Federal Reserve Bank of Cleveland Working Paper Series

Applications of Markov Chain Approximation Methods to Optimal Control Problems in Economics

Thomas M. Phelan and Keyvan Eslami

Working Paper No. 21-04R

May 2022

Suggested citation: Phelan, Thomas M., and Keyvan Eslami. 2022. "Applications of Markov Chain Approximation Methods to Optimal Control Problems in Economics." Working Paper No. 21-04R. Federal Reserve Bank of Cleveland. https://doi.org/10.26509/frbc-wp-202104r.

Federal Reserve Bank of Cleveland Working Paper Series
ISSN: 2573-7953

Applications of Markov Chain Approximation Methods to Optimal Control Problems in Economics*

Thomas Phelan
Federal Reserve Bank of Cleveland

Keyvan Eslami
Ryerson University

May 13, 2022

Abstract

In this paper we explore some benefits of using the finite-state Markov chain approximation (MCA) method of Kushner and Dupuis (2001) to solve continuous-time optimal control problems in economics. We first show that the implicit finite-difference scheme of Achdou et al. (2022) amounts to a limiting form of the MCA method for a certain choice of approximating chains and policy function iteration for the resulting system of equations. We then illustrate that relative to the implicit finite-difference approach, using variations of modified policy function iteration to solve income fluctuation problems both with and without discrete choices can lead to an increase in the speed of convergence of more than an order of magnitude. Finally, we provide several consistent chain constructions for stationary portfolio problems with correlated state variables, and illustrate the flexibility of the MCA approach by using it to construct and compare two simple solution methods for a general equilibrium model with financial frictions.

JEL Codes: C63, E00, G11.

Keywords: Markov chain approximation, Dynamic programming, numerical methods, financial frictions.

*The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of Cleveland or the Federal Reserve System. We appreciate comments from Ben Craig, Jesus Fernandez-Villaverde, Pablo Guerron, Ben Moll, Sergio Ocampo, Guillaume Sublet and seminar participants at various conferences and universities. All remaining errors are our own. Replication materials may be found at https://github.com/tphelanECON/EslamiPhelan_MCA.

1 Introduction

Dynamic optimization problems are ubiquitous in economics, and since closed-form expressions for such problems are available only in isolated special cases, quantitative work requires the use of numerical methods for their solution. In this paper we solve a number of dynamic optimization problems that arise naturally in economic applications by employing the Markov chain approximation (MCA) method of Harold Kushner and Paul Dupuis.[1] The method has several advantages over alternative approaches to continuous-time optimization problems that remain unexploited.[2] To the best of our knowledge, this paper is the first to outline such advantages by means of examples taken from the economics literature.

The most common approach to solving continuous-time optimization problems is the method of finite-differences, which has recently been applied to a number of economic environments by Achdou et al. (2022). In this method, one first employs recursive arguments to establish that the value function is a (viscosity) solution of a partial differential equation known as the Hamilton-Jacobi-Bellman equation, before replacing derivatives with quotients and solving the ensuing finite system of equations. In contrast, the MCA method approximates the solution to the continuous-time control problem by replacing it with a problem in which the state evolves according to a Markov chain assuming finitely many values, and applies discrete-time arguments to this latter problem.

The validity of the approximation method is based on the intuitive idea that if the discretized process is "close" to the original process, then the value function of the discrete problem will be close to the original value function. The criteria necessary for the convergence of the value function of the discrete problem to that of the original problem are referred to as *local consistency* conditions. These amount to the requirement that the increments of the chain possess the first- and second-order conditional moments of the original process, at least up to a term that is second-order in the time increment. One benefit to proceeding in this manner is that arguments from discrete-time dynamic programming already familiar to economists — such as the contraction mapping theorem and Blackwell's conditions — are applicable to this discrete problem and ensure the convergence of various well-known numerical algorithms. Further, even in the presence of non-convexities and in multiple dimensions, the Markov chain may often be chosen so as to eliminate the need for costly root-finding, without sacrificing the global convergence of the algorithm.

In this paper, we first establish a connection between the above approaches by showing that a limiting case of one widely used finite-difference scheme is equivalent to a particular case of the MCA method. Formally, we show that a limiting version of the implicit finite-

[1] For a textbook treatment see Kushner and Dupuis (2001).

[2] For earlier examples of economic applications of these methods see, e.g., Barczyk and Kredler (2014b), Barczyk and Kredler (2014a), and Golosov and Lucas Jr (2007).

difference scheme of Achdou et al. (2022) is equivalent to using the MCA method for a certain chain with negligible timestep and solving the resulting Bellman equation using policy function iteration. Establishing this connection shows that the former algorithm amounts to making two choices, a choice of chain and a choice of solution method, neither of which may be optimal for a given problem. The two applications in this paper illustrate the benefits of tailoring the solution method for the discrete Bellman equation to the problem at hand.

The first class of applications shows the benefits of departing from policy function iteration. It is well known that policy function iteration converges at a quadratic rate near the solution, and so typically requires a small number of iterations for convergence. However, updating the value function requires solving a linear system of equations, which becomes very costly computationally as either the number of gridpoints increases or the sparsity structure of the matrix becomes more complex. One therefore expects that the implicit finite-difference method will slow down rapidly as either the number of gridpoints or the dimension of the state variable increases. Section 3 explores this idea by considering variations of a problem common in economics, in which an infinitely lived risk-averse agent makes a consumption-savings choice in the presence of idiosyncratic risk and/or discrete choices over a durable good. We show that for standard parameters and moderate grid sizes, variations of the *modified* policy function iteration of Puterman and Shin (1978) can lead to an increase in the speed of convergence of more than an order of magnitude relative to policy function iteration, for the same degree of tolerance between successive iterations. Further, we provide a novel variation of modified policy function iteration that remains convergent even when the timestep vanishes, and, as such, is well-suited to discrete-choice problems in which the state may transition instantaneously.

Our second application illustrates the superiority of policy function iteration over value function iteration for two-dimensional stationary portfolio problems by comparing two algorithms for computing competitive equilibria in a macrofinance model with time-varying volatility. Both algorithms consist of two distinct steps. The first is similar to the "static step" of Brunnermeier and Sannikov (2016), and computes policy functions given continuation values before imposing market-clearing and consistency between individual and aggregate laws to update all equilibrium quantities. The second step then takes prices and the aggregate law of motion as given, and updates the continuation values. The purpose of this example is to again illustrate that the choice of chain is separate from the choice of solution method for the discrete problem, and that the practitioner may choose the latter to exploit features of the given problem. Indeed, although we draw upon the finite-difference literature to construct our chains, for the stationary problems appearing in these models, convergence appears more rapid and stable if one uses policy function iteration instead of

value function iteration to update the value function. Employing value function iteration yields an algorithm that is similar to the "false transient" approach used in Bonnans et al. (2004) and d'Avernas and Vandeweyer (2019) and the "iterative method" of Brunnermeier and Sannikov (2016), and requires a delicate choice of timestep that may be avoided when time does not explicitly enter the individual problems. As with the rest of the literature, we are unable to establish convergence of our algorithm to the competitive equilibrium. However, since we show how to efficiently solve individual portfolio problems, we believe that our policy iteration algorithm (or extensions thereof) will prove useful for applications similar to those surveyed in Brunnermeier and Sannikov (2016).[3]

The two classes of examples we present are intended to demonstrate the flexibility that the Markov chain approximation method provides to the modeller. However, they are by no means the only such avenues for flexibility. Multi-grid methods, the Gauss-Seidel algorithm, and control-dependent interpolation intervals, just to name a few, are other refinements that can be readily deployed in either the construction of the approximating chain or the numerical method used to solve the resulting Bellman equation. The purpose of this paper is to show that in the context of continuous-time optimization problems, substantial gains in speed can be obtained simply by judiciously choosing between value function iteration, policy function iteration, and variations of modified policy function iteration, techniques that are likely already familiar to most economists from the discrete-time theory.

2 Motivating example

In this section we outline the MCA method in the context of the stochastic one-sector neoclassical growth model. Although this example may be easily solved via a number of different numerical methods, it serves to give an intuitive account of how the method works and to contrast it with the finite-difference method. As we noted in the introduction, the basic idea here is to approximate the solution by solving a problem in which the state evolves according to a chain that assumes only finitely many values. The value function associated with this problem will be a good approximation to the original value function if for any given control vector, the increments of the chain share the same first and second conditional moments as the original process.

[3]Note that the possibility of gains in speed from such an approach is explicitly conjectured on page 1541 of Brunnermeier and Sannikov (2016).

2.1 Setup

Suppose that a social planner wishes to maximize the expected lifetime utility of an infinitely lived representative agent with the following preferences over consumption

$$U(c) = \mathbb{E}\left[\rho \int_0^\infty e^{-\rho t} u(c_t) dt\right].$$

We assume that capital and consumption goods may be costlessly transformed into one another. The sole state variable is then the capital stock, which evolves according to

$$dk_t = \mu(k_t, c_t)dt + \sigma(k_t)dZ_t, \tag{1}$$

where $Z = (Z_t)_{t\geq 0}$ is a standard Brownian motion, σ is a smooth function, and $\mu(k, c) = f(k_t) - \delta k_t - c_t$ for some smooth function f and constant $\delta > 0$. For simplicity suppose σ vanishes outside of some interval $[\underline{k}, \overline{k}]$ where $\overline{k} > \underline{k} > 0$ and $f(\underline{k}) > \delta \underline{k}$, and that at these boundary points consumption must satisfy $\mu(\underline{k}, c) \geq 0$ and $\mu(\overline{k}, c) \leq 0$. Given any $k_0 \in [\underline{k}, \overline{k}]$ a natural way to solve this problem is to replace (1) with

$$k_{t+\Delta_t} = k_t + \mu(k_t, c_t)\Delta_t + \sqrt{\Delta_t}\sigma(k_t)X_t \tag{2}$$

for some $\Delta_t > 0$, where $(X_t)_{t=0}^\infty$ is an i.i.d. sequence of random variables with mean zero assuming the values ± 1. Standard dynamic programming arguments, such as those outlined in Stokey (1989), imply that the value function for the discrete-time problem is the unique fixed point of the functional equation $B[V] = V$, where

$$B[V](k) = \max_{c\in\Gamma(k)} \Delta_t \rho u(c) + e^{-\rho\Delta_t}\mathbb{E}[V(k + \mu(k, c)\Delta_t + \sqrt{\Delta_t}\sigma(k)X)] \tag{3}$$

where $\Gamma(\underline{k}) = [0, f(\underline{k}) - \delta\underline{k}]$, $\Gamma(\overline{k}) = [f(\overline{k}) - \delta\overline{k}, \overline{c}]$ and $\Gamma(k) = [\underline{c}, \overline{c}]$ otherwise, where $\underline{c}$ and $\overline{c}$ are any values that ensure that capital never leaves $[\underline{k}, \overline{k}]$. Further, one may show that B is a contraction on the space of bounded continuous functions on $[\underline{k}, \overline{k}]$, and so the fixed point may be found by applying it repeatedly to any initial guess.

