

10 – DP Examples and Option Value AGEC 642 - 2024

I. An infinite-horizon stochastic DD problem:

The cow replacement problem (Taken from Miranda and Fackler)

Consider a stochastic infinite-horizon dynamic optimization problem that is near to the hearts of any agricultural economist: when to replace your dairy cow. Here's how dairy cows work. They get pregnant; they give birth; and then they produce milk for a while. That sequence of events is one lactation cycle. Eventually, the cow will be replaced¹ and the cycle starts all over again. In the Miranda and Fackler framework, they assume that a dairy cow can be used up to n_1 lactation cycles and its productivity can be one of n_2 classes. Each cow belongs to a productivity class x yields $q_x y(s)$ tons of milk during the s^{th} lactation cycle so that all cows follow the same pattern for their yields, but the level of their yields varies depending on the cows. The farmer does not know the productivity class of a cow until after its first lactation.

There are two state variables in this case,

- the lactation cycle of the cow: $s = \text{lactation number of cow} \in S_1 = \{1, 2, \dots, n_1\}$
- the quality of the cow: $x = \text{cow quality} \in X = \{1, 2, \dots, n_2\}$

The choice variable is: $z = 0$ (keep cow), or $z = 1$ (replace)

The state equation for the lactation number is simple,

$$s_{t+1} = \begin{cases} s_t + 1 & \text{if } z_t = 0 \\ 1 & \text{if } z_t = 1 \end{cases}$$

The stochastic state equation for the quality variable is

$$x_{t+1} = \begin{cases} x_t & \text{if } z_t = 0 & \text{with probability 1} \\ x_i & \text{if } z_t = 1 & \text{with probability } w_i \end{cases}$$

where w_i is the probability of getting a cow of class i .

The benefit function is

$$\pi(z, x, s) = \begin{cases} pq_x y(s) & z = 0 \\ pq_x y(s) - c & z = 1 \end{cases}$$

where c is the cost of replacing a cow. Note that you pay the cost of replacing the cow *after* milking.

The sequence of activities, e.g., you pay after milking, is important and substantive in discrete-time problems. In continuous-time specifications such sequence issues tend to go away because everything can indeed be simultaneous and/or instantaneous. The order of events is, therefore, an important modelling choice.

¹ The replacement process does not usually end well for the cow.

A formal statement of the cow replacement problem, therefore, is:

$$\max_{z_t=0,1} E \sum_{t=0}^{\infty} \frac{\pi(z_t, x_t, s_t)}{(1+r)^t} \quad \text{s.t.}$$

$$x_{t+1} = \begin{cases} x_t & \text{if } z_t = 0 & \text{with probability } 1 \\ x_i & \text{if } z_t = 1 & \text{with probability } w_i \end{cases}$$

$$s_{t+1} = \begin{cases} s_t + 1 & \text{if } z_t = 0 \\ 1 & \text{if } z_t = 1 \end{cases}$$

