Moving Forward: A Simulation-Based Approach for Solving Dynamic Resource Management Problems

Michael R. Springborn Amanda Faig

Abstract

Standard dynamic resource optimization approaches, such as value function iteration, are challenged by problems involving complex uncertainty and a large state space. We extend a solution technique to address these limitations called approximate dynamic programming (ADP). ADP recently emerged in the macroeconomics literature and is novel to bioeconomics. We demonstrate ADP in solving a simple fishery management model under uncertainty to show: the mechanics of ADP in simplest form; the accuracy of ADP; the value of a nonparametric extension; and readily adaptable, non-specialized code. We then demonstrate ADP's capacity to handle rich bioeconomic problems by solving the fishery management problem subject to four autocorrelated shock processes (governing economic returns and biological dynamics) which entails four sources of stochasticity and five continuous state variables. We find that accounting for multiple autocorrelation has important impacts on harvest policy and generates gains that depend crucially on the structure of harvest cost.

Key words: Dynamic optimization, bioeconomic model, fishery, uncertainty, approximate dynamic programming, reinforcement learning, simulation, nonparametric, autocorrelation, non-stationarity.

JEL Codes: C61, Q22, Q57.

Running Title: "Moving Forward"

Michael R. Springborn is an associate professor, Department of Environmental Science & Policy, University of California Davis, 1 Shields Avenue, Davis, CA 95616 USA (email: mspringborn@ucdavis.edu). **Amanda Faig** is a research associate, Resource Ecology and Fisheries Management Division, Alaska Fisheries Science Center, 7600 Sand Point Way NE, Seattle, WA 98115 USA (email: amanda.faig@noaa.gov).

We thank CDFW-ERP (Grant No. E1383002) for funding for this project.

We thank Isaiah Hull for helpful conversations about the ADP approach as well as Mark Agerton, David Kling, Paul Fackler, Bruno Nkuiya, and Jim Sanchirico.

Introduction

Resource management models usually require the marriage of social-environmental modeling and optimization techniques. For example, a standard analysis of efficient fishery management might involve solving an integrated model of biology and user behavior using dynamic programming tools like value function iteration (VFI). To date, these tools have generated useful intuition by providing reliable solutions to relatively simple problems. However, such standard approaches are not well-positioned to address evolving future needs. The frontier of natural resource management is moving towards an ever richer representation of social-environmental systems, involving an expanding appetite for incorporating more states, increasingly evoking the curse of dimensionality. Furthermore, capacity to address uncertainty is central (LaRiviere et al. 2017). In a recent review of dynamic analysis in fisheries, Clark and Munro (2017) argue that we are forced to face "the question of how best to mitigate the consequences of uncertainty" given the centrality of uncertain returns to investment in resource management.

In this article we illustrate and extend a Monte Carlo simulation-based stochastic optimization approach with very limited previous application to natural resource management known as "approximate dynamic programming" (ADP). Key strengths of ADP are (1) its facility with high-dimension problems, and (2) its use of simulation to capture complex uncertainty in biophysical and economics dynamics. ADP first appeared in the operations research and engineering literatures (Powell 2007; Bertsekas 2011) and is also referred to as reinforcement learning (Sutton and Barto 1998) or neuro-dynamic programming. Judd, Maliar, and Maliar (2011), followed by Maliar and Maliar (2013) and Hull (2015), were the first to develop full applications of this method in the economics literature, specifically to solve high-dimension macroeconomic problems. In short, ADP involves iterating between two steps—in the first step Monte Carlo simulations are run to generate samples, and in the second stage these samples are used to update estimates of a value function or policy function.

Existing approaches for using simulation in bioeconomic optimization have been limited. Taleghan et al. (2015) and Hall et al. (2018) use simulation to construct an approximation of the Markov transition matrix before solving with standard VFI.[1] This approach facilitates complex uncertainty processes for which "exact inference is intractable" (Taleghan et al. 2015). Alternatively, Moxnes (2003) begins with specifying a functional form for the policy (defined over the state space) that is conditioned on a policy parameter vector. Next, multiple paths for stochastic shocks are produced, each over the same finite-horizon. Finally, nonlinear optimization is used to identify the policy parameter vector that maximizes average welfare across the set of simulations.[2] While both of these approaches expand the frontier of bioeconomic optimization,

they have limitations. First, they introduce approximation error through an assumed policy function form (Moxnes 2003) or an approximated Markov transition matrix (Taleghan et al. 2015; Hall et al. 2018). Second, both approaches struggle with scale as additional dimensions in state space are considered. This is due to the standard dimensionality limits of VFI and further in the approach of Moxnes (2003) via simplification required in the policy function.

In natural resource management, the closest existing work to what we present here is Fonnesbeck (2005) who solves a bioeconomic optimization problem using a reinforcement learning approach based on Monte Carlo simulation.[3] Fonnesbeck is similarly motivated by the need to handle large-scale resource management problems. The application is innovative but does not progress beyond a two-state variable problem (also solvable using dynamic programming) and achieves only a "coarse approximation of the true optimal policy".[4]

ADP allows for the application of dynamic programming to particularly "large and complex problems", providing tractability despite high dimensionality (Bertsekas 2011). Demand on computer memory—a common bottleneck in numerical optimization—is limited, since the core of the procedure is forward simulation rather than manipulation of large arrays (as in VFI). In richer problems, transition dynamics can be so complex that inference is intractable, leaving simulation the only practical approach for treating dynamics (e.g., as Taleghan et al. (2015) show in a bioeconomic model of invasive riverine tamarisk control). Instead of handling stochasticity through numerical integration or the construction of potentially large (state and action-dependent) Markov transition matrices, ADP uses Monte Carlo simulations to inform the expectation. This minimizes the complexity and potential inaccuracies these elements can generate.

ADP is of particular value in marine resource management problems for two reasons: the need to address complex uncertainty and facilitate ecosystem-based fisheries management (EBFM). In their recent discussion of dynamic analysis in fisheries, Clark and Munro (2017) highlight the "increasingly popular ecosystem approach to fisheries management" where decision makers manage a portfolio of assets; i.e., multiple important stocks. A special committee of the American Fisheries Society, providing input into the amendment process for the Magnuson–Stevens Fishery Conservation and Management Act in the US, recently reiterated the importance of explicitly "account(ing) for uncertainty and change in the climate and ecosystem" and of implementing EBFM—a "holistic approach" focusing on "multiple species" (Miller et al. 2018). Facilitating these marine resource management demands requires expanding the standard toolkit with emerging techniques like ADP.

The first key objective of this article is to demonstrate, in simplest form, the essential elements of ADP

for solving environmental management problems. To do so we use ADP to solve a basic resource management model, specifically the canonical model of fishery management under uncertainty developed by Reed (1979). The accompanying computer code we provide for this simple model is accessible and readily adaptable to other resource management problems. We provide Matlab code for replicating the analyses in this article at `https://github.com/mspringborn/ADP_fishery_autocorrelation`. This intentionally simple first model allows us to also solve the problem with a standard VFI approach to verify ADP's accuracy and reliability.

Dynamic programming typically requires representation of a value function at its core. Contributing more broadly to the ADP literature, we show the utility of a nonparametric representation of the value function—e.g., using Gaussian process regression—to address weaknesses of existing approaches. Using the simple Reed model, we compare the performance of parametric methods (common in the nascent ADP literature) with our nonparametric approach. While parametric methods are subject to substantial error, the nonparametric approach can accurately identify optimal policies and value functions. To our knowledge, this work is first to use nonparametric regression to improve approximation of the value function in ADP.

The second key objective of this article is to generate new insights in applying ADP to a pressing problem with complex uncertainty and a large state space. Since Reed's (1979) analysis of a single source of uncertainty in stock dynamics, a growing literature has examined various kinds of uncertainty. This includes uncertain stock levels (Clark and Kirkwood 1986; Roughgarden and Smith 1996; Sethi et al. 2005), uncertain harvest implementation (Roughgarden and Smith 1996; Sethi et al. 2005), uncertainty and spatial interactions (Costello and Polasky 2008), uncertainty and capital adjustment (Singh, Weninger, and Doyle 2006), and the effect of attitudes toward risk (Lewis 1981; Kapaun and Quaas 2013). Many studies involve random perturbations to the bioeconomic model (Walters and Hilborn 1978; Nøstbakken and Conrad 2007). Nøstbakken (2006) observes that such studies usually focus on a single source of uncertainty. Some exceptions that consider two or three sources include Hanson and Ryan (1998), Sethi et al. (2005), and Nøstbakken (2006).[5]

Another typical simplification is to assume that shock levels are independent and identically distributed (i.i.d.) over time. A handful of exceptions allow shock levels to evolve in an autocorrelated fashion over time (Parma 1990; Walters and Parma 1996; Spencer 1997; Nøstbakken 2006). However, treatment is limited to a single autocorrelated variable. Limitations to the number of sources of uncertainty are driven by the fact that "the computational challenges are significant" (Rodriguez et al. 2011). On top of this, accounting for autocorrelation in shock levels requires an additional state variable for each stochastic parameter. To show ADP's facility with such complexity, in the latter half of the article we extend the model to a problem with

four autocorrelated shock processes, which, combined with the stock level, requires five continuous state variables.

Rodriguez et al. (2011) argue that accounting for fluctuations in multiple bioeconomic parameters "that arise in any real-world management situation is important." Nøstbakken and Conrad (2007) note that uncertainty can enter in many ways, including biologically—due to stochastic stock-growth dynamics—as well as economically, through fluctuating prices and costs. Earlier work addressing parameter variation has mainly focused on one type or the other. We consider two sources of biological uncertainty and two sources of economic uncertainty, operating simultaneously.

Autocorrelation in stochastic parameters is a common reality in bioeconomic settings. Parma (1990) and McGough, Plantinga, and Costello (2009) observe that autocorrelated productivity shocks are apparent in many fisheries, belying the typical assumption of a stationary stock-recruitment relationship (i.e., with i.i.d. shock levels). Economic parameters, such as price and harvest cost are also likely autocorrelated due to persistent trends in demand (due to evolving consumer tastes, dynamics of substitute goods, etc.) and supply (due to evolving costs in the labor market and input costs like fuel). Nøstbakken (2008) and Deroba and Bence (2008) argue that research is needed to address the implications of autocorrelation for optimal management.

Autocorrelated shock levels can have important effects on optimal management. For example, autocorrelation associated with biological dynamics (e.g., predatory pressure, general mortality, and growth) affects optimal policy (Parma 1990; Spencer 1997; Singh, Weninger, and Doyle 2006) and static policies are less efficient (Walters and Parma 1996). Parma (1990) finds that "escapements are raised when favorable conditions are anticipated and they are lowered when poor environments are expected" and that "feedback responses reinforce recruitment fluctuations and lead to a sequence of boom and bust periods in the fishery." In one of the few studies with autocorrelation on the economic side of the model, Nøstbakken (2006) conducted a real options analysis of a switching policy (from no harvest to a maximum harvest level) under prices that follow geometric Brownian motion (and with uncertainty in stock growth). While she finds that pulse fishing is optimal, she observes that the maximum harvest rate of the fishing fleet dominates the switching decision, rather than price volatility.

Although ignoring autocorrelation when determining optimal policy can lead to inaccurate results, bioeconomic models typically ignore it or consider only a single source. This is likely due, at least in part, to the computational challenge of solving such problems with traditional dynamic optimization methods, since each

autocorrelated parameter carries an additional state variable and source of stochasticity. Here we show that ADP handles integrating over four sources of stochasticity and five continuous state variables with relative ease.

Relative to naively treating shocks to parameters as i.i.d., we find that accounting for autocorrelation leads to large differences in response to observed shocks. We also find that optimal policy is quite sensitive to shocks to stock-recruitment function; thus, previous findings that growth uncertainty has a negligible effect on optimal policy do not hold in general (Sethi et al. 2005). While economic shocks are important, they are typically dominated by biological shocks. Additional rents from autocorrelation-savvy management strongly depend on harvest cost structure. Finally, accounting for autocorrelation is important for avoiding fishery closures; the savvy manager substantially reduces the rate of closures in all cases.