The finite-state Markov chain method of Kushner and Dupuis (2001) approximates the original problem in a fundamentally different way. Instead of (2), we consider an optimal control problem in which the capital stock assumes values in a finite grid $S := \{\underline{k}, \underline{k} + \Delta_k, \ldots, \overline{k} - \Delta_k, \overline{k}\}$, where $\Delta_k = (\overline{k} - \underline{k})/N$ for some $N > 1$. We construct a Markov chain such that the increments possess the same conditional mean and variance as (2) as follows. If $c_t = c$, then the increments of the Markov chain at $k_t = k$ are supported on

$\{k - \Delta_k, k, k + \Delta_k\}$ with probabilities

$$\begin{aligned} p(k, k \pm \Delta_k, c) &= \frac{\Delta_t}{\Delta_k^2}\big(\sigma^2(k)/2 + \Delta_k \mu(k,c)^{\pm}\big) \\ p(k,k,c) &= 1 - p(k, k - \Delta_k, c) - p(k, k + \Delta_k, c) \end{aligned} \tag{4}$$

where $x^{\pm} := \max\{\pm x, 0\}$ for $x \in \mathbb{R}$. Associated with (4) we have the Bellman operator

$$\tilde{B}[V](k) = \max_{c \in \Gamma(k)} \Delta_t \rho u(c) + e^{-\rho \Delta_t} \mathbb{E}\big[V(k')\big], \tag{5}$$

and so using the stochastic dynamic programming theory in Stokey (1989), one may show that the operator in (5) is a contraction and so the fixed point in the space of functions on S may be found by iterating successively on an arbitrary guess. Now note that the conditional mean and variance of the increment $\Delta k_t := k_{t+\Delta_t} - k_t$ are $\mathbb{E}[\Delta k_t | k_t = k] = \Delta_t \mu(k,c)$ and $\mathbb{E}[(\Delta k)^2 | k_t = k] - \mathbb{E}[\Delta k | k_t = k]^2 = \Delta_t\big(\sigma(k)^2 + \Delta_k |\mu(k,c)| - \Delta_t \mu(k,c)^2\big)$, respectively. Indeed, the probabilities in (4) were chosen for precisely this purpose: the σ terms ensure that the chain has the right variance, and their symmetry ensures that they do not affect the mean, while the μ terms ensure that the chain has the right mean. One may then use weak convergence arguments to show that as $\Delta_t, \Delta_k \to 0$ the fixed points of the operators defined in (3) and (5) both converge to the value function of the original problem.

So why is this construction useful, given that both (3) and (5) lead to a discrete-time Bellman equation? The main point here is that in the second discretization, when the agent contemplates the effect of varying her consumption, she need only compare *local* payoffs; the shape of the value function is irrelevant. Using (4) and omitting terms independent of consumption, for $j = 0, 1 \ldots, N$ the problem of the consumer at k_j is

$$\max_{c \in \Gamma(k)} \rho u(c) + e^{-\rho \Delta_t}\big(\mu(k_j, c)^+ V_F(k_j) - \mu(k_j, c)^- V_B(k_j)\big) \tag{6}$$

where $V_F(k_j) = [V(k_{j+1}) - V(k_j)]/\Delta_k$ and $V_B(k_j) = [V(k_j) - V(k_{j-1})]/\Delta_k$ denote forward and backward differences. The crucial difference between (3) and (6) is that in the latter, optimal consumption is available in closed-form, regardless of the shape of either the value function or the production function. To illustrate, suppose that utility is logarithmic and that the production function is $f(k) = \max\{k^{1/3}, 5(k-10)^{1/3}\}$, which is non-concave and has a "kink" at $k = 10$. Figure 1 shows the computed consumption and drift for the parameters $(\rho = 0.1, \delta = 0.075, \underline{k} = 1, \overline{k} = 80, N = 1000, \Delta_t = 10^{-6})$, for both a deterministic ($\sigma = 0$) and a stochastic ($\sigma = 0.2$ in interior, vanishing at boundaries) case. We used policy function iteration with a tolerance in the supremum norm between iterations of 10^{-6}. In both cases, convergence occurs in less than 0.03 seconds using Python and the standard scipy sparse solver (scipy.sparse.linalg.spsolve) to update the value function on an Intel Core i7-8650U

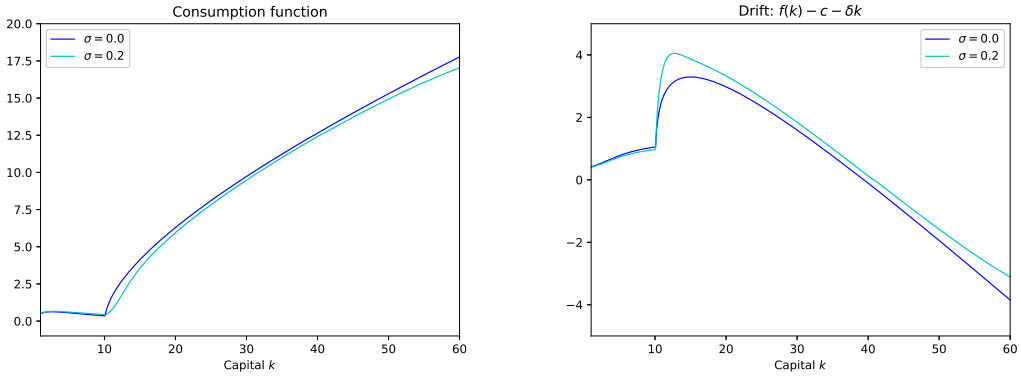

Figure 1: Policy function and drift in a non-concave one-sector growth model

processor, beginning with an initial guess of zero saving.

We now compare the above approach with a class of finite-difference methods that have been applied to a number of economic problems of interest by Achdou et al. (2022), who in turn build upon the earlier application of Candler (2001). We first (heuristically) derive the appropriate partial differential equation in order to motivate the algorithm.[4] By the Principle of Optimality, for any $t, h > 0$ we have

$$V(k,t) = \max_c \mathbb{E}\left[\rho \int_t^{t+h} e^{-\rho(s-t)} u(c(s))ds + e^{-\rho h} V(k(t+h), t+h)\right].$$

Subtracting $V(k,t)$ from both sides, dividing by h and using Ito's lemma gives

$$0 = \max_{c \in \Gamma(k)} \rho u(c) + \mu(k,c)\frac{\partial V}{\partial k} + \frac{\sigma(k)^2}{2}\frac{\partial^2 V}{\partial k^2} + \frac{\partial V}{\partial t} - \rho V.$$

A common approach to solving the above partial differential equation is to approximate the partial derivatives with various choices of difference quotients and solve the resulting finite system of equations. Achdou et al. (2022) proceed in this manner and focus primarily on what they term the *implicit* method. To understand the algorithm, define a rectangular grid $S = \{k_0, \ldots, k_N\} \times \{t_M, \ldots, t_0\}$ for the domain $[\underline{k}, \overline{k}] \times [0, t_0]$ for some $t_0 > 0$ with constant increments Δ_k and Δ_t in each dimension, and imagine we are given a terminal value $V(k, t_0)$. For each $(k_j, t_n) \in S$ write $V^n(k_j) = V(k_j, t_n)$ and define V^{n+1} as the

[4] For details we refer to Achdou et al. (2022) and Tourin (2013) and the references therein. Our goal is not to recapitulate the theory of finite-difference methods but to instead relate it to Markov chain approximation methods.

solution to the linear system

$$\begin{aligned}\rho V^{n+1}(k_j) = \rho u(c_j^n) + \mu(k_j, c_j^n)^+ V_F^{n+1}(k_j) - \mu(k_j, c_j^n)^- V_B^{n+1}(k_j) \\ + \frac{\sigma(k_j)^2}{2} V_C^{n+1}(k_j) + \frac{1}{\Delta_t}[V^n(k_j) - V^{n+1}(k_j)]\end{aligned} \tag{7}$$

for $j = 1, \ldots, |S|$, where $V_C(k_j) = [V(k_{j-1}) - 2V(k_j) + V(k_{j+1})]/\Delta_k^2$ denotes the second-order central difference, V_F and V_B denote forward and backward differences as above, and c_j^n solves

$$\max_{c \in \Gamma(k_j)} \rho u(c) + \mu(k_j, c)^+ V_F^n(k_j) - \mu(k_j, c)^- V_B^n(k_j).$$

In practice, convergence is typically rapid and insensitive to changes in Δ_t when the latter is large. To see why, note the fixed point of (7) solves $0 = \max_{c \in \Gamma(k)} \rho u(c) + T_{\text{IFD}}(c)V$, where

$$T_{\text{IFD}}(c)V(k) = \mu(k, c)^+ V_F(k) - \mu(k, c)^- V_B(k) + \frac{\sigma(k)^2}{2} V_C(k) - \rho V(k). \tag{8}$$

If we set $\Delta_t = \infty$, then the implicit method may be written as follows: fix V_0; find c_0 solving $\max_{c \in \Gamma(k)} \rho u(c) + T_{\text{IFD}}(c)V_0$; find V_1 solving $0 = \rho u(c_0) + T_{\text{IFD}}(c_0)V_1$; replace V_0 with V_1 and repeat until convergence. The method fits within the framework of Puterman and Brumelle (1979), who show that convergence is assured if $T_{\text{IFD}}(c)^{-1} \leq 0$ for all policy functions c. Finally, if we define $T(c; \Delta_t) = [e^{-\rho \Delta_t} P(c; \Delta_t) - I]/\Delta_t$, where $P(c; \Delta_t)$ denotes a matrix populated by the transition probabilites in (4) and I is the identity, then the following allows us to understand both the convergence properties of the implicit method and its relationship to the Markov chain approximation method.

Lemma 1. *For any policy c we have* $\lim_{\Delta_t \to 0} T(c; \Delta_t) = T_{\text{IFD}}(c)$.

Proof. Simply compare the right-hand side of (8) with $T(c; \Delta_t)V$, which is

$$-\frac{(1 - e^{-\rho \Delta_t})}{\Delta_t} V(k) + e^{-\rho \Delta_t} \left(\mu(k, c)^+ V_F(k) - \mu(k, c)^- V_B(k) + \frac{\sigma(k)^2}{2} V_C(k) \right)$$

from which the result follows by taking limits. □

Lemma 1 shows that the implicit finite-difference method of Achdou et al. (2022) amounts to solving the original problem by considering the limit of a particular family of Markov chains and a particular solution method (policy function iteration) for the resulting system of equations. Although Achdou et al. (2022) acknowledge a connection between Markov chain approximation methods and their finite-difference approach, we know of no analysis that provides the student of economics with guidance on which method to use in

any particular situation. The remainder of this document is devoted to this task. We first outline the method formally before turning to economic applications.

2.2 The Markov chain approximation method

This section provides a general overview of the theory developed in Kushner and Dupuis (2001). We focus only on fixing consistent notation and stating the relevant results and definitions necessary to understand subsequent examples and refer the reader to the text for details. We are interested in continuous-time control problems of the following form.