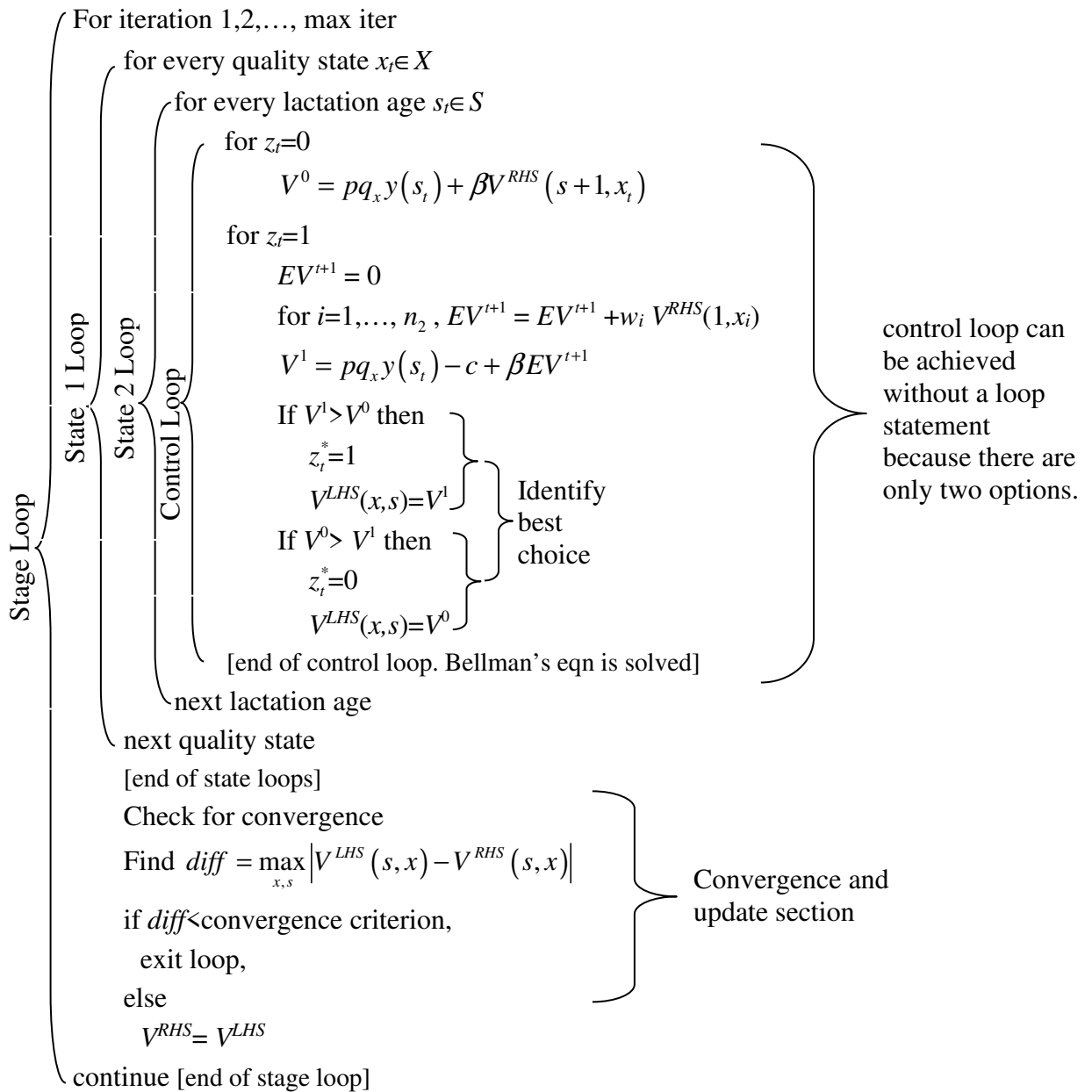
The Bellman's equation for this problem becomes

$$V(s_t, x_t) = \max_{z=0,1} \begin{cases} pq_x y(s_t) + \beta V(s+1, x_t); & \text{if } z_t = 0 \\ pq_x y(s_t) - c + \beta \sum_{i=1}^{n_2} w_i V(1, x_i); & \text{if } z = 1 \end{cases}$$

In this case, the state space will be a two dimensional array of $n_1 \times n_2$ points. This can be solved using two loops in the state space, a loop over x inside a loop over s .

Again, we can look at this process using pseudocode.

Set $V^{RHS}(x,s)=0$ for every state $x \in X, s \in S$.



II. Optimal decisions in a dynamic context

Once you have solved a dynamic programming problem, you obtain an optimal policy function and a value function. The policy function, $z^*(x,t)$ in a finite horizon model and $z^*(x)$ in an infinite horizon model, tells you the decisions, contingent on any particular state variable that you might face. In principle this could be quite useful in giving advice or analyzing optimal choices.

Consider the cow replacement problem. After you have solved the problem, you have a clear rule for replacing cows. Conventional wisdom may be that you should replace a cow after a certain number of lactation cycles, but if the cow is particularly productive, should you wait a little longer? Should you retire the cow earlier if it is at the low end of the productivity distribution? Assuming your problem specification is correct, your solution to the problem above can serve as the basis for strong advice or better management of your herd.

Second, your solution can be used to predict the outcomes in the market. For example, if the price of replacing a cow goes up, what can we expect will happen to the supply of milk in the short run and, over the longer term, what will happen to the supply of milk and the market-clearing price? These questions can only be answered by understanding how the underlying dynamic optimization problem that farmers are either implicitly or explicitly solving.