Methods

The Bioeconomic Model

To illustrate the ADP approach, we replicate the standard stochastic dynamic fishery model originated by Reed (1979). The problem entails a sole owner selecting harvest to maximize the expected present value of profits subject to specified growth dynamics. In each period, t, the manager observes the fish stock available for harvest, x_t, and chooses the level of harvest, h_t. The escapement ($s_t = x_t - h_t$) grows according to a stock-recruitment equation, $G(s_t)$. After growth, the stock is subjected to a random, multiplicative growth shock, z. While the standard Reed model accounts for the shock at the end of the discrete period, to facilitate our solution approach (as described below) we index the shock as accruing at the beginning of the time period. Stochastic stock dynamics are thus given by:

$$x_{t+1} = z_{t+1}G(s_t) = z_{t+1}G(x_t - h_t). \tag{1}$$

The objective is to maximize the present value of long-run profits, where the profit function is given by $\pi(h_t|x_t)$, and β is the discount factor. Profit is assumed to follow:

$$\pi(h_t|x_t) = p \cdot h_t - \int_{x_t - h_t}^{x_t} \left(\frac{c}{X_t}\right) dX_t = p \cdot h_t - c \cdot \ln\left(\frac{x_t}{x_t - h_t}\right), \tag{2}$$

where p is the price per unit of harvest, and c is a constant (Reed 1979).[6] We generally follow Sethi et al. (2005) in our specification of bioeconomic parameters as summarized in the appendix.[7]

We consider a logistic stock-recruitment equation to model population growth:

$$n_{t+1} = G_l(s_t) = s_t \left[1 + R \cdot \left(1 - \frac{s_t}{K}\right)\right], \tag{3}$$

where R is the growth rate, and K is the carrying capacity. The initial stock (pre-harvest) in period $t+1$ is n_{t+1}, which is equivalent to the post-growth population from period t. The standard problem specified here with logistic growth is known to lead to a well-behaved value function. To provide a stronger challenge, we also consider an alternative growth model that generates a convex-concave value function. Modifying the logistic stock-recruitment relationship in equation 3 following Conrad (2010, 77), we add the following additional form exhibiting critical depensation:

$$n_{t+1} = G_c(s_t) = s_t \left[1 + R \cdot \left(1 - \frac{s_t}{K}\right)\left(\frac{s_t}{K_0} - 1\right)\right], \tag{4}$$

where R is the growth rate, K is the carrying capacity, and K_0 is the "critical population level." Because $n_{t+1} < s_t$ when $s_t \in (0, K_0)$, in a deterministic model K_0 is also the "minimum viable population;" i.e., the population size below which extinction is inevitable. However in our model, it is possible for growth shocks to move the population in and out of this negative growth stock range. In contrast to $G_l(s_t)$, $G_c(s_t)$ is not consistently concave—the critical depensation leads to a convex function when $s_t \in (0, (K + K_0)/3)$. Bulte and Kooten (2001, 90) motivate the use of this form arguing that population dynamics are "more complex than usually modeled by economists" and that "it is necessary to expand the ecological underpinnings" of bioeconomic models.

A common form of the Bellman equation for this problem is:

$$\begin{aligned} J(x_t) &= \max_{h_t} \{\pi(h_t|x_t) + \beta E_z [J(x_{t+1})]\} \\ s.t. \quad & x_{t+1} = z_{t+1} G(x_t - h_t), \\ & h_t \leq x_t. \end{aligned} \tag{5}$$

However, by shifting the value function argument to be the *pre-shock* state (n_t), the Bellman can be re-

expressed as:

$$\begin{aligned} V(n_t) &= E_z\left[\max_{h_t}\{\pi(h_t|n_t,z_t)+\beta V(n_{t+1})\}\right], \\ s.t. \quad & n_{t+1}=G(z_t n_t - h_t), \\ & h_t \leq z_t n_t. \end{aligned} \tag{6}$$

This expression results from shifting the accounting of events back by one operation such that the first step is the realization of the shock. This moves the expectations operator outside of the maximization calculation. This way of framing the decision problem is sometimes referred to as using the "post-decision state" (Judd 1998). Here, the approach simply involves positioning the start of a period such that the stochastic event occurs first. This structure allows the maximization problem (conditional on the shock) to be deterministic. Furthermore, this structure allows avoidance of numerical integration to calculate the expectation directly since the Monte Carlo simulations used in ADP serve to capture uncertainty (Hull 2015). These features can significantly simplify solving for the value function as exploited in the solution technique described next.

Solution Method

As an alternative to standard iterative approaches (e.g., VFI) for solving dynamic stochastic models, ADP can be characterized broadly as a stochastic simulation method. Other numerical alternatives include projection methods and perturbation methods. Judd, Maliar, and Maliar (2011) provide a useful overview of the relative advantages and disadvantages across these three classes. Projection methods involve approximating solutions over a given domain using deterministic integration. They are quick and accurate, but slow dramatically when the number of state variables expands. Perturbation methods identify solutions locally using Taylor expansions of optimality conditions. They can handle high-dimensional applications, but accuracy is limited.

Stochastic simulation methods, in general, can handle high-dimensional applications. However, they can be less accurate than projection methods and are subject to potential numerical instability. Judd, Maliar, and Maliar (2011) developed a generalized stochastic simulation algorithm aimed to be accurate, stable, and able to handle high-dimensional applications. They achieve this primarily by normalizing variables, implementing parametric regression tools for ill-conditioned problems (e.g., Tikhonov regularization), and choosing the integration method carefully. Below we present the ADP algorithm for solving the standard problem from the bioeconomic model section, above, to demonstrate its accuracy relative to established methods, develop a nonparametric extension, and convey the intuition and mechanics of the approach in a simple, yet meaningful, setting.

Our ADP algorithm is informed by precursors from Judd, Maliar, and Maliar (2011); Maliar and Maliar (2013); and Hull (2015). The intuition is as follows. We start with a rough guess of the value function. Conditional on this guess, we generate a set of observed draws from the value function, constructed from a draw of the stochastic shock and the resulting optimal management choice. We next use regression to fold the set of value function observations into an updated value function estimate. In our preferred approach we implement this step using nonparametric methods (Gaussian process regression), which is stable and does not impose arbitrary structure on the value function.[8] This process is repeated until the value function converges.

Next, we outline the individual steps of ADP in compact fashion before describing key components in greater detail further below. The first two steps involve setup and initialization—the dynamic core of the algorithm is executed in step 3 and depicted, in part, in figure 1.

An approximate dynamic programming algorithm:

1. Set ADP parameters and functional forms.
   (a) Choose the time horizon, T, for each simulation chain.
   (b) Choose the number of simulations, $\bar{m}$, to complete before executing each regression step.
   (c) Choose a functional form that will determine the step size (or smoothing parameter), $\delta_t \in [0, 1]$.
2. Initialize the value function and the state space.
   (a) Select an initial guess for the value function over the domain of the states, $\bar{V}^{k=0}(n)$.
   (b) Define a discrete grid over the state space, $\hat{n}$.
   (c) Set the regression counter to $k = 1$.
3. For each simulation iteration $m = 1, ..., \bar{m}$ (generating observations to use in regression k), execute the following steps:
   (a) Randomly select an initial state, $n_{t=1}^m$.[9]
   (b) For each period $t = 1, ..., T$ within simulation m, execute the following steps:
      i. Randomly select the shock, z_t^m.
      ii. Choose h^* to maximize the value:

$$\tilde{v}_t^m(n_t^m) = \max_{h_t} \Big\{ \pi(h_t | n_t^m, z_t^m) + \beta \bar{V}^{k-1}(n_{t+1}^m) \Big\}.$$

      iii. Calculate the step size, $\delta_t \in [0, 1]$.
      iv. Compute the expectation with a linear combination of the newly obtained optimized value

and the previously approximated value at state n_t^m:

$$\widetilde{V}_t^m(n_t^m) = \delta_t \tilde{v}_t^m(n_t^m) + (1-\delta_t)\bar{V}^{k-1}(n_t^m).$$

v. Compute the next period state conditional on taking the optimal action: $n_{t+1}^m = G(n_t^m z_t^m - h^*)$.

(c) After completing $\bar{m}$ simulations of T periods each, there are $\bar{m} * T$ observations of the state visited ($\mathbf{n}$) and the associated updated value estimate ($\widetilde{\mathbf{V}}$). Regress $\widetilde{\mathbf{V}}$ on $\mathbf{n}$.[10,11] Let $\bar{V}^k$ represent the fitted model from the regression; i.e., the updated value function estimate.

(d) Check for convergence. Calculate the maximum deviation between the current and former value function estimate across the set $\hat{n}$: $\Delta_k = \max\{\bar{V}^k(\hat{n}) - \bar{V}^{k-1}(\hat{n})\}$. If the average of Δ_k over the last $\bar{k}$ regressions is less than the stopping tolerance, the convergence criterion is met and the program can be terminated.[12] Otherwise, increment the regression counter by one ($k = k+1$) and repeat step 3.

After convergence, the optimal policy function, $h(n_t)$, is computed using the final estimate of the value function above, $\bar{V} = \bar{V^k}$.

[Figure 1 about here.]

Parameter values and further details of the ADP algorithm are presented in the appendix and online-only appendix. Selecting the number of periods per simulation, T, is chosen to balance a tradeoff. On the one hand we want to ensure that the implications of starting at a given state are "felt;" for example the possibility of the population increasing or decreasing. On the other hand, excessive representation of the steady-state region—to which simulation chains congregate given sufficient time—does not provide much additional information. The number of simulation chains, $\bar{m}$, completed between each regression step k is selected to balance another tradeoff: increasing $\bar{m}$ allows for more extensive representation of chains initiated across the state space, but this delays incorporation of new information in the regression step (3c) slowing convergence.

The step size, δ_t, specifies the relative weighting of new versus existing information as implemented in step 3(b)iv. Since the initial guess is likely to be poor, a larger weight on new information is useful at the outset, and speedy transition to the neighborhood of the solution is desirable. However, once the value function estimate has transitioned to a stable neighborhood, a lower weight is desirable to temper the influence of

new *stochastic* realizations of value (as we seek to converge on the expected value function). Application of the step size is alternatively called smoothing, a linear filter or stochastic approximation. It is required solely due to the randomness in observed values of $\tilde{v}_t^m$ and facilitates computation of the expectation (Powell 2011).

A number of alternative step size functions have been explored in the literature, including constant and decreasing weights (Powell 2011). Hull (2015) states that a desirable step size formulation is high during early iterations but falls quickly as observations accrue and subsequently become less informative. We use a step size that is initially high *and constant*, while the value function estimate is moving away from the initial guess and consistently towards higher or lower values. Once the value function estimate is no longer consistently moving in one direction, we switch to a step size function that decreases exponentially until reaching a lower bound (as detailed in the online-only appendix).

The approach outlined above facilitates continuous state, shock, and action variables with relative ease. While we define a discretization of the state space ($\hat{n}$), this vector only serves as a convenience for selecting starting points for each simulation chain and a set of nodes at which we check for convergence in the value function. The state is allowed to vary continuously in the Monte Carlo simulation step.

In figure 2 we present our ADP "dashboard" to provide a snapshot of the algorithm in action after a limited number of regression steps ($k = 96$). Subplot 2A illustrates the last block of simulated value function draws ($\mathbf{n}$ and $\widetilde{\mathbf{V}}$) and the model generated by nonparametric regression ($\bar{V}^k$), described further in the next section. Subplot 2B shows the residuals from the regression which can illustrate if and when bias is introduced into the value function model (discussed further below). Subplot 2C shows the path of the convergence statistic. Instead of using the maximum absolute deviation from the a single regression update (Δ_k)—which is highly variable and could result in a false indication of convergence—we use a rolling average over the last 10 updates. Subplot 2D shows the current estimate of the policy function (in escapement units) given the post-shock level of the stock ($n \cdot z$). Subplot 2E shows the step size, loosely the weight on new information. We maintain a high, constant step size until the threshold for switching to the declining formulation is reached, as depicted in 2F (at $k = 80$ in this example). For the step size statistic, we also use a rolling average to ensure the switch is not driven prematurely by a stochastic draw.

[Figure 2 about here.]

Flexible Representation of the Value Function

We also aim to advance methodology for a historically frustrating component of dynamic optimization: representation of the value function. In step 2 of the algorithm above, the value function must be captured via a lookup table, parametric model, or nonparametric model (Powell 2011, 233). This choice is tied to the regression procedure used in step 3c. The regression step exploits the observed information in $\mathbf{n}$ and $\widetilde{\mathbf{V}}$ to update the value of states surrounding the observations.

Existing applications have used either a lookup table (e.g., Hull 2015) or a parametric model (e.g., Judd, Maliar, and Maliar 2011; Maliar and Maliar 2013). A lookup table for the value function defined at discrete values is simple: no particular structure is imposed on the value over the state variables. However this approach generates discretization bias, especially as dimensions increase and grids become more coarse to ensure tractability. In contrast, parametric (e.g. polynomial) models exploit structure in the value over the state variables (Powell 2011, 304). The advantage is that fewer points are needed and optimization is accelerated by the increased smoothness (Judd 1996). However, as Powell (2011, 316) summarizes, the promise of parametric models is countered by a key handicap, "they are only effective if you can design an effective parametric model, and this remains a frustrating art."[13]

To address the weaknesses of both lookup tables and parametric approximations, we extend the ADP toolbox to implement a nonparametric representation of the value function. This generates a continuous function that allows for very flexible behavior without the need to choose a specific parametric structure. Powell (2011, 316) observes that nonparametric methods hold "tremendous" promise but face substantial challenges. We implement a nonparametric approach that uses data from Monte Carlo simulations to establish both the structure and levels of the function. Specifically, we use Matlab's Gaussian process regression routine (fitrgp), which flexibly captures both the direct contribution of variables and interactions.

Results

We solve the management problem specified in equation 6 using ADP with a nonparametric value function and various parametric alternatives. As a benchmark, we also solve the problem with the standard VFI approach (Judd 1998). This allows us to show that nonparametric ADP reliably and accurately identifies optimal solutions, while parametric approaches are problematic.