Definition 2. Let $(B_t)_{t\geq 0}$ be a standard n-dimensional Brownian motion defined on a probability space $(\Omega, \mathcal{F}, P)$, with $(\mathcal{F}_t)_{t\geq 0}$ the associated natural filtration. For a compact set $U \in \mathbb{R}^m$ define $\mathcal{C}$ to be the set of *admissible controls* — the set of stochastic processes $(u_t)_{t\geq 0}$ adapted to $(\mathcal{F}_t)_{t\geq 0}$ such that $u_t \in U$ for all $t \geq 0$. For some functions $F : \mathbb{R}^n \times U \to \mathbb{R}$, $\mu : \mathbb{R}^n \times U \to \mathbb{R}^n$ and $\sigma : \mathbb{R}^n \times U \to \mathbb{R}^{n\times n}$ we consider the problem

$$\begin{aligned} V(x) = \max_{u\in\mathcal{C}} \mathbb{E}\left[\int_0^\infty e^{-\rho t} F(x_t, u_t)dt\right] \\ dx_t = \mu(x_t, u_t)dt + \sigma(x_t, u_t)dB_t \\ x_0 = x. \end{aligned}$$

We refer to x and u as the *state* and *control* variables, F as the *payoff* function, and μ and σ as the *drift* and *diffusion* functions.

For the income fluctuation problems in this paper we make the following assumption.

Assumption 1. *The functions F, μ, and σ are bounded and Lipschitz.*

Assumption 1 can be weakened slightly without affecting the validity of the approach. However, it covers many examples of interest to us and ensures that a weak solution to the law of motion for the state variable exists and is unique for any admissible control, so that the value function in Definition 2 is well-defined. To reduce a problem of the form in Definition 2 to a finite-state problem, we must specify how to approximate the underlying *state* and *objective* function. The following introduces the notion of a locally consistent approximating Markov chain, which captures the requirement that the first and second conditional moments approximately coincide with their continuous-time counterparts.

Definition 3. A finite-state Markov chain approximation to the processes $(x_t)_{t\geq 0}$ satisfying $dx_t = \mu(x_t, u_t)dt + \sigma(x_t, u_t)dB_t$ for some admissible control $(u_t)_{t\geq 0}$ consists of a family of Markov chains $(\xi^h)_{h>0}$ over finite state spaces $(S_h)_{h>0}$, together with a family of time increment functions $(\Delta^h t(x, u))_{h>0}$, all indexed by scalars $h > 0$, satisfying:

(i) $\lim_{h\to 0} \sup_{x,u} \Delta^h t(x,u) = 0$; and

(ii) $\inf_{x,u} \Delta^h t(x,u) > 0$ for all $h > 0$.

Defining $\Delta_n^h x = \xi_{n+1}^h - \xi_n^h$ the approximation is *locally consistent* if

$$\begin{aligned} \mathbb{E}^h_{x,n,u}[\Delta_n^h x] &= \Delta^h t(x,u)\mu(x,u) + o(\Delta^h t(x,u)) \\ \mathbb{E}^h_{x,n,u}[(\Delta_n^h x - \mathbb{E}^h_{x,n,u}[\Delta_n^h x])^2] &= \Delta^h t(x,u)\sigma(x,u)\sigma(x,u)^t + o(\Delta^h t(x,u)) \end{aligned} \tag{9}$$

where $\mathbb{E}^h_{x,u,n}$ denotes the conditional expectation of the chain ξ^h at time t_n^h given $(\xi_n^h, u_n^h) = (x,u)$, where $u_n^h := u_{t_n^h}, t_0 = 0$ and $\Delta^h t_n = t_{n+1}^h - t_n^h = \Delta^h t(\xi_n^h, u_n^h)$. We refer to (9) as the *mean* and *covariance* consistency requirements, respectively.

We will drop superscripts and subscripts from expectations, since the appropriate operator will be obvious from the context. For each Markov chain ξ^h we will approximate the objective in Definition 2 as

$$\mathbb{E}\left[\int_0^\infty e^{-\rho t} F(x_t, u_t) dt\right] \approx \mathbb{E}\left[\sum_{n=0}^\infty e^{-\rho t_n^h} \Delta^h t_n F(\xi_n^h, u_n^h)\right]. \tag{10}$$

For each $h > 0$, associated with the Markov chain ξ^h and control process $(u_t)_{t\geq 0}$, we define the continuous-time processes $\overline{\xi}^h$ and $\overline{u}^h$ as the right-continuous and piecewise constant processes that coincide with the above chains at the times $(t_n)_{n\geq 0}$. The sum on the right-hand side of (10) is approximately $\mathbb{E}\left[\int_0^\infty e^{-\rho t} F(\overline{\xi}_t^h, \overline{u}_t^h) dt\right]$ with the only difference being the continuous discounting on the intervals $[t_n, t_{n+1})$, which necessarily vanishes as $h \to 0$. The weak convergence arguments of Kushner and Dupuis (2001) are applied to these continuously interpolated processes, so that all approximations to the original process are defined on the same path space. However, for each $h > 0$ the value function we solve numerically corresponds to the control problem with objective (10) and state evolving according to x^h, and so may be solved with discrete-time techniques.

Definition 4. Given a family of Markov chains $\{\xi^h\}_{h>0}$ locally consistent with $dx_t = \mu(x_t, u_t)dt + \sigma(x_t, u_t)dB_t$ for each control, define the approximate value functions

$$\begin{aligned} V^h(x) = \max_{u\in\mathcal{C}} \mathbb{E}&\left[\sum_{n=0}^\infty e^{-\rho t_n^h} \Delta^h t_n F(\xi_n^h, u_n^h)\right] \\ \xi_0^h &= x \end{aligned}$$

for any $h > 0$, where the maximum is once again over the set of all admissible controls.

The finite-state Markov chain approach applies discrete-time dynamic programming arguments to problems of the form in Definition 4 rather than the original problem in

Definition 2. This leads to the Bellman equation for the Markov chain

$$V^h(x) = \max_{u \in U} \Delta_t F(x,u) + e^{-\rho \Delta_t} \mathbb{E}^u[V^h(x')] \tag{11}$$

where x evolves according to the given approximating Markov chain. Our income fluctuation problems will assume discounting and uniformly bounded payoff functions, so that there are no subtleties regarding the applicability of the principle of optimality, and the fixed point of (11) coincides with the sequence problem given in Definition 4. Familiar arguments, such as those outlined in Stokey (1989), show that the right-hand side defines a contraction on the space of functions on the (finite) state space. Finally, Kushner and Dupuis (2001) show that under standard assumptions on the functions defined in the original problem in Definition 2, local consistency ensures convergence of the approximate value functions to the true value function, as stated in Theorem 5.2 of Chapter 10 in Kushner and Dupuis (2001).

Theorem 5. *Under Assumption 1 we have* $V^h(x) \to V(x)$ *as* $h \to 0$.

The MCA method may be used to solve problems in which there are jumps in the state variable. Although we do not strive for the most general framework possible, we outline here the theory necessary to solve a problem of particular interest to economists, in which a risk-averse consumer faces a consumption-savings problem with fluctuating income and may consume discrete amounts of a durable consumption good. The state variable for the agent will consist of her wealth, her income, and the current value of the durable good. This necessitates a discussion of jump processes, since the purchase of the durable good will coincide with a fall in wealth that does not vanish with the length of the time interval, and so cannot be well modelled with a diffusion process. We therefore consider a *jump*-diffusion process of the form

$$dx_t = \mu(x_t, u_{Dt})dt + \sigma(x_t)dZ_t + dJ_t(u_{Jt}) \tag{12}$$

where $(J_t)_{t \geq 0}$ is a jump process defined by

$$J_t = \int_0^t \int_\Gamma q(x_{s-}, u_{Js}, \varphi) N(dsd\varphi). \tag{13}$$

The control vector is written $u_t := (u_{Dt}, u_{Jt})$ to illustrate that some components affect only the drift and others affect only the jumps. In (13) one interprets the integrand $q(x_{s-}, u_{Js}, \varphi)$ as the size of the jumps at time s, with φ denoting the realization of some exogenous uncertainty and supported in some compact set Γ. The quantity N is a Poisson random measure with intensity density $n(dtd\varphi) = \lambda dt \times \Pi(d\varphi)$, meaning $\mathbb{E}[N(A)] = \int_A n(dtd\varphi)$ for all Borel sets A. For the case analyzed in Section 3.4 we may assume that φ is supported at a single point and that the jumps in (12) correspond to purchases of the durable good. It

may help the reader to imagine that the "jumps" are arriving at a constant exogenous rate $\lambda > 0$, but that they coincide with a movement in the state variable at date t only when $q(x_{t-}, u_{Jt}, \varphi) \neq 0$.

To construct a locally consistent Markov chain for (12), one begins with a locally consistent Markov chain for the diffusion component (with transition probabilities denoted by p_D^h) and obtains the approximation by independently drawing from this diffusion process and the jump component. The definition of a locally consistent finite-state Markov chain now includes a component q^h representing the jumps of the Markov chain. The manner in which the transition probabilities in Definition 6 are constructed from the probabilities associated with the diffusion term has an intuitive interpretation. We may view each transition as arising by drawing from a jump process with probability $\lambda\Delta_t$ for some constant Δ_t and drawing from the continuous process with probability $1 - \lambda\Delta_t$.[5]

Definition 6 (Local consistency with jumps)**.** A family of finite-state Markov chains $\{\xi^h\}_{h>0}$ with state spaces $\{S_h\}_{h>0}$ and transition probabilities $\{p^h(x, x')\}_{x,x' \in S_h}$ is locally consistent with the jump diffusion (12) if for each $h > 0$ there exist transition probabilities $\{p_D^h(x, x', u)\}_{x,x' \in S_h}$ and functions $(q^h)_{h>0}$ such that:

(i) The family of Markov chains defined by $\{p_D^h(x, x', u)\}_{x,x' \in S_h}$ is locally consistent with the diffusion process $(z_t)_{t \geq 0}$ defined by $dz_t = \mu(z_t, u_t)dt + \sigma(z_t)dZ_t$;

(ii) The functions $(q^h)_{h>0}$ converge to q uniformly as $h \to 0$; and

(iii) For some $\delta_h(x, u) = o(\Delta_t)$,

$$p^h(x, x', u) = (1 - \lambda\Delta_t - \delta_h(x, u))p_D^h(x, x', u) + (\lambda\Delta_t + \delta_h(x, u))1_{x+q^h(x,u_J)=x'}.$$

The analogue of Theorem 5 for the case of controlled jump-diffusion processes is outlined in Chapter 13 of Kushner and Dupuis (2001).

To solve control problems of the form in Definition 2, we therefore need only solve the Bellman equation (11) for some choice of a locally consistent approximating Markov chain. This is important because the literature on dynamic programming with finite state spaces contains a wealth of techniques for solving finite-state Markov decision problems. Section 3 illustrates the benefits of this viewpoint by solving an income fluctuation problem using modified policy function iteration. Section 4 deals with high correlation among multiple state variables, for which the construction of consistent chains poses some difficulties, and is not considered in the analysis of Achdou et al. (2022).