Third, if you are interested in studying how economic decisions play out in the future, the use of a simulation model might be useful. Simulation models often use an *open-loop* decision process – i.e., they assume that decisions are set in stone prior to the start of the problem. In reality of course, optimal decisions are *closed-loop*, meaning that decision makers respond to new information about the states in which they find themselves. Using a DP solution in your simulation work (i.e., incorporating $z^*(x_t)$ into your model) will realistically incorporate the fact that decision makers react to changing conditions before making decisions.

For example, suppose that you are studying a policy to promote the use of biofuels. This will change the dynamic incentives of individuals throughout the economy – from producers of corn to producers of oil and coal. One cannot simply assume that they will react to new policies in the same way that they have reacted to price changes in the past – the structure of the dynamic choice problem has been changed.² Due to rising computational power, we are increasingly able to add dynamic realism to policy analysis that these economists have promoted.

There are limits however. Many simulation models provide a benefit function and determine how the state variables change over time – the fundamental ingredients necessary for DP model. Hence, while it might appear that one could simply use the simulation model for dynamic optimization, it is common for such models to be much too large for direct use in dynamic programming. If a simulation model contains dozens and even hundreds of state variables, it is computationally impossible to directly solve the associated DP problem. The Approximate Dynamic Programming approach of Powell (2007) is probably the most promising way to approach such large-dimension dynamic programming problems though other approaches have been proposed.³

² This idea is at the heart of Kydland and Prescott's "Rules Rather than Discretion" paper (1977) and addresses the *Lucas Critique* (Lucas 1976). Kydland and Prescott received the Nobel prize in 2004 and Lucas received his Nobel Prize in 1995.

³ Woodward, Wui and Griffin (2005) suggest a work-around in which a smaller DP problem is used to approximate the large set of state variables implicit in the simulation model. Approximate dynamic

III. DP in options

A. A DD-DP Example - A stochastic finite-horizon problem: option pricing (Taken from Miranda and Fackler)

An American put option is a right to sell a specified quantity of a commodity at a particular “strike” price on or before a particular date. For example, suppose you buy the right to sell a bushel of wheat for \$2.75 on or before June 30. If on June 1 the price drops to \$2.50 you could sell your option for the difference, 25¢. But is selling on June 1 the best strategy? Would it make more sense to hold onto the option, hoping that it falls even more before June 30? Certainly the option is worth at least 25¢ when the price is at \$2.50, but is it worth even more than that? The answers to these questions can be found by realizing that the management of such an asset is a dynamic optimization problem.

The Cox-Ross-Rubenstein binomial option pricing model, assumes that the price, p_t , will go up with probability q and down with probability $1-q$. If it goes up, then $p_{t+1} = p_t \cdot \phi$ with $\phi > 1$. If it goes down, $p_{t+1} = p_t / \phi$. Note that the price after n periods can only take on a finite number of values from a maximum of $p_0 \phi^n$ to a minimum of $p_0 (1/\phi)^n$.

A period can be of any length, Δt in years (e.g., $1/365^{\text{th}}$ of a year) so that if r is the annualized discount rate, the discount factor $\beta = \exp(-r\Delta t)$.

While we won't use this here, the parameters ϕ and q are sometimes assumed to take the forms:

$$\phi = \exp(\sigma\sqrt{\Delta t}) > 1 \quad \text{and} \quad q = 1/2 + \frac{\sqrt{\Delta t}}{2\sigma} \left(r - \frac{\sigma^2}{2} \right)$$

where σ is the annualized volatility of the commodity price.

We can use dynamic programming to identify the value of an option that expires at time T with a strike price of p' . In this case the state variable is the price at time t , p_t . The control variable, z_t , indicates whether the right is exercised, $z_t = 1$, or held, $z_t = 0$. So the benefit function is

$$\pi(z, p_t) = \begin{cases} 0 & \text{if } z = 0 \text{ (keep)} \\ p' - p_t & \text{if } z = 1 \text{ (exercise and receive the difference between } p_t \text{ and } p') \end{cases}$$

The stochastic state equation for the price is

$$p_{t+1} = \begin{cases} p_t \phi & \text{with probability } q \\ p_t / \phi & \text{with probability } 1 - q \end{cases}$$

Finally, there is an implicit state variable indicating whether the option is still held, which goes from 1 to 0 as soon as you exercise your option.