Before presenting ADP and VFI solutions below, we first illustrate the problem created by parametric functional form approaches to the ADP value function. In figure 3 we present a snapshot of the key regression

step in the middle of the solution process (i.e., given some arbitrary number of regression steps already completed but well before convergence). The top row illustrates the last block of simulated value function draws ($\mathbf{n}$ and $\widetilde{\mathbf{V}}$) and the model generated by the regression. The first two columns illustrate outcomes under a parametric model (quadratic or quartic) while the final column follows from the nonparametric model. Visually we see that the parametric models admit shapes known to be inappropriate for this model. The value function should be concave and non-decreasing. Both parametric models are decreasing over some range shown in the figure. The quartic model also results in undesirable convexity to approach the expected shape. A cubic model (not presented here) resulted in similar problems. While these polynomials cannot be non-decreasing and concave over the entire positive domain, to be effective approximating functions they would need to satisfy these properties over the relevant domain plotted here.

Such shape property violations are known to lead to unstable fluctuations and substantial errors in standard VFI applications (Cai et al. 2017). The residuals of these regressions show the bias introduced by the parametric models: residuals trend positive and negative in a systematic fashion.[14] Essentially the "data" from the simulated observations is unable to "pull" the function into a more accurate shape given the parametric form. In contrast, the nonparametic value function meets shape expectations, and residuals are centered around zero. As expected, residuals are increasing in stock (n) since stochasticity in the model enters as a multiplicative shock to the population. Observations in figure 3 are sparse at larger stock levels and sometimes concentrated near a particular midpoint. This is expected given the constant escapement policy that emerges (inclusive of post-harvest stock growth) as seen in figure 2D.

[Figure 3 about here.]

Value and Policy Function Solutions

Next we present solution results for the base growth model without the convexity introduced by critical depensation. In figure 4 we show the value functions (left subplot) and policy functions (right subplot) that result from the VFI solution approach and from ADP under the nonparametric and three parametric models. We identify the VFI solution using a lookup table representation of the value function (treating n as a continuous variable using piecewise cubic interpolation). While this approach can be burdensome in terms of computer memory use for multi-state problems, the single state considered here is easily captured. In our application, VFI generates the benchmark solution as a reference to assess the performance of the other approaches.

In figure 4 we see that our nonparametric ADP approach is the only ADP alternative to successfully identify the optimal policy and value function. The VFI and nonparametric ADP results are visually indistinguishable. The parametric value functions—as foreshadowed in figure 3—take unreasonable forms and the escapement policy is biased. Parallel results for the critical depensation growth model—in online-only appendix figure 1—show that other parametric ADP approaches fail, while the nonparametric ADP approach accurately identifies the convex-concave form of the value function that emerges under this growth function. This value function is convex for low stock levels and eventually concave since the growth function is also convex-concave as governed by the critical population level ($K_0 = 25$).

[Figure 4 about here.]

A key metric of performance is the policy function error relative to the benchmark VFI (lookup table) approach given by a constant escapement level. Under both growth models, the nonparametric ADP approach is quite accurate, with an error of 0.2 to 0.4%. In contrast, parametric approaches show error between 1.4 to 67%. The cubic ADP approach fails to capture the true shape of the value function, but (presumably by chance) shows here the least amount of error. However, the error is not consistently low, ranging from 1.4 to 20% depending on the growth model. The most flexible function (quartic) performs worst, with error of 16 to 67%. The error generally increases with the order of the parametric model, albeit with an exception (i.e., cubic value function for the critical depensation growth model). We also consider the value function error, defined as the absolute difference between the left and right-hand sides of equation 6, conditional on the optimal policy. For the ADP solution, checking values of the population that really matter—from the level resulting from the optimal policy down to half that level—we find that the value function error is quite small, 0.01-0.06% (both growth models). To check for sensitivity of these results to parameters, we consider alternative levels of the cost and carrying capacity parameters (c, K) that differ by $+/-10\%$. Results were qualitatively similar: negligible policy and value function error under the nonparametric approach and substantial error from the parametric alternatives.

For the single-state benchmark model, solution run times are slower for ADP.[15] The VFI run time (for both growth models) is 54 seconds. ADP with nonparametric regression takes an average of five minutes and 23 seconds (critical depensation) to six minutes (base growth model). Thus, the core motivation for using ADP is not speed; rather to enable solutions for models—as developed in the next section—with high dimensionality and/or complex uncertainty that makes VFI unwieldy.

Accounting for Autocorrelation in Multiple Shock Levels: Several Continuous States and Sources of Uncertainty

To demonstrate the capacity of ADP to handle several continuous state variables and several sources of uncertainty, we extend the model to incorporate multiple autocorrelated shocks to parameters. In the biological model, in lieu of the single multiplicative growth shock ($z_t \cdot n_t$), we embed shocks to carrying capacity (z_t^K) and to the growth rate (z_t^R) directly in the stock-recruitment function:[16]

$$x_t = G_{l,z}(z_t^K, z_t^R, n_t) = n_t \left[1 + z_t^R r \cdot \left(1 - \frac{n_t}{z_t^K K}\right)\right]. \tag{7}$$

For the critical depensation model we consider shocks to the carrying capacity and the critical population level ($z_t^{K_0}$):

$$x_t = G_{c,z}(z_t^K, z_t^{K_0}, n_t) = n_t \left[1 + R \cdot \left(1 - \frac{n_t}{z_t^K K}\right)\left(\frac{n_t}{z_t^{K_0} K_0} - 1\right)\right]. \tag{8}$$

We also incorporate profit shocks to price (z_t^p) and cost (z_t^c):

$$\pi(z_t^p, z_t^c, h_t | x_t) = z_t^p p \cdot h_t - z_t^c c \cdot \ln\left(\frac{x_t}{x_t - h_t}\right). \tag{9}$$

Equation 9 maintains the assumption of density-dependent harvest cost. We also consider density-*independent* harvest costs given by $z_t^c c \cdot h_t$. Results discussed below show that the cost function form holds a strong influence on the welfare returns to autocorrelation-savvy management.

In the three equations above we assume that the shock levels, z_t^i, can be autocorrelated. We generally follow the autocorrelated, stochastic process specification of Spencer (1997), where the shock levels evolve according to:

$$z_{t+1}^i = \rho \cdot z_t^i + (1 - \rho) + \epsilon_t^i, \tag{10}$$

where ρ is the first-order autocorrelation coefficient and ϵ_t^i is a mean-zero normal random variable. The long-run expected value of z_t^i is 1, and we restrict the range to $z_t^i \in [0.5, 1.5]$. When $\rho = 0$ there is no serial correlation—the process is pure white noise. Below we consider cases without and with autocorrelation: $\rho \in \{0, .95\}$.[17] The standard error of ϵ_t^i is scaled so that the variance of each z_t^i remains the same (0.1) regardless of the value of ρ. Relative to the single uncorrelated shock model first presented, the z_t^i terms are now recast as shock *levels* that incorporate the cumulative effect of stochastic innovations, ϵ_w^i, for $w \leq t$. Let $\mathbf{z_t}$ represent a vector of the four shock levels impacting biology and profit.

When there is no autocorrelation, the value function defined in equation 6 is still a function of only one state, $V(n_t)$. However, under autocorrelation, the value function becomes dependent on all four of the shock levels, $V(n_t, \mathbf{z_t})$. In this five-state variable case, the same solution algorithm is applied, albeit with each step previously applied to the scalar stock level (n_t), extended to the vector $(n_t, \mathbf{z_t})$, and with dynamics for $\mathbf{z_t}$ given by equation 10.

Results Under Autocorrelation

We focus on how the extension to a five-state, autocorrelated shock model affects optimal management, expected welfare, and fishery closures. We solve for the optimal escapement policies when autocorrelation is and is not present. We then assess the performance of these two policies over simulations in which shocks are indeed autocorrelated. We use 5,000 simulations of 70 periods in which the first 20 periods are discarded for burn-in. Simulations are started at uniformly random points in the state space. We consider unit harvest cost functions that are density dependent and density independent. For simplicity, we refer to these two cases as "dependent" and "independent", respectively. We also consider both the base growth and critical depensation growth models. Overall, against a backdrop of autocorrelation, we are interested in the performance of the autocorrelation-savvy manager (relative to an autocorrelation-naive manager) and how this depends on the growth model and cost function. We focus on base growth model (logistic) results, below, and present critical depensation model results in the online-only appendix to simplify the presentation of cases. Overall, outcomes under these alternative growth models are qualitatively similar.

Previewing welfare results, in the dependent case we find that the expected profit gain from autocorrelation-savvy management (over autocorrelation-naive) is either modest or almost zero: the mean profit (across simulations and over time) increased by 0.2%. However, we find that the expected profit gain in the *independent* case is substantial at 10%. We also find that the autocorrelation-savvy manager does a much better job of avoiding fishery closure, again especially in the independent case. Below we mainly present results for this independent case (accept where noted). Additional parallel results for the dependent case are in the online-only appendix. At the end of this section we discuss, in detail, the welfare results introduced, above, and provide intuition for the impact of the cost function, which drives the biggest differences.

For all model variations we consider—cost and growth functions and autocorrelation levels—the constant escapement policy found by Reed (1979) and others holds: optimal escapement is constant over stock levels unless the stock falls below the constant escapement level, at which point optimal escapement is 100%.[18] Thus, the optimal policy for each scenario can be summarized by the constant escapement level. In figure 5 we summarize optimal constant escapement levels with autocorrelation ($\rho = 0.95$) and without ($\rho = 0$) for a

set of shock level cases. We consider cases where shock levels are at their mean ("all $z^i = 1$") and in which various shock levels are at their minimum ("min z^i") or maximum ("max z^i"), while remaining shocks are at their mean. Finally, we depict minimum and maximum escapement ("min esc," "max esc") cases in which all shock levels are at an extreme (high or low) level. (See online-only appendix figure 2 for the dependent case.)

Under no autocorrelation (dark bars in figure 5) the optimal action depends only on the stock (x) and *economic* shock levels. Here the policy is not a function of biological shocks—they are already incorporated into x and provide no information about next period's state. Under no autocorrelation, current shock levels do not inform future shock levels and, therefore, do not affect expected future values. In contrast, under autocorrelation (light bars) all current shock levels inform future shock levels and, thus, have an effect on value and optimal management.

As expected, escapement is high when shock levels result in a low price and high harvest cost (with and without autocorrelation). Escapement is also relatively high when the carrying capacity is high and the growth rate low. Overall, shock levels drive wide variation in optimal escapement. When all shock levels are considered together the potential for variation is greater in the autocorrelation case, which is not surprising since policy is now sensitive to economic *and* biological shock levels: in the autocorrelation case, optimal escapement can vary by more than a factor of five depending on the shock level (e.g., from under 20 to over 125 at high stock levels). But when we consider the response to a high individual shock level (with other shock levels at their mean), variation in escapement is lower under autocorrelation. This is because under autocorrelation the decision maker expects the condition to persist, while under no autocorrelation the condition is transient. For example, a high price will be immediately exploited under no autocorrelation (since the expected price in future periods is the mean), while under autocorrelation the decision maker expects prices to remain high (and exploitable) over multiple periods.

Comparing the relative impact of various shock levels, optimal escapement is more sensitive to price shock levels (z^p) than cost (z^c); in part because shocks enter multiplicatively against a price parameter that is larger than the cost parameter ($p > c$). Policy is also much more sensitive to the shock level for carrying capacity (z^K) than for growth rate (z^R). This might be expected given growth rate affects only new recruits, while carrying capacity can affect both (when stock exceeds the carrying capacity).

[Figure 5 about here.]

In figure 6, given autocorrelation we show how optimal escapement share varies (shading intensity) with biological shock levels (axes) for various stocks (panels). As noted above, optimal escapement is higher given high carrying capacity (z^K) and low growth rate (z^R). Accounting for multiple sources of autocorrelated uncertainty here provides a more nuanced result than obtained by Parma (1990). In her analysis, Parma finds that "escapements are raised when favorable conditions are anticipated and they are lowered when poor environments are expected." We find a similar result with respect to carrying capacity. But in the case of growth rate, escapement is higher under $unfavorable$ conditions.

To illustrate the potential policy error from ignoring autocorrelation, we show the isocline where the (fixed) autocorrelation-naive policy ($A^*_{\rho=0}$) is equal to the autocorrelation-savvy policy. $A^*_{\rho=0}$ is reported in the subtitle to each subplot. The difference in optimal escapement share is small when the shock levels are near their mean ($z^i = 1$) and large when z^K is high and z^R is low (or vice versa). Thus, we expect gains for an autocorrelation-savvy manager (relative to naive) to be high when z^K and z^R diverge strongly from each other or from their mean.

[Figure 6 about here.]

In figure 7 we show how optimal escapement share for the autocorrelation-savvy manager varies with *economic* shock levels (axes) for various stocks (panels). Generally (as noted above) escapement share increases as economic conditions are *less favorable*, as z^p falls and z^c increases. Considering a harvest optimization problem with (uncorrelated) price and stock uncertainty, Hanson and Ryan (1998) find that price volatility has only a modest impact on the optimal policy. In contrast, we find that price fluctuations have a large impact on optimal escapement under no autocorrelation (figure 5, $\rho = 0$). Under autocorrelation ($\rho = 0.95$), the role of z^p attenuates but is still quite substantial especially for low prices. These observations also hold for the dependent case (online-only appendix figure 2).

[Figure 7 about here.]

Results in figures 6 and 7 show that the optimal policy response to z^K depends on z^R, and similarly the optimal response to z^p depends on z^c. This also holds for the dependent case (see the online-only appendix). It also holds across each shock level pair, such as z^K and z^p (not shown). Thus if autocorrelation is present in more than one parameter shock level (as modeled here), then all should be accounted for since the optimal response to one shock level depends on the others.