[5]The following definition is less general than that given in Section 5.6.2 of Kushner and Dupuis (2001), but is sufficient to cover the example in Section 3.4.

3 Income fluctuation problems

We first consider the setting of an income fluctuation problem in order to illustrate the benefits of departing from policy function iteration when solving various forms of income fluctuation problems. Sections 3.1 and 3.2 first recapitulate the theory behind modified policy function iteration and then explain our extension. Section 3.3 then considers an income fluctuation problem in which income is the product of two diffusion processes and the agent cares only about non-durable consumption. Section 3.4 considers a variation of an income fluctuation problem with discrete choices over a durable good and applies the normalization used prior to Lemma 1 to establish that an analogue of modified policy function iteration is applicable even when the timestep is sent to zero.

3.1 Modified policy function iteration

Recall that in the one-sector growth model of Section 2, we constructed a locally consistent chain for the capital process and solved the resulting system of equations using policy function iteration. The algorithm converged in a small number of iterations, which is unsurprising given that policy function iteration is known to converge locally at a quadratic rate. However, the updating step in policy function iteration requires solving a linear system of equations. The cost of this operation grows rapidly when the number of gridpoints increases or the sparsity structure of the transition matrix becomes more complex, both of which occur as the dimension of the problem grows.

Our first application will illustrate the benefits of employing the *modified* policy function iteration algorithm of Puterman and Shin (1978). This algorithm generalizes value function iteration by updating the value function a fixed number of times between successive updates of the policy function. In this case convergence is known to occur only at a linear rate, and so will typically require more iterations than policy function iteration. However, at no point in the algorithm do we need to solve a linear system of equations. Further, we show that an analogue of modified policy function iteration remains applicable even when the timestep vanishes, and so is well suited to settings in which the state variable may change by a large amount instantaneously, as is the case in problems in which one must choose consumption in a fixed finite set. To the best of our knowledge, this *generalized* policy function iteration is novel, and in practice appears quite useful in a wide range of applications.

We briefly recapitulate here the arguments and algorithms of Puterman and Brumelle (1979) and Puterman and Shin (1978) in order to fix ideas and to explain our generalization. Suppose we have a controlled finite-state Markov chain with state space S of cardinality $|S|$ and time increment $\Delta_t \in \mathbb{R}^{|S|}$, and that at each point $x \in S$ the control u may assume values in some subset U of Euclidean space, with the associated transition probabilities given by

$P : S^2 \times U \to [0,1]$. As the preceding notation indicates, we allow for the possibility that the time increment is state-dependent. The Bellman equation for a discrete-state problem with flow payoff function f and discount rate $e^{-\rho\Delta_t(x)}$ is

$$V(x) = \max_{u \in U} \Delta_t(x) f(x,u) + e^{-\rho\Delta_t(x)} \sum_{x' \in S} P(x,x',u) V(x') \qquad x \in S. \tag{14}$$

Writing $F(x,u) = \Delta_t(x) f(x,u)$ for all $x \in S$ and $u \in U$, we can write this as

$$0 = \max_{\hat{u} \in U^{|S|}} F(\hat{u}) + T(\hat{u}) V =: B(V) \tag{15}$$

where $T(\hat{u}) := \text{diag}(\beta) P(\hat{u}) - I$ for $\beta(x) := e^{-\rho\Delta_t(x)}$ and the second equality in (15) defines B. Policy function iteration is then the following.

Algorithm 1 (Policy function iteration)**.** Choose an arbitrary control u_0 and denote by V_0 the associated value function. We then iterate as follows:

(i) Choose $\hat{u}_0$ to solve $B(V_0) = F(\hat{u}_0) + T(\hat{u}_0) V_0$.

(ii) Define $V_1 = -T(\hat{u}_0)^{-1} F(\hat{u}_0)$ as the value of adhering to $\hat{u}_0$ forever.

(iii) Replace V_0 with V_1 in Step (i) and repeat until convergence.

If we write $\hat{u}(V)$ for the control that attains the maximum in (15) then the updating rule in Step (ii) of Algorithm 1 may be written

$$V_{n+1} = -T(\hat{u}(V_n))^{-1} F(\hat{u}(V_n)) = V_n - T(\hat{u}(V_n))^{-1} B(V_n). \tag{16}$$

Puterman and Brumelle (1979) show that policy function iteration is essentially an abstract version of Newton's method and inherits some of the same properties, such as rapid (quadratic) convergence near the solution. However, as we noted earlier, the updating step in policy function iteration requires the solution of a linear system of equations of a size as large as the number of gridpoints. Computational time therefore grows rapidly as one increases the grid size or dimension, motivating the search for alternatives to policy function iteration. To this end, note that if we abbreviate $\hat{u}_n := \hat{u}(V_n)$ then $-T(\hat{u}_n)^{-1} = \sum_{j=0}^{\infty} (\text{diag}(\beta) P(\hat{u}_n))^j$ and (16) becomes $V_{n+1} = V_n + \sum_{j=0}^{\infty} (\text{diag}(\beta) P(\hat{u}_n))^j B(V_n)$. Modified policy function iteration simply truncates this sum at a finite integer k.

Algorithm 2 (Modified policy function iteration)**.** Fix an initial guess V_0 satisfying $B(V_0) \geq 0$. We then iterate as follows:

(i) Choose $\hat{u}_0$ to solve $B(V_0) = F(\hat{u}_0) + (\text{diag}(\beta) P(\hat{u}_0) - I) V_0$.

(ii) Define

$$V_1 = V_0 + \sum_{j=0}^{k} (\text{diag}(\beta)P(\hat{u}_0))^j B(V_0). \tag{17}$$

(iii) Replace V_0 with V_1 in Step (i) and repeat until convergence.

Puterman and Shin (1978) show that for any integer k and initial guess V_0 satisfying $B(V_0) \geq 0$, the sequence of iterates $(V_n)_{n=1}^{\infty}$ produced by Algorithm 2 converges monotonically to the solution to the equation $B(V) = 0$. The case $k = 0$ in Algorithm 2 corresponds to value function iteration, while policy function iteration in Algorithm 1 arises as we take the limit $k \to \infty$. Also note that we can always find an initial guess V_0 satisfying $B(V_0) \geq 0$ by setting $V_0 = -T(u_0)^{-1}F(u_0)$ for an arbitrary policy u_0.

3.2 Generalized policy function iteration

Both Algorithm 1 and Algorithm 2 approximated a root of the function B defined in equation (15) in which $T(\hat{u}) := \text{diag}(\beta)P(\hat{u}) - I$ for some discount parameter β and transition matrix P. For these choices of B and T, the process by which the two algorithms updated the value function could be written as

$$V_1 = V_0 + \sum_{j=0}^{k} (I + T(\hat{u}_0))^j B(V_0) \tag{18}$$

without any explicit reference to the underlying Markov chain. Now note that both the root of equation (15) and the control that attains the maximum on the right-hand side are unaffected if we scale F and T state-by-state. This suggests the possibility of designing analogues of Algorithm 2 for scaled versions of equation (15). Such transformations will not affect Algorithm 1 (since scaling a linear system leaves the solution unaffected) but will affect the updating step (18) for finite k.[6] One benefit of this generalization is that it will allow us to apply an analogue of Algorithm 2 to an operator similar to that employed in the implicit finite-difference method of Achdou et al. (2022).

We now suppose that we wish to find a solution to an equation of the form (15), where T is no longer necessarily of the form $\text{diag}(\beta)P - I$ for some transition probability P and vector β. Given a normalizing function $C : S \times U^{|S|} \to \mathbb{R}$, we can define a *generalized* policy function iteration algorithm as follows.

Algorithm 3 (Generalized policy function iteration)**.** Fix an initial guess V_0 satisfying $B(V_0) \geq 0$. We then iterate as follows:

[6] The possibility of applying a version of the above algorithm to more general operators than those of the form $T = \beta P - I$ is mentioned on page 64 of Puterman and Brumelle (1979). We unfortunately cannot access the references that elaborate on this point.

(i) Choose $\hat{u}_0$ to solve $B(V_0) = F(\hat{u}_0) + T(\hat{u}_0)V_0$.

(ii) Define

$$V_1 = V_0 + \sum_{j=0}^{k} (I + \tilde{T}(\hat{u}_0))^j \tilde{B}(V_0) \tag{19}$$

where $\tilde{T}(\hat{u}_0)(x, x') := T(\hat{u}_0)(x, x')/C(x, \hat{u}_0)$ and $\tilde{B}(V_0)(x) := B(V_0)/C(x, \hat{u}_0)$ for all $x, x' \in S$.

(iii) Replace V_0 with V_1 in Step (i) and repeat until convergence.

As the following shows, the normalizing function may be chosen so that the resulting algorithm retains the attractive monotonicity properties derived by Puterman and Shin (1978) in the discounted Markovian case.

Lemma 7. *If the function C is chosen such that $I + \tilde{T}(\hat{u}_n) \geq 0$ for all $n \geq 0$, then the sequence defined by (19) is monotone increasing if $B(V_0) \geq 0$.*

Proof. For any $V, W \in \mathbb{R}^{|S|}$ we have $B(W) = \max_{\hat{u} \in U^{|S|}} F(\hat{u}) + T(\hat{u})W \geq F(\hat{u}(V)) + T(\hat{u}(V))W$ and so $B(W) \geq B(V) + T(\hat{u}(V))(W - V)$. It follows that for any $n \geq 1$ and $(V, W) = (V_n, V_{n+1})$ we have

$$\begin{aligned} B(V_{n+1}) &\geq B(V_n) + T(\hat{u}_n)(V_{n+1} - V_n) \\ &= B(V_n) + T(\hat{u}_n) \sum_{j=0}^{k} (I + \tilde{T}(\hat{u}_n))^j \tilde{B}(V_n) \\ &= (I + \tilde{T}(\hat{u}_n))^{k+1} B(V_n) \geq 0. \end{aligned}$$

Since $I + \tilde{T}(\hat{u}_n) \geq 0$, this shows that $B(V_n) \geq 0$ implies $B(V_{n+1}) \geq 0$ for all $n \geq 0$, from which the conclusion follows. □

Note that we can always find an initial guess V_0 satisfying $B(V_0) \geq 0$ by setting $V_0 = -T(u_0)^{-1} F(u_0)$ for an arbitrary policy u_0. In what follows, for each $n \geq 0$ we will simply choose $C(\cdot, \hat{u}_n)$ to be the least value such that $I + \tilde{T}(\hat{u}_n)(x, x') \geq 0$ for all $x, x' \in S$. To illustrate one important application of Algorithm 3 that we will employ in our examples, let $F(x, u) = f(x, u)$ and define

$$T(\hat{u}) = \lim_{\Delta_t \to 0} \frac{1}{\Delta_t} (e^{-\rho \Delta_t} P(\hat{u}, \Delta_t) - I) \tag{20}$$

where $P(\cdot, \Delta_t) : U^{|S|} \to \mathbb{R}^{|S| \times |S|}$ denote the transition probability functions arising from a discretization using the MCA method with constant timestep Δ_t. As we noted in Section 2, this operator in (20) often coincides with that appearing in the implicit finite-difference

method of Achdou et al. (2022). In this case, Algorithm 3 is effectively constructing a particular approximation to $T(\hat{u})^{-1}$ in such a way as to guarantee the monotonicity (and hence convergence) of the iterates, without needing to ever directly solve a linear system.