The Bellman's equation, therefore, which is an expression of the value of the option at any time t and price p_t is written:

programming will be discussed toward the end of AGECE 642; a clear application and extension of this approach is provided in Springborn and Faig (2019).

$$V(p_t, t) = \max_{z=0,1} \left\{ \begin{array}{l} p' - p_t; \text{ if } z = 1 \\ 0 + \beta [qV(p_t \phi, t+1) + (1-q)V(p_t / \phi, t+1)]; \text{ if } z = 0 \end{array} \right\}$$

with $V(p_{T+1}, T+1) = 0$. Notice that if $z=1$, $V(\cdot, t+1) = 0$.

We know that the Bellman's equation always can be written, $V(\cdot) = E\{u(\cdot) + \beta V(\cdot)\}$. So what $u(\cdot)$ in the equation above?

What would the grid space for the state variable p need to look like in order to solve this problem numerically? It would need to incorporate all possible prices that might be reached in $T/\Delta T$ periods between 0 and T . Suppose that there are 10 periods, $p_0 = 3.0$, and $\phi = 1.1$. We would need to evaluate the full range of price increases at $3.0 \cdot 1.1$, $3.0 \cdot (1.1)^2$, $3.0 \cdot (1.1)^3$, ..., $3.0 \cdot (1.1)^{10}$, and the full range of price decreases, $3.0/(1.1)$, $3.0/(1.1)^2$, ..., $3.0/(1.1)^{10}$. Given the assumed multiplicative structure, these 21 points span the complete space of possible prices that can be observed in 10 periods.

B. Using dynamic programming to value an asset, the case of the American put option

As we know, one of the two main outputs every time the Bellman's equation is solved is the value function itself. This has economic meaning – it tells us the value of the state variable, assuming that from that point on the asset(s) are used optimally. Let's consider how this might be useful.

In the option pricing model above, the value function tells us the value of an option at time t , given a strike price, p' , and the current price, p_t . Suppose that you look at the market and see that the actual price of such an option is less than this price. If you trust your model, then you should buy the option. That is, if your estimates of q and ϕ are better than the values that others have, you can make money on that informational advantage. In effect, this is an argument that the market is not completely rational or you have better information. You would want to also consider the possibility that your estimates of the probability distributions are wrong. In any case, estimating the true value of the option based on your assessment of prices in the future could be very useful and requires dynamic optimization.

Let's look at another example: the case of an investor considering whether to purchase a plot of land that is valuable entirely for the timber that can be harvested from the land. The timber could be harvested today, but given future growth and variability in prices over time, you believe that it would be better to put off harvesting until some future date. Using a dynamic programming problem with two state variables, the harvestable stock of timber and the price, you could evaluate the value of the land with its standing timber for either a single or infinite stream of timber harvests. Once this DP problem is solved, you know the value of the stand that is available today, $V(x_t)$. If the asking price is less than the $V(x_t)$, then it looks like a good investment.

In addition to the level of the value function, the slope of the value function can also be a useful result of a DP model. Remember this is equivalent to the co-state variable, μ_t , in an

optimal control framework; it is the shadow price of the resource. So if you can vary your state variable holdings marginally, e.g., by purchasing or selling at a market price, then μ is maximum price you would be willing to pay increase your holdings and the minimum amount you would be willing to accept to sell a portion of your holdings. This might be useful in benefit-cost analysis or simple decision-making by a business.

C. *Real options and quasi-option value (a slight detour)*⁴

(A link to a video walking through this example is provided in the supplementary material for this lecture.)

Two ideas that are closely related to dynamic programming are “real options” of Dixit and Pindyck (1994) and quasi-option value of Arrow and Fisher (1974) and Henry (1974), which was elaborated by Fisher and Hanemann (1987). The problem that these authors consider is particularly important for the consideration of an irreversible investment, which to some degree defines just about any investment opportunity. The short paper by Mensink and Requate (2005) does an excellent job of clarifying these two concepts and I will build on their analysis here.⁵

Mensink and Requate (MR) consider the problem of making a binary choice, (e.g. to build something) $d \in \{0,1\}$, in period 1 or 2 with $d_1 + d_2 \leq 1$. The benefits that accrue in period 1, $B_1(d_1)$, are only a function of that period’s choice. The benefits in period 2, $B_2(d_1, d_2, \theta)$, are a function of whatever choice was made in 1, the choice in period 2, d_2 , and θ , a random variable. The decision in period 1 can be thought of as a state variable at the start of period 2.