In table 1 we report summary statistics from 5,000 Monte Carlo simulations of 70 periods where the first 20 periods are discarded for burn-in. In each simulation, paths for each shock level, z_t^i, are produced based on a uniformly random starting point and assuming autocorrelation ($\rho = 0.95$). Then the autocorrelation-savvy ($\rho = 0.95$) and autocorrelation-naive ($\rho = 0$) policies are applied separately, but in parallel, against the same series of starting points and shock levels. We examine several outcomes: escapement share, stock, profit, harvest, and closure of the fishery. We consider both density-independent and density-dependent harvest costs. We first take the mean (over the final 50 periods) for each simulation and then, across simulations, report the mean and 90% confidence interval (in parentheses). We also show the percentage difference in the mean statistic.

[Table 1 about here.]

On average, the escapement share and stock decrease, while profit and harvest increase under the autocorrelation-savvy manager. This holds for both cost functions (and both growth models—see online-only appendix table 1). In figure 8 we show histograms of these outcomes from table 1 for the independent case, illustrating the spread of mean simulated outcomes. Each "observation" in the set summarized in a histogram is the mean from a given simulation. Dark shading represents outcomes for the autocorrelation-naive manager ($\rho = 0$), and white represents outcomes for the autocorrelation-savvy manager ($\rho = 0.95$). Muted shading appears when they overlap. In the top left panel we see that for the autocorrelation-savvy manager, the distribution of escapement shares narrows and shifts lower. Correspondingly, the distribution of harvest levels also narrows and increases (bottom left). The distribution of stock levels narrows and shifts down (top right).[19] Finally, in the bottom right we present results for enhancing profit via autocorrelation-savvy management. Gains come from both increasing returns during lean years (shrinking the left-hand tail) and from exploiting boom years (expanding right-hand tail).

[Figure 8 about here.]

The distribution of outcomes in the dependent case are qualitatively similar. The only key exception is that profit gains for the savvy manager in the dependent case are much smaller. In the base growth model case they are essentially zero (table 1). This occurs despite the fact that harvest is substantially higher (14.5%) for the savvy versus naive manager. We do not expect profit gains to be as high as harvest gains because the auto-correlation savvy manager is fishing to a lower stock level with higher density-dependent

unit harvest costs (relative to the naive manager). Still, the near zero profit gains present a puzzle that leads to a central insight in our analysis: gains to autocorrelation-savvy management are substantial in the independent case, but negligible in the dependent case because differences in the cost structure have implications for (1) avoiding instances of errant high harvest for the naive manager, and (2) facilitating instances of ideal high harvest for the savvy manager.

First, we consider the naive manager's capacity to avoid instances of errant high harvest. In the dependent case, increasing marginal harvest cost (as the stock is fished down) provides an "economic brake," saving (to some degree) the naive manager from errantly over-harvesting the stock when economic shock levels suggest high harvest (high z^P and low z^C) but, in addition, expected biological shock levels (naively ignored) suggest moderation.[20] In the independent case, there is no such economic brake impeding errant high harvest by the naive manager.[21] Figure 5 shows that more aggressive harvest (low escapement) from realizing the lowest cost (z^c) or highest price (z^p) shock levels for the naive manager ($\rho = 0$) is not ideal—this response is tempered for the savvy manager ($\rho = 0.95$). These results show that an autocorrelation-naive manager can over-respond to economic shocks—they may believe they are optimally responding to idiosyncratic shocks when they are actually errantly driving down the stock (given the ignored autocorrelated shock levels to come). This effect is also stronger without the "economic brake" of stock-dependent cost serving to limit harvest.

Second, we consider how cost functions differentially allow instances of ideal high harvest for the savvy manager. In the dependent case, increasing marginal harvest cost (in the stock) reduces the opportunity for the savvy manager to set an aggressive harvest level when shock levels might suggest it. In contrast, in the independent case, the savvy manager has the opportunity to aggressively draw down the stock when prudent (e.g., given low z^K and high z^p). For the savvy manager, more aggressive harvest in the independent case—specifically when z^K is low—can be seen in comparing the $\rho = 0.95$ case in figures 5 and online-only appendix figure 2.

As a final comparison, we consider differences in the propensity for fishery closure. These results should be interpreted with caution since we have not imposed an explicit cost of closure in our objective function (other than the effect on fishery profit).[22] However, the results illustrate how naive management might lead to excessive interruption of harvesting activity. We define closure as instances in which the escapement share exceeds 99.9%. We report the percentage of periods the fishery is closed in the final line of table 1. In all scenarios, autocorrelation-savvy management results in a substantially lower closure percentage. In the independent case, the savvy manager achieves exceedingly rare closures (0.3% of periods), while ignoring

autocorrelation leads to fishery closure in 4.4% of periods. In the dependent case closure is more frequent overall (3.5-11.0%), but savvy management reduces the rate by approximately one third.

Discussion

The OECD has argued that figuring out "how uncertainty and lack of information should be taken into account" is among the short list of most important issues to work on in fisheries (OECD 2012). However, modeling such ecosystem realities can generate "overwhelming complexities very rapidly" (Clark and Munro 2017). Standard approaches for solving dynamic resource optimization models, such as VFI, have limited capacity to handle increasingly detailed problems involving a large number of states and complex sources of multiple uncertainty. Another drawback is that they do not integrate conveniently with the prevalent use of Monte Carlo simulation analysis in the natural sciences. In contrast, ADP incorporates a larger set of continuous states and complex sources of stochasticity with ease in a familiar forward simulation framework. When uncertainty takes the form of stochastic draws observed between actions, ADP eliminates the need for numerical integration (replaced by repeated simulation).

We add to the broader ADP literature with an extension to nonparametric methods. Parametric approximation of the value function can lead to instabilities in the solution procedure and large error in the policy function. We implement a nonparametric approach to capture the value function without imposing a functional form. The flexible approach is stable and adapts to convex-concave dynamics (critical depensation growth model) and cost function alternatives with relative ease. Our code implementing ADP for a simple resource model can be adapted to the dynamics and uncertainties of other systems. Using a standard fishery management model under uncertainty, we show that ADP is reliable and accurate. Expanding on our simple base model, we show that ADP readily handles five continuous states and four sources of uncertainty.

In addition to illustrating a new optimization tool to resource economics, we contribute to the understanding of fisheries management under uncertainty. We consider uncertainty in price, cost, carrying capacity, and growth rate. In contrast to the typical assumption of idiosyncratic shocks, we allow for shock levels that are autocorrelated over time. In our base model we find that policy is sensitive to all four shocks: escapement increases under high carrying capacity, low growth rate, low price, and high cost. Policy is much more sensitive to changes in carrying capacity than other variables. These findings qualify insights from Sethi et al. (2005, 317) who conclude that, "growth...uncertainties have only a small effect on optimal policy (and) profits...even when uncertainties are high." We find that shocks to growth function that are autocorrelated have a strong impact on profit and the optimal escapement policy, which can vary by a factor of three.

We find that accounting for autocorrelation leads to large differences in response to observed shocks. For example, in our central case (density-independent harvest cost and base growth model) an autocorrelation-naive manager makes big adjustments in escapement due to price and cost fluctuations (exceeding +/-50% in some cases). In contrast, a savvy manager with knowledge that shock levels will persist somewhat makes much smaller adjustments, less than half as large. Overall, both biological and economic shock levels are important: optimal savvy policy is most sensitive to carrying capacity shock levels (in both directions) and low price shock levels. We find that accounting for multiple autocorrelation simultaneously is important since (1) the optimal response to one shock level depends on each of the others, and (2) results from the literature on a single autocorrelated shock no longer necessarily hold when additional shocks are considered. Illustrating the latter result, Parma (1990) finds that accounting for autocorrelation leads to raised escapement when favorable conditions are anticipated. In contrast, we find escapement is higher when the carrying capacity is $favorable$ and the growth rate $unfavorable$.

We also find that rents from autocorrelation-savvy management strongly depend on harvest cost structure. When harvest costs are density-*dependent*, savvy policy generates no or modest profit gains (relative to a naive manager). When harvest costs are density-*independent*, savvy policy leads to substantial profit gains. Driving this result is the fact that differences in harvest cost structure influence the degree to which (1) a naive manager avoids instances of errant high harvest; and (2) a savvy manager takes advantage of conditions for ideal high harvest.

Finally, we find that accounting for autocorrelation is important for avoiding fishery closures; the savvy manager substantially reduces the rate of closures in all cases. A naive manager, who believes they are optimally responding to idiosyncratic shocks, can over-respond to economic shocks, errantly driving down the stock when not ideal; especially without the "economic brake" of stock-dependent cost.

Two other important sources of uncertainty include state measurement error (imperfect knowledge of the state) and implementation error (imperfect achievement of the control) (Sethi et al. 2005). In such cases, the state achieved after implementing the action may no longer be deterministic (as in the post-decision state framework used in this article). ADP can still be used in such cases, but in modified form. One option is to reintroduce numerical integration in the decision step to evaluate the expectation. Another option is to add the most recent action to the vector of state variables and shift time period accounting slightly to allow shocks that occur just after a decision is made to enter at the beginning of the following period.

Outside of economics, the use of simulation models to examine environmental systems in recent decades

has followed the expansion of computing technology and interest in complex processes (Peck 2001). Natural science research has a long history of using simulations to understand the properties of diverse applications, such as fisheries, forestry, agriculture, and climate change (Petrovskii and Petrovskaya 2012). Within the economics literature, Moxnes (2003) argues that simulation models are appealingly familiar to decision makers. However, common *economic* optimization techniques used in these systems are not structured to take advantage of forward simulation. Iterative techniques, like VFI, take an approach that is either explicitly or stylistically consistent with backward induction (for finite-horizon and infinite-horizon problems, respectively).[23] Powell (2011, 233) argues that ADP should be seen as a "powerful tool for the simulation community" due to the similarity in "culture." This suggests that ADP holds strong promise for facilitating necessary collaboration between natural scientists and economists to tackle rich bioeconomic problems.

Notes

[1]The Markov transition matrix contains the probability of reaching each discrete state in the next period given the current state and action taken.

[2]Moxnes (2003) focuses on a particularly challenging form of uncertainty; measurement error leading to uncertainty in the level of the state variable when choosing the action. This kind of uncertainty has been explored more recently in several papers (Springborn and Sanchirico 2013; Kling, Sanchirico, and Wilen 2016; MacLachlan, Springborn, and Fackler 2016). While such uncertainty is not the focus here, ADP's facility with high dimensionality would be advantageous given the additional states dedicated to beliefs in these models.

[3]Specifically, Fonnesbeck (2005) uses a tabular Q-learning algorithm based on temporal difference learning.

[4]Potapov (2009, 37), inspired by Monte Carlo techniques from reinforcement learning and neuro-dynamic programming, also develops an approach for generating approximate solutions to large-scale problems. The approach is promising for a specialized set of spatially extended systems but takes many hours to arrive at a solution.

[5]Sethi et al. consider uncertainty in stock size, growth, and implementation of harvest quotas; Hanson and Ryan (1998) and Nøstbakken (2006) consider price and stock uncertainty.

[6]The marginal cost expression arises from the common assumptions of constant cost (c_E) per unit effort

(E_t) and harvest production function given by $h_t = qE_t x_t$, where q is a catchability coefficient. Taking total cost ($c_E E_t$) and the production function, we can restate total cost as $c_E E_t = (c_E h_t)/(qx_t) \equiv ch_t/x_t$.

[7]Sethi et al. (2005) assume constant (density-independent) costs in their main analysis (subsumed into a constant *net* price per unit harvest) but then consider the density-dependent costs specified in equation 2 in their sensitivity analysis. While Sethi et al. use a uniform random growth shock, we use a normal random variable (Bulte and Kooten 2001) truncated to a finite domain.

[8]Gaussian process regression is a nonparametric kernel-based approach that seeks a distribution over possible functions that are consistent with the data. For a comprehensive overview, see Rasmussen and Williams (2005).

[9]We select starting states randomly from a uniform distribution. Alternatively, starting states can be selected randomly from $\hat{n}$ or, to ensure representation of simulation chains originating across the state space, to multiple replicates of $\hat{n}$ (depending on the number of nodes in the discretization).

[10]If the state space is treated as discretized throughout, it is feasible to avoid this regression step and simply replace the left-hand side of the expression in step 3(b)iv with $\bar{V}^k(n_t^m)$ directly (e.g., algorithm 1 in Hull (2015)). Time for regression estimation is saved, but information obtained at a particular point is not then used to update points around it (Powell 2011).

[11]Others, such as Hull (2015), have found it useful to weight observations in the regression. In our application we found that weighting did not help. We assessed weighting applied as follows: (1) for each node other than the first and last—which bound the state-space—assign a target node weight ($\hat{w}_n$) of $\frac{1}{\hat{n}-1}$; (2) for the first and last nodes assign a target node weight of $\frac{1}{2*(\hat{n}-1)}$; (3) for each node, count the number of observations that is closest to that node ($\tilde{w}_n$), and finally (4) for each of the $\bar{m} * T$ observations, assign it a weight of $\frac{\tilde{w}_n}{\hat{w}_n}$, where n is the node nearest to the observation.