3.3 Non-durable consumption

We first consider the problem of an infinitely lived agent who faces idiosyncratic income risk and may save in a risk-free bond. As emphasized by Achdou et al. (2022), this is a natural application for an economist, as such problems are an integral component of Bewley-Huggett-Aiyagari incomplete markets models, which form the backbone of much of modern macroeconomics. Suppose that preferences over consumption are given by

$$U(c) = \mathbb{E}\left[\rho \int_0^\infty e^{-\rho t} \frac{c_t^{1-\gamma}}{1-\gamma} dt\right] \tag{21}$$

and that wealth evolves according to

$$da_t = [ra_t - c_t + y_t]dt \tag{22}$$

where $r > 0$ is fixed and $(y_t)_{t\geq 0}$ denotes the income of the agent. We also assume that the agent faces a borrowing constraint of the form $a_t \geq \underline{a}$ for all $t \geq 0$ and some $\underline{a}$, where for simplicity in what follows we assume $\underline{a} = 0$, so that the agent is unable to borrow. The state-dependence of the constraint set means that this problem does not technically fit into the framework of Section 2.2. However, for simplicity we proceed as if the agent receives a large negative payoff from leaving the prescribed state space. We assume $y_t = e^{z_t}$ where $z_t := z_{1t} + z_{2t}$ for z_1 and z_2 satisfying

$$dz_{it} = -\theta_i z_{it} dt + \sigma_i dZ_{it} \tag{23}$$

for $i = 1, 2$, where $Z := (Z_{1t}, Z_{2t})_{t\geq 0}$ is two-dimensional Brownian motion. We introduce two income shocks simply to illustrate the benefits of departing from policy function iteration when the grid grows large, but one can imagine such a problem arising when one wishes to model, e.g., consumer behavior in the presence of both persistent and transitory shocks. When approximating (23) we assume the volatility vanishes near the boundary, but omit this from the notation for brevity. Since income is bounded, there is no loss in assuming that both assets and consumption must be bounded, so that Assumption 1 is satisfied. To construct our chain we must specify the state space, the transition probabilities, and the (possibly state-dependent) timestep. Write $\overline{a}$ for the maximum level of wealth in the discretization, and $\underline{z}_1, \overline{z}_1, \underline{z}_2$ and $\overline{z}_2$ for the lower and upper bounds for the income processes.

For a vector of integers $N = (N_a, N_1, N_2)$, define the state increments

$$(\Delta_a, \Delta_1, \Delta_2) = ((\overline{a} - \underline{a})/N_a, (\overline{z}_1 - \underline{z}_1)/N_1, (\overline{z}_2 - \underline{z}_2)/N_2)$$

and the individual grids $S_a = \{\underline{a} + \Delta_a, \ldots, \overline{a} - \Delta_a\}$ and $S_i = \{\underline{z}_i + \Delta_i, \ldots, \overline{z}_i - \Delta_i\}$ for $i \in \{1, 2\}$ and define the state space $S_h := S_a \times S_1 \times S_2$. We choose our transition probabilities so that if the chain is at point $x := (a, z_1, z_2) \in S_h$ at time t then the possible values at time $t + \Delta_t$ lie in the set

$$\Delta(x) := \{(a, z_1, z_2), (a \pm \Delta_a, z_1, z_2), (a, z_1 \pm \Delta_1, z_2), (a, z_1, z_2 \pm \Delta_2)\}.$$

One may check that the following defines a locally consistent chain for any c,

$$\begin{aligned} p(a \pm \Delta_a, z_1, z_2) &= \frac{\Delta_t}{\Delta_a}[ra - c + e^{z_1+z_2}]^{\pm} \\ p(a, z_1 \pm \Delta_1, z_2) &= \frac{\Delta_t}{\Delta_1^2}\left(\frac{\sigma_1^2}{2}\chi_1(z_1) + \Delta_1[-\theta_1 z_1]^{\pm}\right) \\ p(a, z_1, z_2 \pm \Delta_2) &= \frac{\Delta_t}{\Delta_2^2}\left(\frac{\sigma_2^2}{2}\chi_2(z_2) + \Delta_2[-\theta_2 z_2]^{\pm}\right) \end{aligned} \tag{24}$$

where $\chi_i(z_i) := 1_{z_i \notin \{\underline{z}_i + \Delta_i, \overline{z}_i - \Delta_i\}}$ and $i = 1, 2$, provided these quantities lie in the unit interval. The borrowing constraint is imposed by requiring $c \leq ra + y$ when $a = \underline{a} + \Delta_a$. To ensure that the process remains on the grid, we impose $c \geq ra + y$ for $a = \overline{a} - \Delta_a$.[7] The Bellman equation is then

$$V(x) = \max_{c \geq 0} \Delta_t \frac{c^{1-\gamma}}{1-\gamma} + e^{-\rho\Delta_t}\mathbb{E}[V(x')] \tag{25}$$

for all $x \in S_h$. Optimal consumption solves

$$\max_{c \geq 0} \frac{c^{1-\gamma}}{1-\gamma} + e^{-\rho\Delta_t}\left([ra - c + e^{z_1+z_2}]^+ V_F^a - [ra - c + e^{z_1+z_2}]^- V_B^a\right)$$

which again requires no non-linear root-finding.

Numerical illustration: We now compare computational times for the algorithms in Section 3.1 for a fixed set of parameters. To the extent possible we adopt the parameters of Appendix F in Achdou et al. (2022), who consider a two-dimensional problem and compare the performance of their implicit finite-difference scheme with the endogenous grid method of Carroll (2006). Arguments analogous to those provided for the one-sector neoclassical growth model in Section 2 reveal that the implicit finite-difference method of Achdou et al.

[7] Note, however, that for a sufficiently large upper limit this will not bind.

(2022) is asymptotically equivalent to using policy function iteration for the probabilities (24) as the timestep vanishes. To contrast the approach adopted in this paper with that of Achdou et al. (2022), we therefore compare policy function iteration with the algorithms of Section 3.1.

Following Achdou et al. (2022) we fix $\gamma = 2$, $r = 0.03$, and $\rho \approx 0.0526$, corresponding to a discrete-time discount parameter of $\beta = 0.95$. With a single income state variable Achdou et al. (2022) target an annual autocorrelation of 0.95, which implies that $\theta = -\ln(0.95) \approx 0.0513$. Since the stationary solution to (23) is Gaussian with mean zero and variance $\nu^2 := \sigma^2/(2\theta)$, their choice of $\nu = 0.2$ implies that $\sigma \approx 0.064$. To illustrate the effect of changing the sparsity structure of the transition matrix on the performance of different methods, we consider two choices for the income process. In each case we choose the parameters of Achdou et al. (2022) for the first component of income. In the first case we set $z_2 = 0$ so that the problem becomes two-dimensional, while in the second we choose $(\theta_2, \sigma_2) = (\theta_1, \sigma_1)$. For each choice we solve the above problem using the modified and generalized modified policy function iteration algorithms of Section 3.1 for a number of grid sizes and relaxation steps. In the former case we must also specify a timestep with the property that the probabilities in (24) remain bounded within the unit interval. This can be found either by experimentation or by imposing *a priori* bounds on consumption and checking ex-post that they do not bind. To understand the latter approach, note that the transition probabilities defined in (24) will lie within the unit interval at a point (a, z_1, z_2) if

$$\Delta_t\big(|ra - c + e^{z_1+z_2}|/\Delta_a + \sigma_1^2/\Delta_1^2 + |\theta_1 z_1|/\Delta_1 + \sigma_2^2/\Delta_2^2 + |\theta_2 z_2|/\Delta_2\big) < 1.$$

If we impose the requirement that $c \in [0, \kappa(ra + e^{z_1+z_2})]$ for some $\kappa \geq 2$ and all (a, z_1, z_2), then it will suffice to choose a timestep no greater than the (grid-dependent) upper bound

$$\overline{\Delta}_t(\kappa) := \big((\kappa - 1)|ra + e^{z_1+z_2}|/\Delta_a + \sigma_1^2/\Delta_1^2 + |\theta_1 z_1|/\Delta_1 + \sigma_2^2/\Delta_2^2 + |\theta_2 z_2|/\Delta_2\big)^{-1}.$$

In our numerical experiments, we will conjecture (and verify ex-post) that consumption never exceeds two times interest and labor income, and so use the above $\overline{\Delta}_t(\kappa)$ with $\kappa = 2$. It is possible to experiment with choices of the timestep that ensure more rapid convergence, but doing so may lead the above expressions for probabilities to lie outside the unit interval. Note that this experimentation is unnecessary in the case of generalized modified policy function iteration, in which the timestep has been sent to zero, and it is for this reason that we record the output of both exercises. When implementing the generalized modified policy function iteration of Section 3.1, we choose our normalizing constant C at each iteration to be the least value necessary in order to ensure the non-negativity of $I + \tilde{T}$ everywhere.

We now record the speed of convergence for policy function iteration (recall this is the implicit method of Achdou et al. (2022) for $\Delta_t = \infty$), value function iteration, and modified value function iteration with $k = \{10, 50, 100\}$, grids chosen to encompass four standard deviations in each dimension of log income, and a tolerance between successive iterations in the supremum norm of 10^{-6}. In order to make an apples-to-apples comparison across algorithms the initial guess is always the value function associated with zero net savings. All figures are in seconds and are the average of ten runs, and the updating step in the policy function iteration was conducted using the standard scipy sparse solver (scipy.sparse.linalg.spsolve) in Python on an Intel Core i7-8650U processor. Table 1 gives the time until convergence in seconds for both modified and generalized policy function iteration in two dimensions.

Grid size	PFI	VFI	$k = 10$	$k = 50$	$k = 100$	$k = 200$
(200, 10)	0.037	3.681	0.417	0.120	0.076	0.063
(300, 15)	0.080	8.774	0.986	0.285	0.189	0.156
(400, 20)	0.142	15.806	1.792	0.533	0.354	0.284
(500, 25)	0.274	32.265	3.635	1.070	0.690	0.526

Table 1: Time until convergence for $2D$ problem: MPFI

Grid size	PFI	$k = 0$	$k = 10$	$k = 50$	$k = 100$	$k = 200$
(200, 10)	0.045	1.363	0.148	0.059	0.053	0.057
(300, 15)	0.097	3.690	0.388	0.138	0.109	0.105
(400, 20)	0.172	7.869	0.813	0.251	0.211	0.189
(500, 25)	0.329	15.594	1.598	0.498	0.362	0.350

Table 2: Time until convergence for $2D$ problem: Generalized MPFI

In this example there is no gain from departing from policy function iteration. In contrast, Table 3 repeats this exercise for a three-dimensional problem. For the third and fourth choices of grids, modified policy function iteration is faster than policy function iteration. Table 4 gives the analogous results for the case of generalized modified policy function iteration.