We will present the framework both generally in the context of a simple example, the decision to build a hydroelectric dam. There are 2 periods, 1 and 2, representing the present and the future. The cost to build the dam is \$18 million. The benefit in period 1 is \$2. The benefit in period 2 is either $\underline{\theta} = \$10$ or $\bar{\theta} = \$20$. The net benefit of doing nothing is always zero. We will assume that the probability of $\underline{\theta}$ is 0.5. The periodic values are shown in the table below.

Cost	$B_1(0)$	$B_1(1)$	$B_2(\text{not built})$	$B_2(\text{built}, \underline{\theta})$	$B_2(\text{built}, \bar{\theta})$
18	0	2	0	20	10

The *ex ante* value at the start of period 2 is

$$V_2^O(d_1) = \max_{d_2} E_{\theta} [B_2(d_1, d_2, \theta)].$$

⁴ Aram Avanesyan, who took my class in 2012, provided comments contributed to this section.

⁵ Mensink and Requate (2005) is actually a comment on Fisher (2000), who mistakenly argued that the real option and quasi-option value were equivalent. Arguments similar to Fisher’s had also been made in previous versions of these lecture notes. I thank Gabriel Power for bringing this paper to my attention. I have changed notation substantially from MR’s presentation to be more consistent with other notation used in these notes.

This means that the decision maker will make an optimal choice, \max_{d_2} , but that choice is made before she knows the value of θ . MR call this the *open-loop second period expected value*.

In our example, the open-loop second period expected value is

$$V_2^O(d_1=0) = \max_{d_2} \begin{cases} 0.5 \cdot (10-18) + 0.5 \cdot (20-18) = -3 & \text{if } d_2 = 1 \\ 0 & \text{if } d_2 = 0 \end{cases}$$

$$= 0$$

and

$$V_2^O(d_1=1) = 0.5 \cdot 10 + 0.5 \cdot 20 = 15.$$

MR distinguish V_2^O from the *closed-loop second period expected value*, V_2^C ,

$$V_2^C(d_1) = E_\theta \max_{d_2} [B_2(d_1, d_2, \theta)].$$

In the closed-loop case, the decision maker observes θ before making the second period decision ($E_\theta \max$). The expectation is taken to find the average payoff across all possible values of θ , even though the decision, d_2 , is made without uncertainty.

In our example, no decisions are made once the dam is already built in period 1, so $V_2^C(d_1=1) = V_2^O(d_1=1) = \15 . If $d_1=0$, however, we know that the best option is to build if $\theta = \bar{\theta}$ yielding net benefits of +2 and not build if $\theta = \underline{\theta}$ because net benefits would be -8. Hence,

$$V_2^C(d_1) = 0.5 \cdot \max_{d_2 \in \{0,1\}} \left(\underbrace{0}_{d_2=0}, \underbrace{10-18}_{d_2=1} \right) + 0.5 \cdot \max_{d_2 \in \{0,1\}} \left(\underbrace{0}_{d_2=0}, \underbrace{20-18}_{d_2=1} \right) = +1$$

Note that it always the case that a decision maker is better off observing the state before making a decision; i.e., $V_2^C(d_1) \geq V_2^O(d_1)$.

Now consider the decision maker's problem in period 1. If it is not possible to observe θ before making the decision in period 2, then the problem in period 1 is

$$V_1^O = \max_{d_1 \in \{0,1\}} B_1(d_1) + V_2^O(d_1) = \max_{d_1 \in \{0,1\}} \tilde{V}_1^O(d_1),$$

where $\tilde{V}_1^O(d_1)$ is the net present value in the open loop problem if d_1 is chosen.

