpha\beta}, \\ Z &= \frac{1}{1-\alpha\beta}, \\ L &= \frac{1-\alpha}{\psi(1-\alpha\beta) + (1-\alpha)}. \end{aligned}$$