Grid size	PFI	VFI	$k = 10$	$k = 50$	$k = 100$	$k = 200$
(45, 15, 15)	0.762	7.410	0.847	0.249	0.195	0.158
(60, 20, 20)	4.828	23.362	2.652	0.797	0.540	0.420
(75, 25, 25)	20.045	58.819	6.635	1.944	1.237	1.088
(90, 30, 30)	68.306	110.348	12.535	3.757	2.432	2.062

Table 3: Time until convergence for $3D$ problem: MPFI

Grid size	PFI	$k = 0$	$k = 10$	$k = 50$	$k = 100$	$k = 200$
(45, 15, 15)	0.902	9.763	1.103	0.363	0.281	0.330
(60, 20, 20)	4.924	33.238	3.650	1.136	0.868	0.722
(75, 25, 25)	21.632	94.989	10.582	2.956	1.997	1.752
(90, 30, 30)	67.251	203.795	22.390	6.432	4.112	3.398

Table 4: Time until convergence for $3D$ problem: Generalized MPFI

Obviously, absolute (rather than relative) running times depend on the operating system and are potentially subject to substantial idiosyncratic variation. However, in light of the results in Table 3 and Table 4, we believe we are being conservative when we state that the change in solution methods can reduce computational times by an order of magnitude.

It is worth emphasizing that the generalized modified policy function iteration has the advantage that one need not worry about the timestep being chosen such that the probabilities lie in the unit interval, which in more general settings than the above may be non-trivial to ensure. As a result, it does not appear that one approach is always better than the other and we believe both to be of interest.

3.4 Durable consumption and discrete choice

One interesting property of the generalized modified policy function iteration given in Section 3.1 is that it remains applicable even as the timestep vanishes. This is useful for discrete-choice models in which there are large and instantaneous changes in wealth. To illustrate, we now consider a variation of the problem of Section 3.3 in which the agent has preferences over non-durable consumption that may assume a continuum of values, as well as a durable good that may assume only finitely many values. This is similar to a continuous-time version of the model of Fella (2014), who extends the endogenous grid method of Carroll (2006) to allow for both adjustment costs and discrete choices. The

MCA method may be applied to this case without any delicate choices of grids or any need for interpolation of the function. In this case the amount by which wealth changes upon purchase of the durable good does not vanish as the size of the grid tends to zero, and so we are not able to restrict attention to adjacent transitions. However, as we shall see, this causes no major difficulties, as the optimal policy in the updating step still admits a closed-form expression.

We assume that the agent has preferences over non-durable and durable consumption represented by the function

$$U(c, D) := \mathbb{E}\left[\rho \int_0^\infty e^{-\rho t} u(c_t, D_t) dt\right] \tag{26}$$

for some u and denote the values of durable consumption by $S_D := \{\underline{D}, \underline{D} + \Delta_D, \ldots, \underline{D} + N_D \Delta_D\}$ for some $\underline{D}, N_D$ and Δ_D. Note that in contrast with the case of non-durable consumption, this grid S_D is a primitive of the problem, and not a choice made in the discretization.

As in Section 3.3 we suppose that income is of the form $y_t = e^{z_t}$ for some mean-reverting $(z_t)_{t \geq 0}$, and model the choice of the durable good as follows. At any instant the agent makes a binary choice indicating whether she wishes to change the durable good. For simplicity, we assume that the agent may only purchase (and not sell) the durable good. Further, the opportunities to change the durable good are assumed to only arrive stochastically at some constant rate $\lambda > 0$. As $\lambda \to \infty$ this approximates a situation in which the durable good may change instantaneously. If $\overline{p}$ denotes the price of the durable good, then for some constants θ and σ the laws of motion for assets, log income, and durable consumption are

$$\begin{aligned} da_t &= [r a_t + e^{z_t} - c_t] dt + dJ_{1t}(u_{Jt}) \\ dz_t &= -\theta z_t dt + \sigma dZ_t + dJ_{2t}(u_{Jt}) \\ dD_t &= dJ_{3t}(u_{Jt}) \end{aligned} \tag{27}$$

where $(J_t)_{t \geq 0} = (J_{1t}, J_{2t}, J_{3t})_{t \geq 0}$ is a process of the form (13), with $u_{Jt} \in \{0, 1\}$ a control indicating a desired increase in the durable good and the jumps are $q_t(0) = (0, 0, 0)$ and $q_t(1) = (-\overline{p}\Delta_D, 0, \Delta_D)$. The formulation in (27) implicitly assumes that the agent may only change durable consumption by one unit at a time. However, if we choose λ to be a large number this approximates a situation in which the agent is unrestricted in her choice. As discussed in Section 2.2, to construct an approximating Markov chain for this jump-diffusion process, we treat the diffusion and jump components separately. We first define the discretized problem for a positive timestep before considering operators that arise in the limit as we send the timestep to zero, as per the discussion following Lemma 7.

First, define the discrete grid $S := S_a \times S_z \times S_D$, where $S_a := \{\underline{a} + \Delta_a, \ldots, \overline{a} - \Delta_a\}$ and

$S_z := \{\underline{z} + \Delta_z, \ldots, \overline{z} - \Delta_z\}$ for some integers $N_a, N_z \geq 1$ and bounds $\underline{a}, \overline{a}, \underline{z}$ and $\overline{z}$, where $(\Delta_a, \Delta_z) := ((\overline{a} - \underline{a})/N_a, (\overline{z} - \underline{z})/N_z)$. To ensure that income remains on the grid, we impose $\overline{p}\Delta_D = K\Delta_a$ for some integer $K \geq 1$. We first define the transition probabilities for wealth and income,

$$p(a \pm \Delta_a, z, D) = \frac{\Delta_t}{\Delta_a}[ra + e^z - c]^{\pm}$$
$$p(a, z \pm \Delta_z, D) = \frac{\Delta_t}{\Delta_z^2}\left(\frac{\sigma^2}{2}\chi(z) + \Delta_z[-\theta z]^{\pm}\right)$$

where $\chi(z) := 1_{z \notin \{\underline{z} + \Delta_z, \overline{z} - \Delta_z\}}$ and define the transitions for the durable good

$$p(a - \overline{p}\Delta_D, z, D + \Delta_D) = \lambda 1_{u_{Jt}=1}\Delta_t.$$

The Bellman equation is then $0 = \max_{c, u_J} u(c, D) + T(c, u_J; \Delta_t)V$, where

$$T(c, u_J; \Delta_t)V = \frac{1}{\Delta_t}\left(e^{-\rho\Delta_t}\mathbb{E}[V(a', z', D')] - V(a, z, D)\right)$$

and the optimal policy for the durable good is $u_J = 1_{V(a - \overline{p}\Delta_D, z, D + \Delta_D) > V(a, z, D)}$. For our example we follow Fella (2014) and assume preferences of the form $u(c, D) = \ln c + \eta \ln(D + \iota)$ for some $\eta, \iota > 0$, so that the problem of finding optimal consumption is identical to the problem in Section 3 with $\gamma = 1$. We now define $\tilde{T}(c, u_J) = \lim_{\Delta_t \to 0} T(c, u_J; \Delta_t)$. Simplification gives

$$\begin{aligned}
\tilde{T}(c, u_J)V = &\frac{1}{\Delta_a}[ra + e^z - c]^+[V(a + \Delta_a, z, D) - V(a, z, D)] \\
&+ \frac{1}{\Delta_a}[ra + e^z - c]^-[V(a - \Delta_a, z, D) - V(a, z, D)] \\
&+ \frac{1}{\Delta_z^2}\left(\frac{\sigma^2}{2}\chi(z) + \Delta_z[-\theta z]^+\right)[V(a, z + \Delta_z, D) - V(a, z, D)] \\
&+ \frac{1}{\Delta_z^2}\left(\frac{\sigma^2}{2}\chi(z) + \Delta_z[-\theta z]^-\right)[V(a, z - \Delta_z, D) - V(a, z, D)] \\
&+ \lambda(V(a + q_a(u_J), z + q_z(u_J), D + q_D(u_J)) - V(a, z, D)) - \rho V(a, z, D)
\end{aligned} \tag{28}$$

where we have used the notation $q = (q_a, q_z, q_D)$ for the jumps. Finally, when implementing the generalized modified policy function iteration of Section 3.1, we again choose our normalizing function C to be the least value necessary in order to ensure the non-negativity of $I + \tilde{T}$ pointwise throughout the iterative process.

Numerical illustration: For simplicity we retain the same parameters for the income process as in Section 3. Our preferences are of the form used on page 339 of Fella (2014) and are ordinally equivalent to $\eta = 1/0.77 - 1$. We also follow Fella (2014) in our choice of

Grid size	PFI	$k = 0$	$k = 10$	$k = 50$	$k = 100$	$k = 200$
(50, 10, 10)	0.237	1.061	0.137	0.118	0.130	0.154
(100, 20, 10)	1.974	9.157	0.982	0.481	0.519	0.610
(150, 30, 10)	8.178	35.887	3.883	1.210	1.293	1.524
(200, 40, 10)	22.962	87.559	9.347	2.554	2.373	2.797
(250, 50, 10)	45.126	184.858	19.771	5.429	3.683	4.381

Table 5: Time until convergence for durable consumption problem: Generalized MPFI

interest rate $r = 0.06$, discount parameter $\rho = -\ln(0.93)$, and $\iota = 0.01$ and set the upper bound for durable consumption to be roughly 10 times the unconditional average of income. For simplicity we set $\overline{p} = 1$. Finally, for the arrival rate of the jumps we choose $\lambda = 52$, which may be interpreted as the assumption that when the agent elects to purchase a unit of the durable good, it takes on average roughly a week for the transaction to go ahead.

Table 5 gives the average time until convergence for ten runs of the generalized modified policy function iteration, with a tolerance in the supremum norm between successive iterations of 10^{-6}. The initial guess is always the value function associated with zero net savings and no purchase of the durable good. The speed gains from departing from policy function iteration appear lower here than for the case with non-durable consumption with two income shocks (although they are still substantial). We suspect that this is because the infrequency of the durable good purchases implies that the updating matrix is more sparse, which reduces the computational cost of using a direct solver for policy function iteration.

Figure 2 plots slices of the policy function for non-durable consumption together with a slice of the value function, and Figure 3 plots the change in durable consumption at several values of the durable good.