In our example, this means

$$V_1^O = \max_{d_1 \in \{0,1\}} \left(\underbrace{0}_{d_1=0}, \underbrace{-16+15}_{d_1=1} \right) = \max_{d_1 \in \{0,1\}} \left(\underbrace{0}_{\tilde{V}_1^O(0)}, \underbrace{-1}_{\tilde{V}_1^O(1)} \right) = 0, \quad (1)$$

On the other hand, if information is observed before making the period 2 decision, then the problem becomes

$$V_1^C = \max_{d_1 \in \{0,1\}} B_1(d_1) + V_2^C(d_1) = \max_{d_1 \in \{0,1\}} \tilde{V}_1^C(d_1).$$

$$V_1^C = \max_{d_1 \in \{0,1\}} \left(\underbrace{0+1}_{d_1=0}, \underbrace{-16+15}_{d_1=1} \right) = \max_{d_1 \in \{0,1\}} \left(\underbrace{+1}_{d_1=0}, \underbrace{-1}_{d_1=1} \right) = +1.$$

Finally, MR define the open-loop value that would result if one made the decision in period 1 and locked in this decision in period 2:

$$\bar{B}(0) = B_1(0) + E_\theta B_2(0, 0, \theta) \text{ and } \bar{B}(1) = B_1(1) + E_\theta B_2(1, 0, \theta),$$

though, as above, $\bar{B}(1) = \tilde{V}_1^C(1) = \tilde{V}_1^O(1)$.

In our numerical example, $\bar{B}(0) = 0$ and $\bar{B}(1) = -1$.

If we carried out simple benefit-cost comparison, the decision rule would amount to simply comparing $\bar{B}(1)$ and $\bar{B}(0)$. So they define $NPV = \max\{\bar{B}(1), \bar{B}(0)\}$, which equals 0 in our numerical example. However, using the NPV rule would be a mistake since it does not take into account two factors. First, that you can make a decision in period 2; the ability to delay is captured in both V_1^C and V_1^O . Second, NPV does not take into account you might get more information in period 2; that is captured only in V_1^C .

MR state that Dixit and Pindyck (1994) defines the option value (OV^{DP}) as

$$OV^{DP} = V_1^C - NPV = \max\{\tilde{V}_1^C(0), \tilde{V}_1^C(1)\} - \max\{\bar{B}(0), \bar{B}(1)\},$$

which equals +1 in our numerical example. That is, OV^{DP} is the difference between the optimal closed-loop rule and the expected present value following the NPV rule. This is the basic idea behind the real option literature – that there is value to having the *ability* to wait and take into account future information.

Twenty years before Dixit and Pindyck, Arrow and Fisher (1974) and Henry (1974) introduced what came to be known as quasi-option value, which focuses only on the benefit of taking into account additional information. That is, the benefit that is created by taking account of new information in period 2, i.e.

$$OV^{AFH} = \tilde{V}_1^C(0) - \tilde{V}_1^O(0),$$

which equals +1 in our numerical example.

We can now decompose the difference between OV^{DP} and OV^{AFH} . MR define the *pure postponement value* as

$$PPV = \tilde{V}_1^O(0) - \tilde{V}_1^O(1)$$

i.e. the additional value created by waiting until period 2 to make a decision about whether to make a decision to build.⁶ Note that the PPV can be negative if there are benefits foregone by delaying action. In our numerical example, $PPV = +1$ (see equation (1)).

⁶ Note that when there is no discounting, this would be zero. However, with discounting there might be benefits to waiting to make a decision.

Note that if the optimal closed loop choice in period 1 is $d_1=1$, then $V_1^C = V_1^O = NPV$. On the other hand, when the optimal closed loop choice in period 1 is $d_1=0$, and the optimal open loop choice is $d_1=1$, then $V_1^C \neq V_1^O$, instead $V_1^C = \tilde{V}_1^C(0)$ and $NPV = \bar{B}(1) = \tilde{V}_1^O(1) = V_1^O$. This does not hold in our numerical example, but is the more interesting case.