[12]Using the average maximum deviation over several iterations ($\bar{k} > 1$) helps avoid premature stopping that may result when a pair of regressions happen to produce similar results.

[13]A lookup table can also be used for a continuous (rather than discrete) state, whereby the table entries represent values at particular points along a continuous function.

[14]For the quadratic and quartic models, residuals are, on average, below zero at very low stock levels then above zero for some range.

[15]Comparisons were conducted using Matlab (Version 2018a) with parallel computing (eight workers) running on a PC with an Intel Core i7-4770 Quad-Core processor (3.4 GHZ), 16GB installed RAM, and a Windows 10 Enterprise operating system.

[16]In the basic Reed model a growth shock occurs after the growth function has been applied. To accommodate shocks within the growth function itself, we make a small modification to the specification of the optimization problem in equation 6. Instead of applying the growth function, G, as the last step in a period, ADP requires that we realize growth as the first step in a period since ADP allows us to avoid use of integration as long as shocks occur in a period before the decision is taken.

[17]In their respective single-shock variable models, Spencer (1997) uses $\rho = 0.6$ and Walters and Parma (1996) consider $\rho \in (0, .3, .8)$. Walters and Parma use a tighter normal distribution for shocks (variance of 0.05) but without evident truncation.

[18]Sethi et al. (2005) find non-constant escapement induced by measurement uncertainty in the fish stock. In their case this is expected, since the measurement shock is multiplicative with respect to the current stock, generating a marginal opportunity cost of harvest curve that shifts depending on the current stock level. No such mechanism for a shifting marginal opportunity cost of harvest curve is present in our setting.

[19]Given this outcome, applying the $\rho = 0$ policy when $\rho = 0.95$ could be considered a precautionary policy, on average. However, there are exceptions. For example, when all shock levels are at their mean, except the growth rate shock level is at its minimum (least productive) level, the naive manager makes no adjustment, while the savvy manager increases escapement (figure 5).

[20]The naive manager ignores the fact that current high or low biological shock levels are likely to persist.

[21]For the naive manager, more aggressive harvest in the independent case can be seen in comparing the $\rho = 0$ case of figure 5 and online-only appendix figure 2—in the independent case the minimum escapement level is around 20 compared to about 25 in the dependent case. This result is also evident across the span of z^p and z^c in online-only appendix figures 5 and 6: the escapement share is typically higher in the dependent case, especially at low stock levels.

[22]Our profit comparison also ignores the welfare benefits to *consumers* of harvest that increases under the savvy policy.

[23]Rust (1997) used Monte Carlo sampling to handle integration in estimating the value function but within

an otherwise standard VFI approach.

List of Figures

Fig. 1: Elements of the Dynamic Core of the ADP Algorithm (step 3) for Simulation m in Block k Note: Given the current value function estimate $\bar{V}^{k-1}$, the starting state $n_{t=1}^{m}$, and a stochastic shock (not pictured), the optimal action is chosen resulting in the value $\tilde{v}_{t=1}^{m}$ (first open circle). Applying the step size δ_t (to handle the expectation) provides the "data point" $\tilde{V}_{t=1}^{m}$ (first solid circle). After applying the optimal action the state updates to $n_{t=2}^{m}$ and the sequence repeats, generating additional $\tilde{V}_{t}^{m} \in \widetilde{\mathbf{V}}$ until T periods have been simulated for each block $m = 1, ..., \bar{m}$ (only the first 3 data points are shown, $t = 1, 2, and\ 3$ for block m). Regressing the states ($\mathbf{n}$) on the values ($\widetilde{\mathbf{V}}$) (solid circles) generates the updated, fitted value function $\bar{V}^{k}$.

Fig. 2: Example of the ADP Dashboard for the Base Model after $k = 96$ Regression Updates of the Value Function

Fig. 3: Simulated Value Function Draws ($\mathbf{n}$ and $\widetilde{\mathbf{V}}$) and the Model Generated by the Regression ($\bar{V^{k}}$) (A:C); Residuals (D:F) Given Parametric Models (first two columns); the Nonparametric Model (last column)

Fig. 4: For the Base Model, a Comparison of Value Functions (left) and Policy Functions (right) using the Nonparametric and Parametric Models (ADP approach) and a Lookup Table (VFI approach)

Fig. 5: Optimal Constant Escapement levels (vertical axis) for Various Shock Level Cases (horizontal axis) without Autocorrelation ($\rho = 0$, dark bars) and with Autocorrelation ($\rho = 0.95$, light bars) Note: Shock levels for the minimum escapement and maximum escapement cases ("min esc," "max esc") are at levels that minimize or maximize escapement levels (assuming medium and high stock levels). For remaining cases all shock levels are at their expected level ("all $z^i = 1$") except for the shock labeled. Harvest cost is density-independent.

Fig. 6: Optimal Escapement Share (shading intensity) as a Function of Biological Shock Levels (z^K, z^R) given Autocorrelation ($\rho = 0.95$) and Stock Levels (x) Varying from Low (top left plot) to High (bottom right plot) Note: Optimal escapement share without autocorrelation ($\rho = 0$) appears in the title of each panel ($A^*_{\rho=0}$). The dashed white line is the isocline at which $A^*_{\rho=0.95} = A^*_{\rho=0}$. Harvest cost is density-independent.

Fig. 7: Optimal Escapement Share (shading intensity) as a Function of Economic Shock Levels (z^p, z^c) with Autocorrelation ($\rho = 0.95$) and Stock Levels (x) Varying from Low (top left plot) to High (bottom right plot) Note: Harvest cost is density-independent.

Fig. 8: Comparison of Simulation Outcome Frequencies under Autocorrelation and Density-independent

Harvest Cost when Escapement Policy Applied either Accounts for Autocorrelation ($\rho = 0.95$) or Does Not ($\rho = 0$) Note: Growth occurs either according the base model (simple logistic growth). Results are based on 5,000 simulations of 70 periods in which the first 20 periods are discarded for burn-in.

References

Bertsekas, D. P. 2011. *Dynamic Programming and Optimal Control,* 3rd ed., Volume II. MIT. `http://web.mit.edu/dimitrib/www/dpchapter.pdf`.

Bulte, E. H., and G. C. van Kooten. 2001. "Harvesting and Conserving a Species when Numbers are Low: Population Viability and Gambler's Ruin in Bioeconomic Models." *Ecological Economics* 37(1):87–100.

Cai, Y., K. L. Judd, T. S. Lontzek, V. Michelangeli, and C.-L. Su. 2017. "A Nonlinear Programming Method for Dynamic Programming." *Macroeconomic Dynamics* 21(2):336–61.

Clark, C., and G. Kirkwood. 1986. "On Uncertain Renewable Resource Stocks: Optimal Harvest Policies and the Value of Stock Surveys." *Journal of Environmental Economics and Management* 13(3):235–44.

Clark, C. W., and G. R. Munro. 2017. "Capital Theory and the Economics of Fisheries: Implications for Policy." *Marine Resource Economics* 32(2):123–42.

Conrad, J. M. 2010. *Resource Economics*. Cambridge, UK: Cambridge University Press.

Costello, C., and S. Polasky. 2008. "Optimal Harvesting of Stochastic Spatial Resources." *Journal of Environmental Economics and Management* 56(1):1–18.

Deroba, J. J., and J. R. Bence. 2008. "A Review of Harvest Policies: Understanding Relative Performance of Control Rules." *Fisheries Research* 94(3):210–23.

Fonnesbeck, C. J. 2005. "Solving Dynamic Wildlife Resource Optimization Problems Using Reinforcement Learning." *Natural Resource Modeling* 18(1):1–40.

Hall, K. M., H. J. Albers, M. A. Taleghan, and T. G. Dietterich. 2018. "Optimal Spatial-Dynamic Management of Stochastic Species Invasions." *Environmental and Resource Economics* 70(2):403–27.

Hanson, F. B., and D. Ryan. 1998. "Optimal Harvesting with Both Population and Price Dynamics." *Mathematical Biosciences* 148(2):129–46.

Hull, I. 2015. "Approximate Dynamic Programming with Post-decision States as a Solution Method for Dynamic Economic Models." *Journal of Economic Dynamics and Control* 55(2015):57–70.

Judd, K. L. 1996. "Approximation, Perturbation, and Projection Methods in Economic Analysis." In *Handbook of Computational Economics*, ed. by H. Amman, D. Kendrick, and J. Rust, 509–85. The Netherlands: Elsevier Science B.V.

———. 1998. *Numerical Methods in Economics*. Cambridge, MA: MIT Press.

Judd, K. L., L. Maliar, and S. Maliar. 2011. "Numerically Stable and Accurate Stochastic Simulation Approaches for Solving Dynamic Economic Models." *Quantitative Economics* 2(2):173–210.

Kapaun, U., and M. F. Quaas. 2013. "Does the Optimal Size of a Fish Stock Increase with Environmental Uncertainties?" *Environmental and Resource Economics* 54(2):293–310.

Kling, D. M., J. N. Sanchirico, and J. E. Wilen. 2016. "Bioeconomics of Managed Relocation." *Journal of the Association of Environmental and Resource Economists* 3(4):1023–59.

LaRiviere, J., D. Kling, J. N. Sanchirico, C. Sims, and M. Springborn. 2017. "The Treatment of Uncertainty and Learning in the Economics of Natural Resource and Environmental Management." *Review of Environmental Economics and Policy* 12(1):92–112.

Lewis, T. R. 1981. "Exploitation of a Renewable Resource under Uncertainty." *Canadian Journal of Economics* 14(3):422–39.

MacLachlan, M. J., M. R. Springborn, and P. L. Fackler. 2016. "Learning about a Moving Target in Resource Management: Optimal Bayesian Disease Control." *American Journal of Agricultural Economics* 99(1):140–62.

Maliar, L., and S. Maliar. 2013. "Envelope Condition Method Versus Endogenous Grid Method for Solving Dynamic Programming Problems." *Economics Letters* 120(2):262–66.

McGough, B., A. J. Plantinga, and C. Costello. 2009. "Optimally Managing a Stochastic Renewable Resource under General Economic Conditions." *The BE Journal of Economic Analysis & Policy* 9(1):1–29.

Miller, T., C. Jones, C. Hanson, S. Heppel, O. Jensen, P. Livingston, K. Lorenzen, K. Mills, W. Patterson, P. Sullivan, and R. Wong. 2018. "Scientific Considerations Informing Magnuson-Stevens Fishery Conservation and Management Act Reauthorization." *Fisheries* 43(11):533–41.

Moxnes, E. 2003. "Uncertain Measurements of Renewable Resources: Approximations, Harvesting Policies and Value of Accuracy." *Journal of Environmental Economics and Management* 45(1):85–108.

Nøstbakken, L. 2006. "Regime Switching in a Fishery with Stochastic Stock and Price." *Journal of Environmental Economics and Management* 51(2):231–41.

———. 2008. "Stochastic Modelling of the North Sea Herring Fishery under Alternative Management Regimes." *Marine Resource Economics* 23(1):65–86.

Nøstbakken, L., and J. M. Conrad. 2007. "Uncertainty in Bioeconomic Modelling." In *Handbook of Operations Research in Natural Resources*, ed. by A. Weintraub, C. Romero, T. Bjørndal, R. Epstein, and J. Miranda, vol. 99. International Series In Operations Research & Management Science. 217–35. Boston, MA: Springer.

OECD (Organisation for Economic Co-operation and Development). 2012. *Rebuilding Fisheries: The Way Forward*. Paris, France: OECD Publishing.

Parma, A. M. 1990. "Optimal Harvesting of Fish Populations with Non Stationary Stock Recruitment Relationships." *Natural Resource Modeling* 4(1):39–76.

Peck, S. L. 2001. "Ecological Modeling: A Guide for the Nonmodeler." *Conservation in Practice* 2(4):36–9.

Petrovskii, S., and N. Petrovskaya. 2012. "Computational Ecology as an Emerging Science." *Interface Focus* 2(2):241–54.

Potapov, A. 2009. "Stochastic Model of Lake System Invasion and its Optimal Control: Neurodynamic Programming as a Solution Method." *Natural Resource Modeling* 22(2):257–88.

Powell, W. B. 2007. *Approximate Dynamic Programming: Solving the Curses of Dimensionality.* Hoboken, NJ: John Wiley & Sons.

———. 2011. *Approximate Dynamic Programming: Solving the Curses of Dimensionality,* 2nd ed. Hoboken, NJ: John Wiley & Sons.

Rasmussen, C. E., and C. K. Williams. 2005. *Gaussian Processes for Machine Learning.* Cambridge, MA: MIT Press.

Reed, W. 1979. "Optimal Escapement Levels in Stochastic and Deterministic Harvesting Models." *Journal of Environmental Economics and Management* 6(4):350–63.

Rodriguez, A. A., O. Cifdaloz, J. M. Anderies, M. A. Janssen, and J. Dickeson. 2011. "Confronting Management Challenges in Highly Uncertain Natural Resource Systems: A Robustness–vulnerability Trade-off Approach." *Environmental Modeling & Assessment* 16(1):15–36.

Roughgarden, J., and F. Smith. 1996. "Why Fisheries Collapse and What to do About It." *Proceedings of the National Academy of Sciences* 93(10):5078–83.