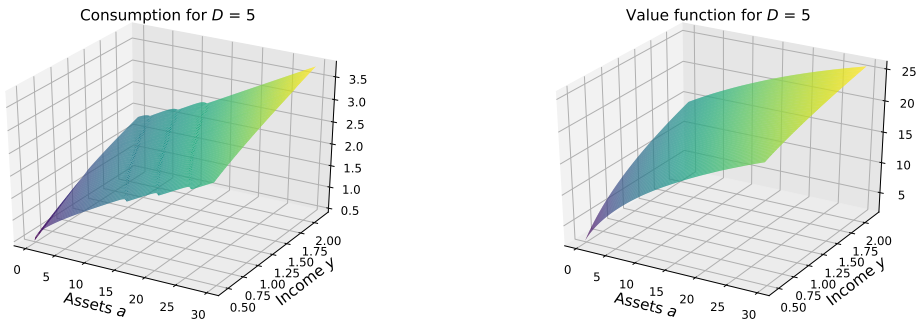

Figure 2: Slices of non-durable consumption and value function

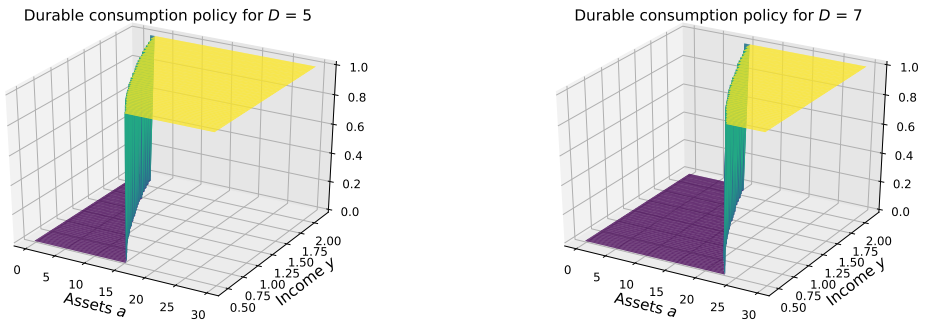

Figure 3: Slices of durable consumption policy function

Non-durable consumption appears to be increasing everywhere in the state space, except for small regions in which the agent wishes to increase their durable consumption. At these points, it appears that an increase in wealth and income may lead to a decrease in non-durable consumption, as the agent anticipates a "large" purchase of the durable good. However, we emphasize once again that this lack of monotonicity of the policy function causes no difficulties for the algorithm, as the policy functions at each stage of the iteration remain given in closed-form.

4 Portfolio problems and financial frictions

This section considers a general equilibrium model with financial frictions in the spirit of Brunnermeier and Sannikov (2014). To illustrate the flexibility of the MCA approach, we allow for time-varying volatility correlated with aggregate (depreciation) shocks. A growing number of models in the macrofinance literature possess such high correlation, posing difficulties for the construction of locally consistent chains. Section 4.1 outlines the environment; Section 4.2 describes how to construct locally consistent chains with high correlation among state variables; and Section 4.3 compares two algorithms for approximating competitive equilibria.

4.1 Setup

Agents may be one of two types, indexed $i \in \{E, H\}$, and referred to as *experts* and *households*, respectively. There is a unit mass of each type indexed by $j \in [0, 1]$. Both types of agents are infinitely lived with the same flow utility function but differ in their discount

rates, with preferences over sequences of consumption of the form

$$U_i(c) = \mathbb{E}\left[\rho_i \int_0^\infty e^{-\rho_i t} \frac{c_t^{1-\gamma}}{1-\gamma} dt\right]$$

for some $\rho_H, \rho_E > 0$ with $\rho_E > \rho_H$ and $\gamma \in (0,1)$. Aggregate capital in the economy at time $t \geq 0$ is denoted k_t, and the amount held by the jth agent of type i is denoted k_{it}^j. When the jth agent of type i invests a (possibly negative) fraction ι_{it}^j of her capital in new capital, the flow of output minus investment is $(\Pi_i - \iota_{it}^j)k_{it}^j dt$ and the law of motion of her capital stock is

$$dk_{it}^j = \iota_{it}^j k_{it}^j dt + \sigma_t k_{it}^j dZ_t \tag{29}$$

where $Z = (Z_t)_{t\geq 0}$ is a Brownian motion common to all agents. The increments of the Brownian motion in (29) may then be thought of as representing stochastic depreciation shocks. The parameters Π_E and Π_H are exogenous and assumed to satisfy $\Pi_E > \Pi_H$. Together with above assumption $\rho_E > \rho_H$, we are therefore assuming that experts are more productive but also more impatient, which will ensure that they do not ultimately own all of the wealth in the economy. The linearity in the investment production technology implies that the price of capital is constant (and here is unity). The volatility $(\sigma_t)_{t\geq 0}$ evolves according to

$$d\sigma_t = \theta(\overline{\sigma} - \sigma_t)dt + \sigma_\sigma dZ_t \tag{30}$$

for some positive $\theta, \overline{\sigma}$ and σ_σ, where the Brownian motions in (29) and (30) coincide. Agents may trade a risk-free bond in zero net supply with (endogenously determined) return denoted $(r_t)_{t\geq 0}$. We also assume that agents have access to a risk-free storage technology with exogenous and constant real return $\underline{r}$. An agent with wealth a_t must choose capital k_t, bond b_t and storage h_t holdings satisfying $k_t + b_t + h_t = a_t$. We assume that wealth, consumption, capital and storage holdings are non-negative, but that bond holdings may assume either sign. We then have the following definition.

Definition 8. Given a process for the interest rate $(r_t)_{t\geq 0}$, the problem of an agent of type $i \in \{E, H\}$ at time $t \geq 0$ with state (a, σ) is

$$\begin{aligned}
V_{it}(a,\sigma) = \max_{(c_\tau, h_\tau, k_\tau)_{\tau\geq t}} & \mathbb{E}\left[\rho_i \int_t^\infty e^{-\rho_i(\tau - t)} \frac{c_\tau^{1-\gamma}}{1-\gamma} d\tau\right] \\
da_\tau &= [r_\tau a_\tau + (\underline{r} - r_\tau)h_\tau - c_\tau + (\Pi_i - r_\tau)k_\tau]d\tau + \sigma_\tau k_\tau dZ_\tau \\
d\sigma_\tau &= \theta(\overline{\sigma} - \sigma_\tau)d\tau + \sigma_\sigma dZ_\tau \\
(a_t, \sigma_t) &= (a, \sigma) \\
a_\tau, c_\tau, h_\tau, k_\tau &\geq 0.
\end{aligned} \tag{31}$$

The homotheticity of utility and the log-linearity of the law of motion for wealth together imply the following lemma, which motivates our subsequent search for a recursive formulation of the problem.

Lemma 9 (Homogeneity)**.** *For any process $(r_t)_{t\geq 0}$ and $i \in \{E, W\}$ there exist processes $(\overline{V}_{it})_{t\geq 0}$ and $(\overline{c}_{it}, \overline{h}_{it}, \overline{k}_{it})_{t\geq 0}$ such that $V_{it}(a,\sigma) = \overline{V}_{it}(\sigma)a^{1-\gamma}/(1-\gamma)$ and $c_{it}(a,\sigma) = \overline{c}_{it}(\sigma)a$, $h_{it}(a,\sigma) = \overline{h}_{it}(\sigma)a$ and $k_{it}(a,\sigma) = \overline{k}_{it}(\sigma)a$, respectively, for all $t, a, \sigma \geq 0$.*

In what follows we abuse notation slightly and write V_i for $\overline{V}_i$. Using the linearity of policy functions in Lemma 9, aggregate consumption, storage and capital demand may be written as functions of policy functions and the wealth share of experts,

$$x_t := \frac{\int_0^1 a_{Et}^j dj}{\int_0^1 a_{Et}^j dj + \int_0^1 a_{Ht}^j dj}. \tag{32}$$

Appendix B.2 shows that the wealth share evolves according to a diffusion process of the form $dx_t = \mu_x(x_t, \sigma_t)x_t dt + \sigma_x(x_t, \sigma_t)x_t dZ_t$ for some μ_x and σ_x depending on the policy functions of each agent. We therefore focus on *Markov equilibria* in which equilibrium quantities are functions only of (x, σ).

Definition 10. For any functions r, μ_x and σ_x, state (a, x, σ) and $i \in \{E, W\}$, the problem of the ith type of agent may be written

$$\begin{aligned}
W_i(a,x,\sigma) = \max_{(\overline{c}_t,\overline{h}_t,\overline{k}_t)_{t\geq 0}} \mathbb{E}\left[\int_0^\infty \rho_i e^{-\rho_i t}\frac{(\overline{c}_t a_t)^{1-\gamma}}{1-\gamma}\right]dt \\
da_t &= [r(x_t,\sigma_t) + (\underline{r} - r(x_t,\sigma_t))\overline{h}_t - \overline{c}_t]a_t dt + \overline{k}_t a_t dR_t \\
dx_t &= \mu_x(x_t,\sigma_t)x_t dt + \sigma_x(x_t,\sigma_t)x_t dZ_t \\
d\sigma_t &= \theta(\overline{\sigma} - \sigma_t)dt + \sigma_\sigma dZ_t \\
(a_0, x_0, \sigma_0) &= (a, x, \sigma) \\
a_t, c_t, h_t, k_t &\geq 0
\end{aligned}$$

where $dR_t = (\Pi_i - r(x_t,\sigma_t))dt + \sigma_t dZ_t$.

Lemma 9 implies that the value functions for type $i \in \{E, H\}$ assume the form $V_i(x,\sigma)a^{1-\gamma}/(1-\gamma)$ for some function V_i, with associated policy functions of the form $\left(\overline{c}_i(x,\sigma)a, \overline{h}_i(x,\sigma)a, \overline{k}_i(x,\sigma)a\right)$ for all (a, x, σ).

Definition 11 (Markov equilibrium)**.** A Markov equilibrium consists of functions for the risk-free rate and drift and diffusion for the wealth share, together with value functions V_i and policy functions $(\overline{c}_i, \overline{h}_i, \overline{k}_i)$ for $i \in \{E, H\}$ solving the problem in Definition 10, such

that the bond market clears,

$$0 = \big(1 - \overline{k}_E(x,\sigma) - \overline{h}_E(x,\sigma)\big)x + \big(1 - \overline{k}_H(x,\sigma) - \overline{h}_H(x,\sigma)\big)[1-x]$$

for all (x,σ), and the law of motion for the wealth share is consistent with individual policy functions.