When $V_1^C \neq V_1^O$, OV^{DP} can be written

$$OV^{DP} = \tilde{V}_1^C(0) - \bar{B}(1)$$

which can then be decomposed as follows

$$\begin{aligned} OV^{DP} &= \tilde{V}_1^C(0) - \tilde{V}_1^O(1) \\ &= \tilde{V}_1^C(0) - \tilde{V}_1^O(1) + [\tilde{V}_1^O(0) - \tilde{V}_1^O(0)] \\ &= \underbrace{\tilde{V}_1^C(0) - \tilde{V}_1^O(0)}_{OV^{AFH}} + \underbrace{\tilde{V}_1^O(0) - \tilde{V}_1^O(1)}_{PPV} \end{aligned}$$

Hence, we see that the option value concept of Dixit and Pindyck can be decomposed into two parts:

$$OV^{DP} = OV^{AFH} + PPV.$$

Why do we care?

To get a sense of the importance of considering OV, consider a stylized version of the problem considered by Arrow and Fisher (1974): that of a government planner considering whether to construct a hydroelectric dam. This project will yield a stream of future benefits in the form of recreational opportunities on the reservoir and hydroelectric power (for simplicity we're ignoring any long-term environmental or social costs). The magnitude of these future benefits will be largely determined by future population growth in the region. The project is expensive, however, involving large financial resources for construction.

If we simply build if the PV of the benefits exceeds the costs, then we use the mistaken NPV rule. We should take into account that the decision can be made in the future and there will be better information when we make that decision – we will have a better estimate of the population in the future. The option value, tells us the additional value from simply leaving our options open.

This final equation above is interesting mostly for historical reasons. We see that in 1974 AFH identified that optimal choices should take into account new information. Nearly 20 years later this general idea was picked up in the real options literature and applied to the many irreversible decisions that occur in business and individual choices. Obviously the concept was not completely ignored in the intervening 20 years, but it did not come to the forefront until Dixit and Pindyck developed it in detail.

Some authors interpret OV as a value that should be added to benefit-cost studies, leading to the correct decision.⁷ An alternative perspective is that if you're able to calculate OV, then you must have considered the correct DP problem, so it is probably better to just model your decision in this correct manner (Bishop).

IV. References

- Arrow, Kenneth J. and Anthony C. Fisher. 1974. Environmental Preservation, Uncertainty and Irreversibility. *Quarterly Journal of Economics* 88:312-319.
- Bishop, Richard C. "Option Value: An Exposition and Extension." *Land Economics* 58(Feb 1982):1-15.
- Dixit, Avinash K. and Robert S. Pindyck. 1994. *Investment Under Uncertainty*. Princeton, N.J.: Princeton University Press.
- Fisher, Anthony C, and W Michael Hanemann. 1987. Quasi-option Value: Some Misconceptions Dispelled. *Journal of Environmental Economics and Management* 14(2):183-90.
- Henry, Claude. "Investment Decisions Under Uncertainty: The "Irreversibility Effect."." *American Economic Review* 64(Dec 1974):1006-12.
- Kydland, Finn E., and Edward C. Prescott. 1977. Rules Rather Than Discretion: The Inconsistency of Optimal Plans. *Journal of Political Economy* 85(3):473-91.
- Lucas, Robert (1976). "Econometric Policy Evaluation: A Critique." Carnegie-Rochester Conference Series on Public Policy 1: 19–46.
- Mensink, Paul, and Till Requate. 2005. The Dixit-Pindyck and the Arrow-Fisher-Hanemann-Henry Option Values Are Not Equivalent: A Note on Fisher (2000). *Resource and Energy Economics* 27(1):83-88.
- Miranda, Mario J., and Paul L. Fackler. 2002. *Applied Computational Economics and Finance*. Cambridge, Mass.: MIT Press.
- Springborn, Michael R., and Amanda Faig. 2019. Moving Forward: A Simulation-Based Approach for Solving Dynamic Resource Management Problems. *Marine Resource Economics* 34(3):199-224.
- Woodward, R.T., Griffin, W.L. and Wui, Y.-S. 2005. *DPSIM Modeling: Dynamic Optimization in Large Scale Simulation Models*. Chapter 14 pages 275-294, in Scarpa R. and Alberini, A. (eds.). *Applications of simulation methods in environmental and resource economics*. Springer Publisher, Dordrecht, The Netherlands.

⁷ Indeed, Arrow and Fisher (1974) state at the very end of their paper, "Essentially, the point is that the expected benefits of an irreversible decision should be adjusted to reflect the loss of options it entails" (p. 319).