Rust, J. 1997. "Using Randomization to Break the Curse of Dimensionality." *Econometrica* 65(3):487–516.

Sethi, G., C. Costello, A. Fisher, M. Hanemann, and L. Karp. 2005. "Fishery Management under Multiple Uncertainty." *Journal of Environmental Economics and Management* 50(2):300–18.

Singh, R., Q. Weninger, and M. Doyle. 2006. "Fisheries Management with Stock Growth Uncertainty and Costly Capital Adjustment." *Journal of Environmental Economics and Management* 52(2):582–99.

Spencer, P. D. 1997. "Optimal Harvesting of Fish Populations with Nonlinear Rates of Predation and Autocorrelated Environmental Variability." *Canadian Journal of Fisheries and Aquatic Sciences* 54(1):59–74.

Springborn, M., and J. N. Sanchirico. 2013. "A Density Projection Approach for Non-trivial Information Dynamics: Adaptive Management of Stochastic Natural Resources." *Journal of Environmental Economics and Management* 66(3):609–24.

Sutton, R. S., and A. G. Barto. 1998. *Reinforcement Learning: An Introduction.* Cambridge, MA: MIT Press.

Taleghan, M. A., T. G. Dietterich, M. Crowley, K. Hall, and H. J. Albers. 2015. "PAC Optimal MDP Planning with Application to Invasive Species Management." *Journal of Machine Learning Research* 16(Dec):3877–903.

Walters, C. J., and R. Hilborn. 1978. "Ecological Optimization and Adaptive Management." *Annual Review of Ecology and Systematics* 9(1):157–88.

Walters, C., and A. M. Parma. 1996. "Fixed Exploitation Rate Strategies for Coping with Effects of Climate Change." *Canadian Journal of Fisheries and Aquatic Sciences* 53(1):148–58.

Appendix
Additional model parameters

We present bioeconomic model parameters in table 2, followed by ADP and VFI solution parameters in table 3.

[Table 2 about here.]

[Table 3 about here.]

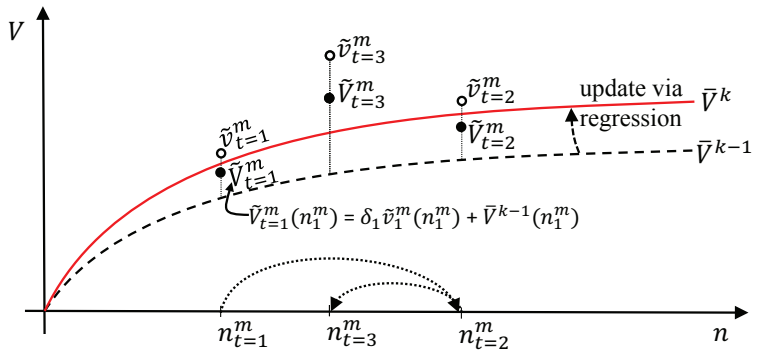

Figure 1. Elements of the Dynamic Core of the ADP Algorithm (step 3) for Simulation m in Block k
Note: Given the current value function estimate $\bar{V}^{k-1}$, the starting state $n^m_{t=1}$, and a stochastic shock (not pictured), the optimal action is chosen resulting in the value $\tilde{v}^m_{t=1}$ (first open circle). Applying the step size δ_t (to handle the expectation) provides the "data point" $\tilde{V}^m_{t=1}$ (first solid circle). After applying the optimal action the state updates to $n^m_{t=2}$ and the sequence repeats, generating additional $\tilde{V}^m_t \in \widetilde{\mathbf{V}}$ until T periods have been simulated for each block $m = 1, ..., \bar{m}$ (only the first 3 data points are shown, $t = 1, 2, and\ 3$ for block m). Regressing the states ($\mathbf{n}$) on the values ($\widetilde{\mathbf{V}}$) (solid circles) generates the updated, fitted value function $\bar{V}^k$.

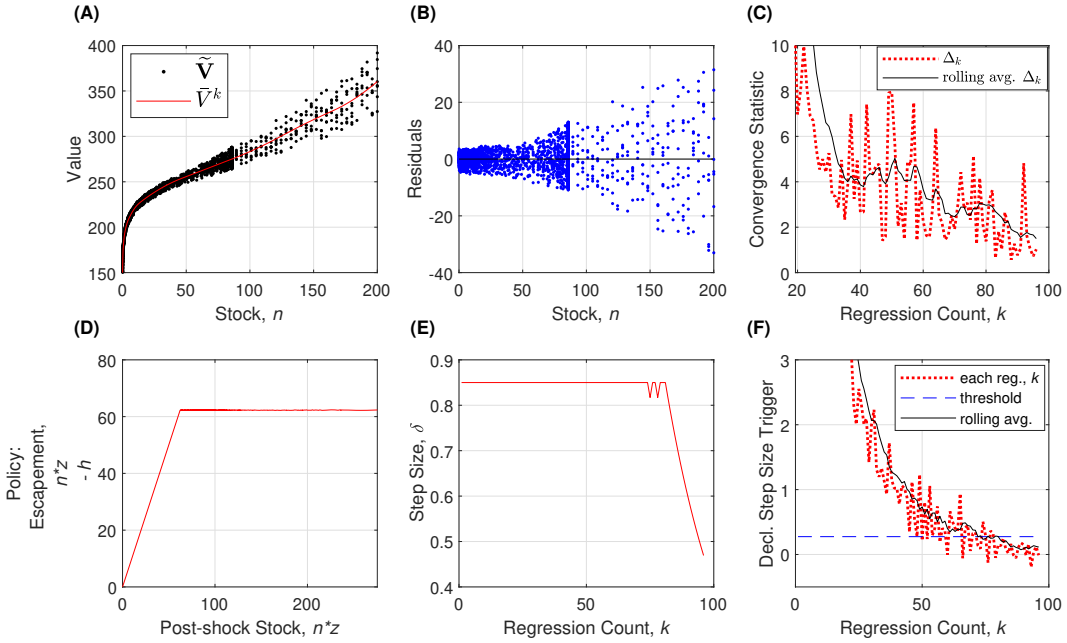

Figure 2. Example of the ADP Dashboard for the Base Model after $k = 96$ Regression Updates of the Value Function

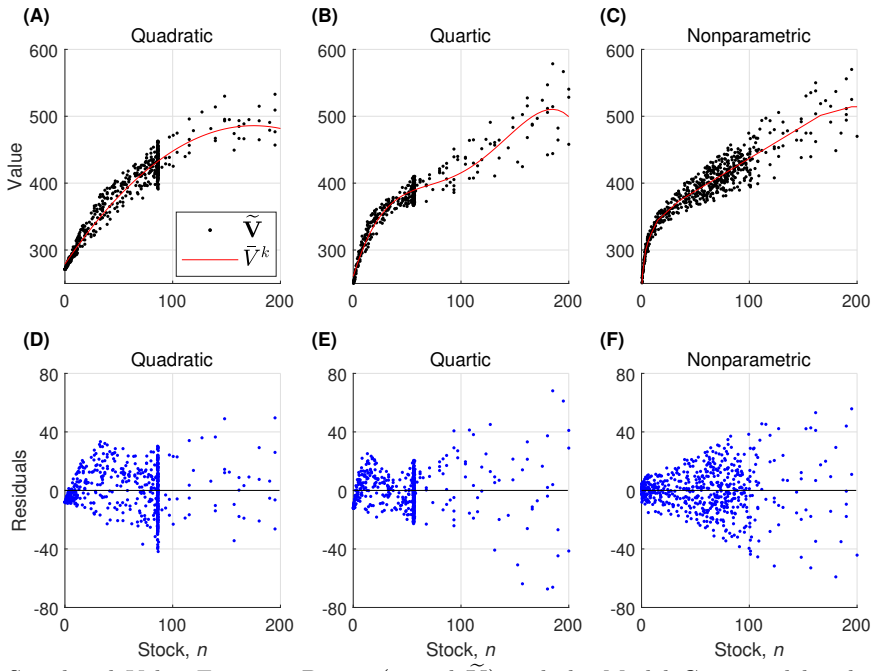

Figure 3. Simulated Value Function Draws ($\mathbf{n}$ and $\widetilde{\mathbf{V}}$) and the Model Generated by the Regression ($\bar{V}^k$) (A:C); Residuals (D:F) Given Parametric Models (first two columns); the Nonparametric Model (last column)

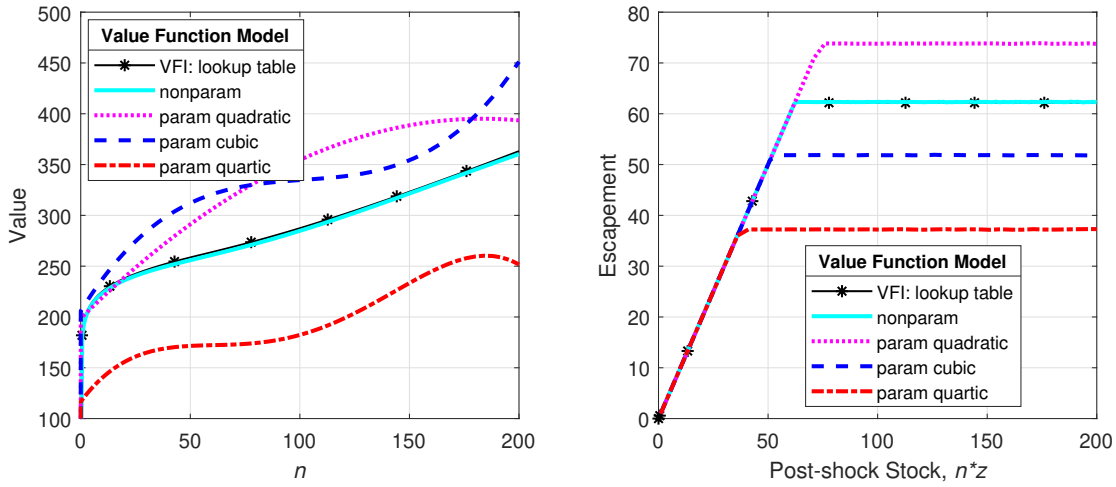

Figure 4. For the Base Model, a Comparison of Value Functions (left) and Policy Functions (right) using the Nonparametric and Parametric Models (ADP approach) and a Lookup Table (VFI approach)

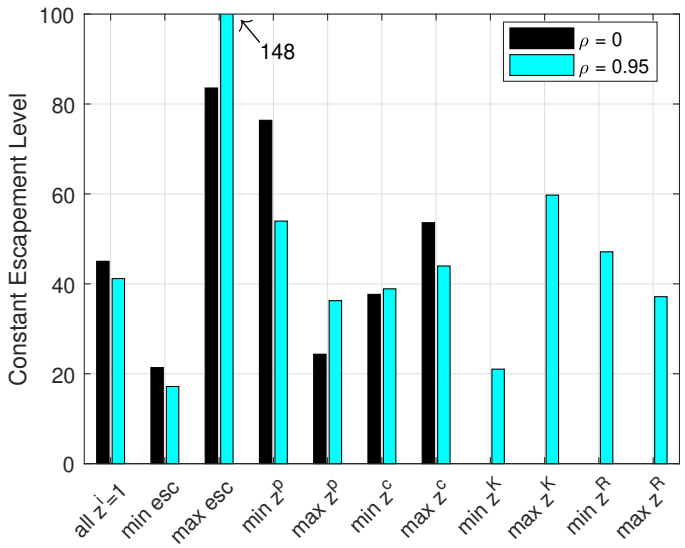

Figure 5. Optimal Constant Escapement levels (vertical axis) for Various Shock Level Cases (horizontal axis) without Autocorrelation ($\rho = 0$, dark bars) and with Autocorrelation ($\rho = 0.95$, light bars)
Note: Shock levels for the minimum escapement and maximum escapement cases ("min esc," "max esc") are at levels that minimize or maximize escapement levels (assuming medium and high stock levels). For remaining cases all shock levels are at their expected level ("all $z^i = 1$") except for the shock labeled. Harvest cost is density-independent.

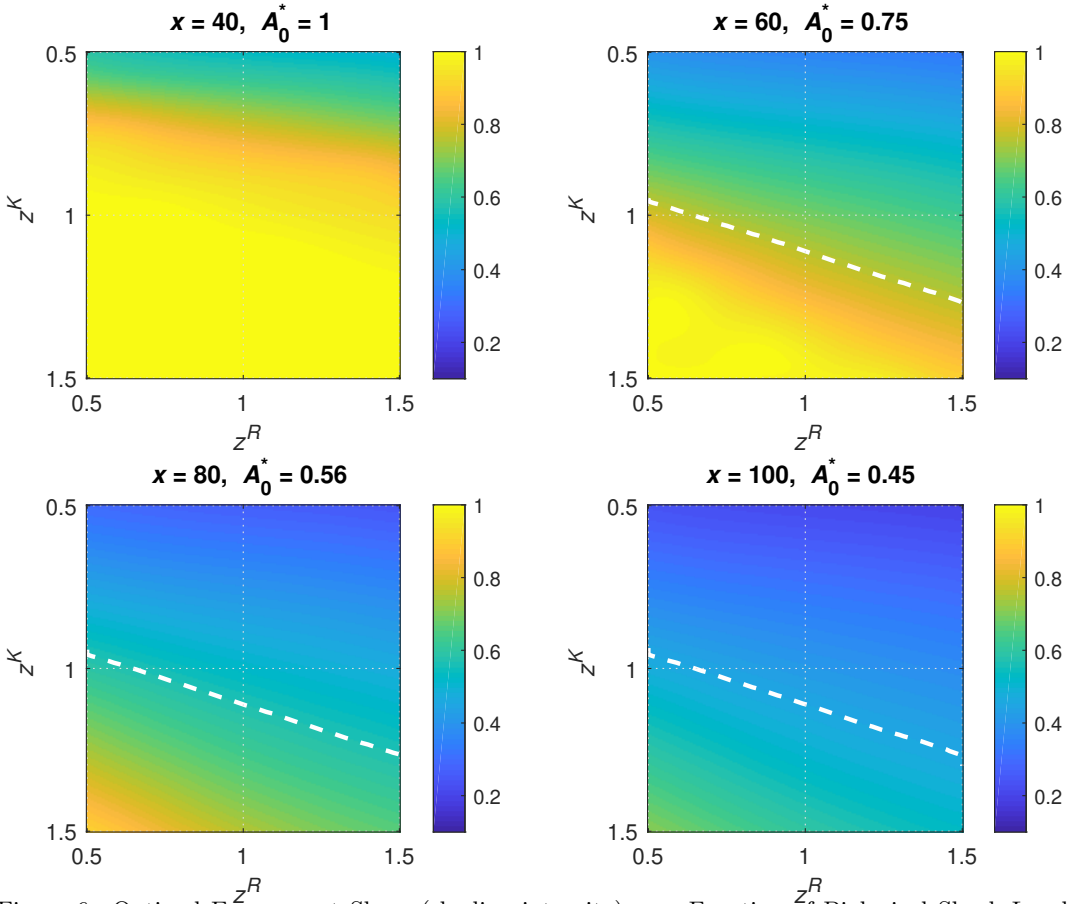

Figure 6. Optimal Escapement Share (shading intensity) as a Function of Biological Shock Levels (z^K, z^R) given Autocorrelation ($\rho = 0.95$) and Stock Levels (x) Varying from Low (top left plot) to High (bottom right plot)
Note: Optimal escapement share without autocorrelation ($\rho = 0$) appears in the title of each panel ($A^*_{\rho=0}$). The dashed white line is the isocline at which $A^*_{\rho=0.95} = A^*_{\rho=0}$. Harvest cost is density-independent.

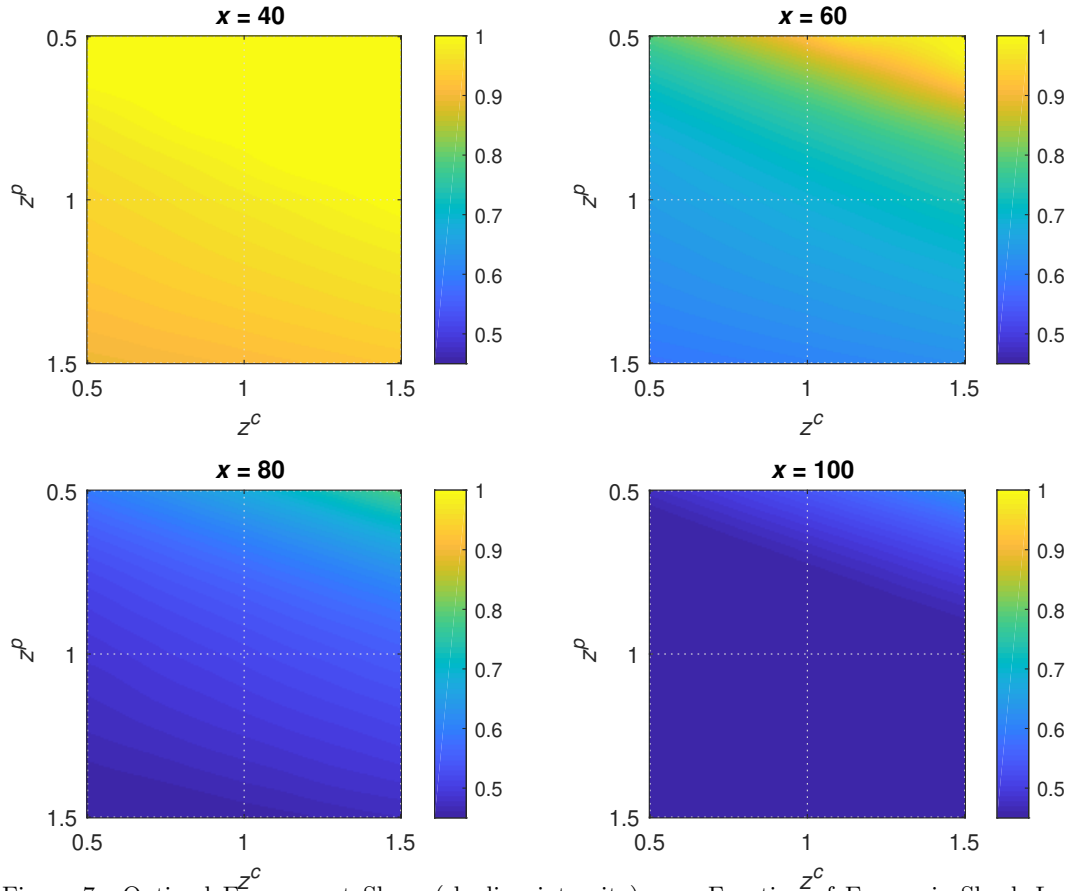

Figure 7. Optimal Escapement Share (shading intensity) as a Function of Economic Shock Levels (z^p, z^c) with Autocorrelation ($\rho = 0.95$) and Stock Levels (x) Varying from Low (top left plot) to High (bottom right plot)
Note: Harvest cost is density-independent.

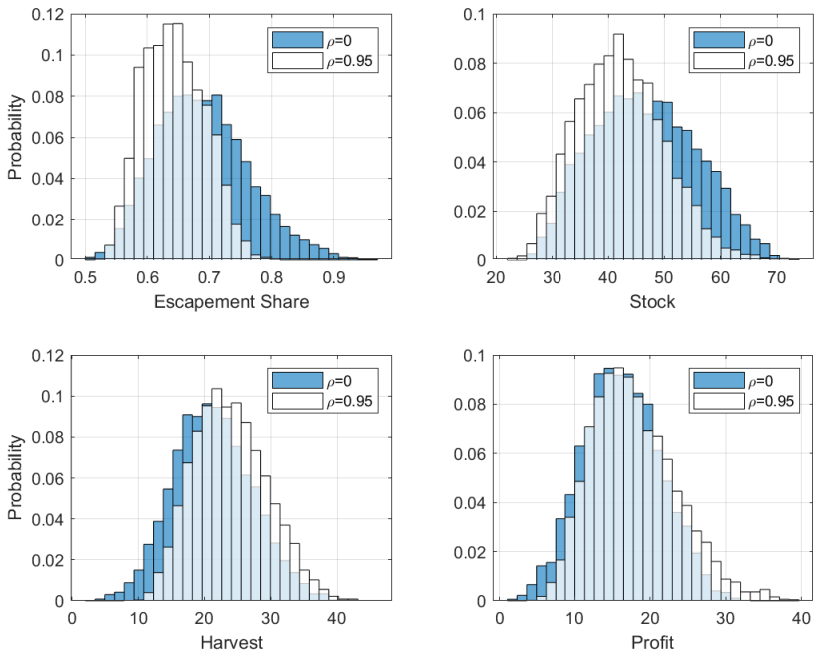

Figure 8. Comparison of Simulation Outcome Frequencies under Autocorrelation and Density-independent Harvest Cost when Escapement Policy Applied either Accounts for Autocorrelation ($\rho = 0.95$) or Does Not ($\rho = 0$)
Note: Growth occurs either according the base model (simple logistic growth). Results are based on 5,000 simulations of 70 periods in which the first 20 periods are discarded for burn-in.

Table 1. Summary Statistics for 5,000 Monte Carlo Simulations with Autocorrelation ($\rho = 0.95$) in Shock Levels

	Harvest Cost Model					
	Density-independent			Density-dependent		
	Naive	Savvy	Diff. (%)	Naive	Savvy	Diff. (%)
Escapement share	0.69 (0.49, 0.98)	0.64 (0.51, 0.82)	-6.6	0.76 (0.55, 1)	0.70 (0.55, 0.91)	-7.58
Stock	46.5 (25.8, 72)	42.3 (23.5, 63.4)	-9.1	59.0 (36.6, 82.2)	51.9 (32.2, 77.1)	-12.0
Profit	16.2 (0.5, 33.2)	17.8 (4.5, 35.7)	10.0	11.0 (0, 27.3)	11.0 (0.3, 27.8)	0.2
Harvest	21.3 (1.1, 40.8)	23.9 (8.9, 40.9)	12.1	20.1 (0, 41.5)	23.0 (5.3, 41.7)	14.5
Closure freq.	4.4%	0.3%		11.0%	3.5%	

Note: Escapement policy applied is either autocorrelation-savvy (assumes $\rho = 0.95$) or naive (assumes $\rho = 0$). Growth occurs according to the base model (simple logistic growth). Harvest cost is either density-independent or density-dependent. From each simulation mean (taken over time), reported statistics include the mean and 90% confidence intervals (in parentheses) across simulations. The difference (percentage) reflects the relative change in expected outcomes due to shift from autocorrelation-naive to autocorrelation-savvy management.

Appendix Table 2. Bioeconomic Model Parameters with Sources

	Description	**Value**	**Source**
p	price per unit harvest	1	Sethi et al. (2005)
c	harvest cost coefficient, density-dependent (independent)	0.75 (35)	Assuming costs are typically 1/3 of revenues
β	discount factor	1/(1.05)	Sethi et al. (2005)
R	intrinsic growth rate	1	Sethi et al. (2005)
K	carrying capacity	100	Sethi et al. (2005)
K_0	critical population level	25	Conrad (2010)
	shock bounds	$z_t \in [0.5, 1.5]$	Sethi et al. (2005)
μ	shock mean	1	
σ^2	shock variance	0.1	

Note: Blank indicates assumed value.

Appendix Table 3. Illustrative ADP and VFI Solution Parameters for Results Presented in Figure 4

	Description	**Value**
T	Simulation time horizon	2
$\bar{m}$	Number of simulations between regressions	570
$\{\delta_{min}, \delta_{max}\}$	Step size bounds	$\{1 \times 10^{-3}, 0.85\}$
γ	Rate of step size decline	2×10^{-4}
$\bar{k}$	Number of regression updates considered in convergence metric	10
	ADP stopping tolerance for convergence	4×10^{-3} (0.325)
	VFI stopping tolerance for convergence	1×10^{-4}

Note: When parameters used for critical depensation model differ, they are provided in parentheses.

Online Appendix
Additional ADP specifications

For the declining step size function, we use

$$\delta_t = \max\{\delta_{max} \exp(-\gamma(m_{tot} - m_{dec})), \delta_{min}\},$$

where δ_{max} is the initial weight given to new information, m_{dec} is the simulation counter value when the switch to the declining function occurs, m_{tot} is the simulation counter, γ is the rate at which the weight decays as simulations accrue, and δ_{min} is the step size lower bound. As our specific trigger to switch to a declining step size, we consider the mean absolute deviation in the value function estimate between updates (across nodes in the state space), i.e. the mean of the vector $|\bar{V}^k(\hat{n}) - \bar{V}^{k-1}(\hat{n})|$, as a percentage of the mean of the value estimate vector $\bar{V}^{k-1}$. The switch is made once the rolling average of this metric over the last 6 updates falls below 0.275%. This threshold is identified by running the ADP algorithm and setting this trigger at the point the metric stops falling, which captures the point at which the value function estimate is no longer moving consistently to higher or lower values.

As a practical matter we discretize the action space, using a vector of possible harvest levels. However, because we face no memory constraints—as we would from large, multi-dimensional arrays in a value or policy function iteration approach—we can use an exceedingly dense discretization (e.g. 1,000 nodes). Additional code and computing time would enable an explicitly continuous action space but would provide no additional utility over such a dense discretization. In step 2a of the ADP algorithm an initial guess for the value function must be set before the first block of simulations and nonparametric regression has been run. We use a simple linear function for this initial guess.

Additional value and policy function results: growth with critical depensation

Paralleling value and policy function solutions for the base growth model in the main text, figure 4, here we present these functions for the critical depensation case.

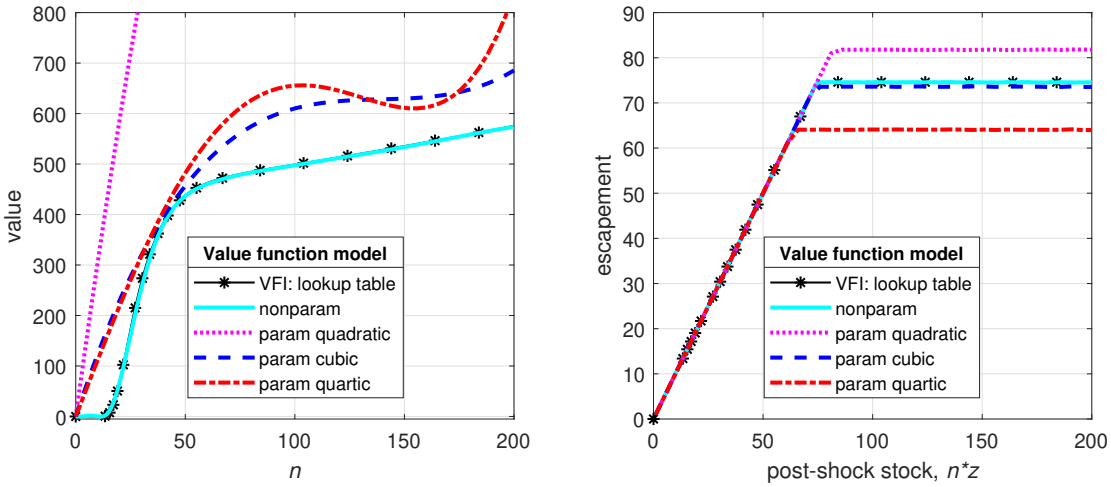

Figure 1. For the critical depensation model, a comparison of value functions (left) and policy functions (right) using the nonparametric and parametric models (ADP approach) and a lookup table (VFI approach).

Additional simulation results: growth with critical depensation

Paralleling table 1 in the main text for the base growth model (in the results under autocorrelation section of the main text) in this section we present simulation results for the critical depensation case.

Table 1. Summary Statistics for 5,000 Monte Carlo Simulations Given High Autocorrelation ($\rho = 0.95$) in Shock Levels

	Harvest Cost Model					
	Density-independent			Density-dependent		
	Naive	Savvy	Diff. (%)	Naive	Savvy	Diff. (%)
Escapement share	0.68 (0.41, 1)	0.66 (0.4, 0.96)	-2.6	0.70 (0.42, 1)	0.68 (0.41, 1)	-2.51
Stock	62.6 (51.8, 71.4)	59.4 (35.5, 72.7)	-5.2	65.7 (50.5, 75.3)	62.4 (38.3, 76.3)	-4.9
Profit	26.5 (0, 70.3)	28.4 (0.9, 81.1)	7.1	22.8 (0, 67.3)	23.8 (0, 77)	4.5
Harvest	35.9 (0, 90.2)	37.5 (1.4, 100.1)	4.5	36.8 (0, 96.1)	38.1 (0, 102.4)	3.4
Closure freq.	10.4%	3.5%		14.3%	10.1%	

Note: Escapement policy applied is either autocorrelation-savvy (assumes $\rho = 0.95$) or naive (assumes $\rho = 0$). Growth occurs with critical depensation. Harvest cost is either density-independent or density-dependent. From each simulation mean (taken over time), reported statistics include the mean and 90% confidence intervals (in parentheses) across simulations. The difference (percentage) reflects the relative change in expected outcomes due to shift from ignoring to accounting for autocorrelation.

Additional base growth model results under autocorrelation: density-dependent harvest cost

Paralleling figures for density-independent harvest cost (in the results under autocorrelation section of the main text) in this section we present results for density-dependent harvest cost.

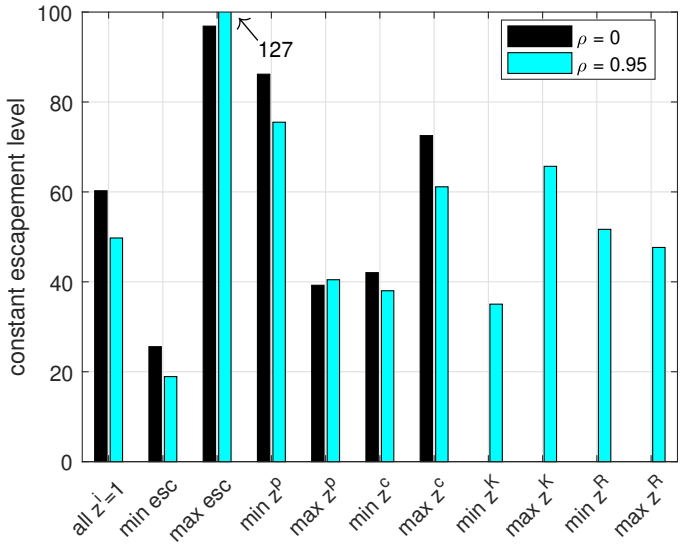

Figure 2. Optimal constant escapement Levels (vertical axis) for various shock level cases (horizontal axis) without autocorrelation ($\rho = 0$, dark bars) and with autocorrelation ($\rho = 0.95$, light bars). Shock levels for the minimum escapement and maximum escapement cases ("min esc", "max esc") are at levels that minimize or maximize escapement levels (assuming medium and high stock levels). For remaining cases all shock levels are at their expected level ("all $z^i = 1$") except for the shock labeled. Harvest cost is density-dependent.

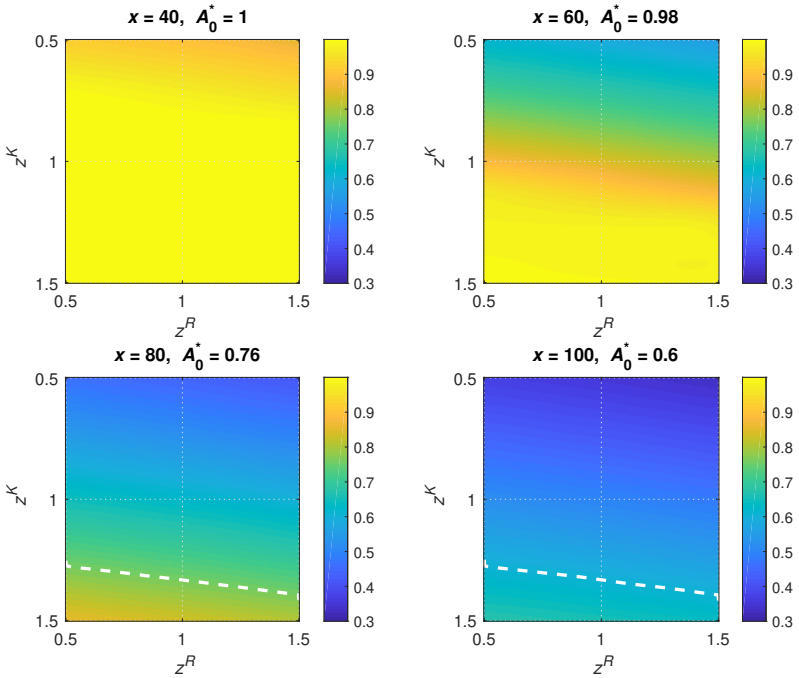

Figure 3. Optimal escapement share (shading intensity) as a function of biological shock levels (z^K, z^R) with autocorrelation ($\rho = 0.95$), base growth model, density-dependent harvest cost, and stock levels (x) varying from low (top left plot) to high (bottom right plot). Optimal escapement share for the no serial correlation case ($\rho = 0$) appears in the title of each figure ($A^*_{\rho=0}$). The dashed white line is the isocline at which $A^*_{\rho=0.95} = A^*_{\rho=0}$, except where $A*_0$ is close to 1.

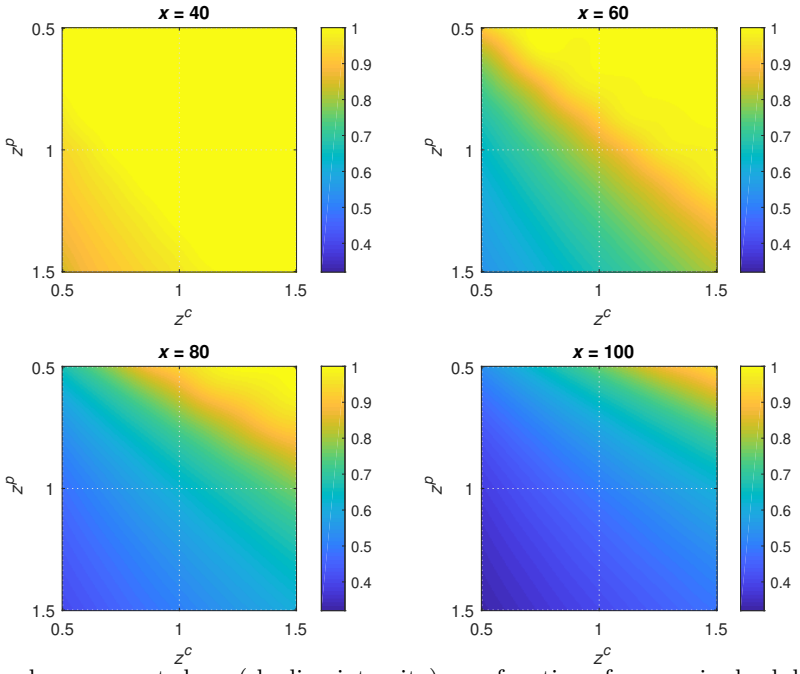

Figure 4. Optimal escapement share (shading intensity) as a function of economic shock levels (z^p, z^c) with autocorrelation ($\rho = 0.95$), base growth model, density-dependent harvest cost, and stock levels (x) varying from low (top left plot) to high (bottom right plot).

Additional results for policy functions under no autocorrelation

In this section we present additional escapement policy function results under no autocorrelation with density-independent harvest cost (figure 5) and density-dependent harvest cost (figure 6).

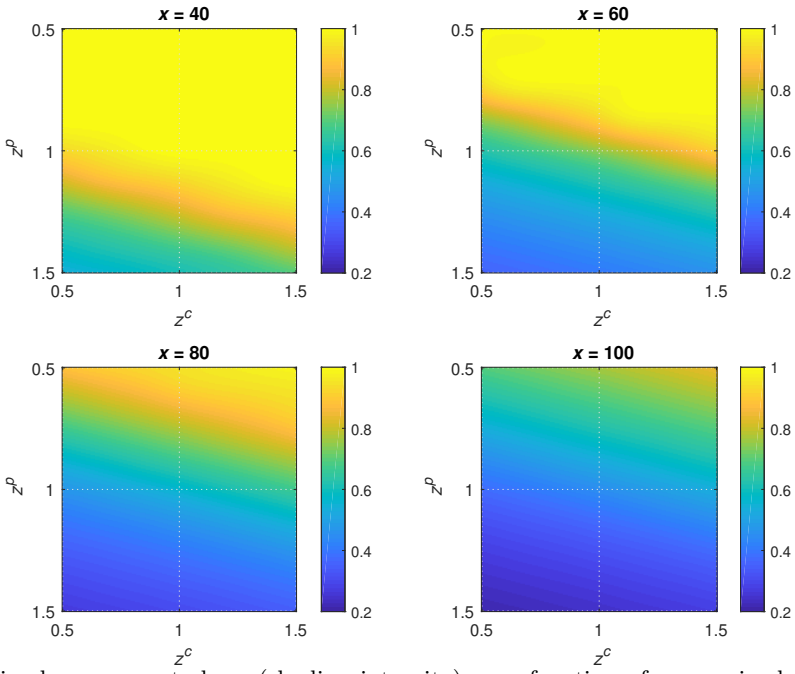

Figure 5. Optimal escapement share (shading intensity) as a function of economic shock levels (z^p, z^c) given *no* autocorrelation ($\rho = 0$), base growth model, density-independent harvest cost, and stock levels (x) varying from low (top left plot) to high (bottom right plot).

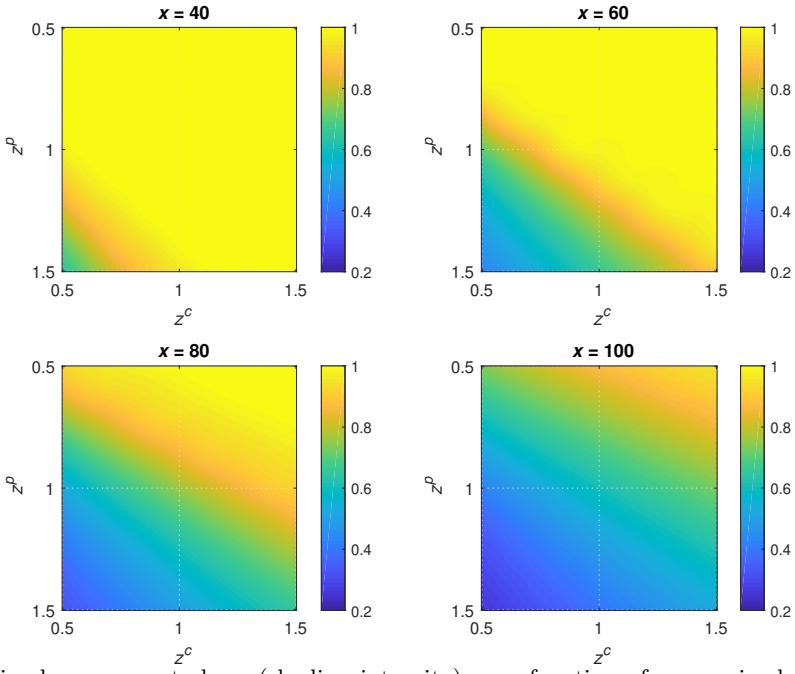

Figure 6. Optimal escapement share (shading intensity) as a function of economic shock levels (z^p, z^c) given *no* autocorrelation ($\rho = 0$), base growth model, density-dependent harvest cost, and stock levels (x) varying from low (top left plot) to high (bottom right plot).