The presence of the storage technology implies that we cannot have $r < \underline{r}$, for this would imply the existence of an arbitrage opportunity by borrowing the bond and employing the storage technology. We may therefore focus (without loss) on the case $r \geq \underline{r}$ and omit storage choice from the agent's problems.[8] We assume for simplicity that households are unproductive with capital, so that $\Pi_H < \underline{r}$, and also assume that $\Pi_E > \underline{r}$ to rule out the trivial case in which only the storage technology is utilized. Given the policy functions in the Markov equilibrium, investment may then be determined residually from the goods market clearing condition

$$\overline{c}_E(x,\sigma)x + \overline{c}_H(x,\sigma)(1-x) = [\Pi_E - \iota_E]\overline{k}_E(x,\sigma)x + [\overline{h}_E(x,\sigma)x + \overline{h}_H(x,\sigma)(1-x)]\underline{r}. \quad (33)$$

4.2 Chain construction with high correlation

We first explain how to solve the individual problem in Definition 10. This poses some difficulties, as standard constructions fail to work when there is high correlation between state variables. To see why, suppose that $(X_t)_{t\geq 0}$ is a two-dimensional process satisfying $dX_t = \mu(X_t)dt + \sigma(X_t)dZ_t$ where $\mu : \mathbb{R}^2 \to \mathbb{R}^2$, $\sigma : \mathbb{R}^2 \to \mathbb{R}^{2\times m}$ and $(Z_t)_{t\geq 0}$ is m-dimensional Brownian motion. Now define coefficients $a_{ij}(X) = (\sigma\sigma^T)_{ij}$ for $i,j = 1,2$. Suppose that μ and σ vanish outside of some domain $[-M,M]^2$, and set $S := S_1 \times S_2$ where S_1 and S_2 are uniform grids with increments Δ_1 and Δ_2, respectively. Denote an arbitrary element of S by $x = (x_1, x_2)$ and consider a Markov chain such that if the chain is at point x at time t, then the possible values at time $t + \Delta_t$ are

$$\Delta(x) := \{(x_1,x_2), (x_1 \pm \Delta_1, x_2), (x_1, x_2 \pm \Delta_2), (x_1 \pm \Delta_1, x_2 \pm \Delta_2), (x_1 \pm \Delta_1, x_2 \mp \Delta_2)\}.$$

It is easy to check that if for $i = 1,2$ we have

$$a_{ii} - \sum_{j\neq i} |a_{ij}|\Delta_i/\Delta_j \geq 0, \quad (34)$$

[8] This does not mean storage is irrelevant as it may be utilized in equilibrium if $r = \underline{r}$. However, the choice of storage will be indeterminate if $r = \underline{r}$ and zero otherwise and so may be omitted from the agent's problem. The presence of the storage technology just puts a lower bound on r.

everywhere, then for sufficiently small $\Delta_t > 0$ the probabilities

$$\begin{aligned}
p(x_1 \pm \Delta_1, x_2) &= \frac{\Delta_t}{\Delta_1^2}\left(\frac{1}{2}[a_{11} - |a_{12}|\Delta_1/\Delta_2] + \Delta_1 \mu_1^{\pm}\right) \\
p(x_1, x_2 \pm \Delta_2) &= \frac{\Delta_t}{\Delta_2^2}\left(\frac{1}{2}[a_{22} - |a_{12}|\Delta_2/\Delta_1] + \Delta_2 \mu_2^{\pm}\right) \\
p(x_1 \pm \Delta_1, x_2 \pm \Delta_2) &= \frac{\Delta_t}{\Delta_1 \Delta_2}\frac{1}{2}a_{12}^{+} \\
p(x_1 \pm \Delta_1, x_2 \mp \Delta_2) &= \frac{\Delta_t}{\Delta_1 \Delta_2}\frac{1}{2}a_{12}^{-}
\end{aligned}$$

define a locally consistent chain. Relative to the uncorrelated case, this procedure removes some mass from the left/right/up/down transitions and adds it to the diagonal elements. However, the above construction will fail to work whenever (34) fails for some point in the domain, since the expressions for probabilities may be negative, leading to the inapplicability of standard dynamic programming arguments. Indeed, in this case it may be impossible to satisfy the local consistency requirements using only local transitions. To illustrate, consider a drift-free diffusion process of the form $(dx_{1t}, dx_{2t}) = (\sigma_1(x)dZ_t, \sigma_2(x)dZ_t)$ for some functions σ_1 and σ_2, where $(Z_t)_{t\geq 0}$ is one-dimensional Brownian motion. In this case we have

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2 \\ \sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}.$$

For the above expressions for probabilities to be non-negative we need $\sigma_1(x)^2/\Delta_1 \geq |\sigma_1(x)\sigma_2(x)|/\Delta_2$ and $\sigma_2(x)^2/\Delta_2 \geq |\sigma_1(x)\sigma_2(x)|/\Delta_1$. This is only true if $|\sigma_1(x)|/\Delta_1 = |\sigma_2(x)|/\Delta_2$ and so cannot be assured to hold everywhere for arbitrary σ_1 and σ_2.[9] A different construction is necessary, one that may call for non-local transitions. The remainder of this section is devoted to this construction, explaining the general process before turning to the problem in Definition 10.

We first illustrate the construction of a locally consistent chain for a drift-free process of the form $(dx_{1t}, dx_{2t}) = (\sigma_1(x)dZ_t, \sigma_2(x)dZ_t)$ on a domain of the form $[0, M_1] \times [0, M_2]$, for functions σ_1, σ_2 and constants $M_1, M_2 > 0$.[10] To this end, we fix integers $N_1, N_2 \geq 1$ and define $\Delta_i = M_i/N_i$ for $i = 1, 2$ and $S_h = \{\Delta_1, \ldots, M_1 - \Delta_1\} \times \{\Delta_2, \ldots, M_2 - \Delta_2\}$. We also write $(\overline{\sigma}_1, \overline{\sigma}_2) := (\sigma_1/\Delta_1, \sigma_2/\Delta_2)$ and $w := \overline{\sigma}_1/\overline{\sigma}_2$. An arbitrary member of S_h is of the form $(i\Delta_1, j\Delta_2)$ for $i \in \{1, \ldots, N_1 - 1\}$ and $j \in \{1, \ldots, N_2 - 1\}$. In constructing our chain we consider two cases that differ in the number of possible transitions. First, we

[9] Note that if σ_1 and σ_2 are constant then it is always possible to do this by appropriately adjusting the step sizes in each dimension.

[10] This construction is based on Section 5.9 of Kushner and Dupuis (2001).

suppose that at any $(x_1, x_2) = (\Delta_1 i, \Delta_2 j) \in S_h$, the transitions assume only three values, so $(\Delta x_1, \Delta x_2) \in \{(0,0), \pm(\Delta_1 m_1, \Delta_2 m_2)\}$ for some non-negative integers (m_1, m_2). We assume that the state may only leave the grid from a point adjacent to the boundary, and that the number of increments the state may move cannot exceed a fixed integer $\overline{m} \geq 1$. These two requirements translate into the restrictions

$$\begin{aligned} m_1 &\leq \min\{\overline{m}, \min\{i-1, N_1 - 1 - i\}\} \\ m_2 &\leq \min\{\overline{m}, \min\{j-1, N_2 - 1 - j\}\}. \end{aligned} \tag{35}$$

The set of non-zero integer pairs satisfying (35) is denoted $\Gamma(i,j)$. If the non-zero values occur with equal (possibly state-dependent) probability $\overline{p} \in (0, 1/2)$, then the mean consistency requirement is satisfied and the covariance consistency requirements are

$$\begin{aligned} 2\overline{p}\Delta_1^2 m_1^2 &= \Delta_t \Delta_1^2 \overline{\sigma}_1^2 + o(\Delta_t) \\ 2\overline{p}\Delta_1\Delta_2 m_1 m_2 &= \Delta_t \Delta_1 \Delta_2 \overline{\sigma}_1 \overline{\sigma}_2 + o(\Delta_t) \\ 2\overline{p}\Delta_2^2 m_2^2 &= \Delta_t \Delta_2^2 \overline{\sigma}_2^2 + o(\Delta_t). \end{aligned} \tag{36}$$

First, note that if we can find a pair of integers (m_1, m_2) satisfying $m_2 = m_1 \overline{\sigma}_2 / \overline{\sigma}_1$, then (36) will be satisfied with zero o terms if $\Delta_t = 2\overline{p} m_1^2 / \overline{\sigma}_1^2 = 2\overline{p} m_2^2 / \overline{\sigma}_2^2$. However, there may exist no such integer pair. Geometrically, this corresponds to a situation in which a shock to the diffusion process pushes the state in a direction in which there are no gridpoints. In general we choose a non-zero integer pair (m_1, m_2) to minimize $|m_2 - m_1 \overline{\sigma}_2 / \overline{\sigma}_1|$ and adjust Δ_t on a case-by-case basis so that it is non-zero and that either the first or third requirement in (36) holds: if $m_1 \geq m_2$, set $\Delta_t = 2\overline{p} m_1^2 / \overline{\sigma}_1^2$ and if $m_2 > m_1$, set $\Delta_t = 2\overline{p} m_2^2 / \overline{\sigma}_2^2$.

Figure 4 depicts this process for $|m| = 2$, at a point located more than $|m|$ points away from the boundary of the grid. The slope of the line is $\overline{\sigma}_2 / \overline{\sigma}_1$, the triangles and the square represent the transitions satisfying (35), and the square represents the (m_1, m_2) selected. This picture suggests that the approximation may be made more accurate by placing some probability on the triangle on the other side of the cross. To formalize this, expand the possible transitions to five points:

$$(\Delta x_1, \Delta x_2) \in \{(0,0), \pm(\Delta_1 m_{11}, \Delta_2 m_{12}), \pm(\Delta_1 m_{21}, \Delta_2 m_{22})\}$$

for some quadruple of integers $(m_{11}, m_{12}, m_{21}, m_{22})$ not all zero. For some $\overline{p} \in (0, 1/2)$, we declare the probability of $(\Delta x_1, \Delta x_2) \in \{(\Delta_1 m_{11}, \Delta_2 m_{12}), (\Delta_1 m_{21}, \Delta_2 m_{22})\}$ equal to $\overline{p}$, and given a point $(i\Delta_1, j\Delta_2) \in S_h$ and recalling $w := \overline{\sigma}_1 / \overline{\sigma}_2$, define

$$\begin{aligned} (m_{11}, m_{12}) &= \operatorname{argmin} \left\{ |j' - i'/w| \mid (i', j') \in \Gamma(i,j) \right\} \\ z &= \min(1, w)(m_{12} - m_{11}/w). \end{aligned} \tag{37}$$

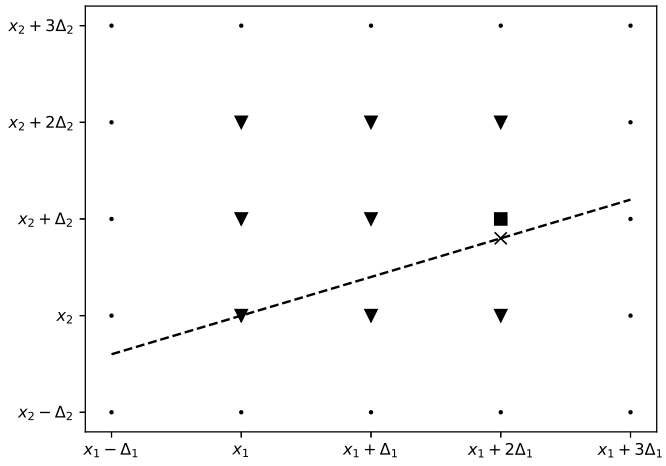

Figure 4: Optimal transition selection

The first pair of integers (m_{11}, m_{12}) will be the pair selected in the above three-point approximation, while $|z|$ may be interpreted as the distance between the two black dots in Figure 4 (note that z may be positive or negative). The second point, (m_{21}, m_{22}), will depend upon the signs of z and $w - 1$, as these determine whether the adjacent point lies to the left, right, above, or below the original point, and may be summarized as

$$\begin{aligned} m_{21} &= m_{11} + 1_{w\leq 1}(2 \times 1_{z>0} - 1) \\ m_{22} &= m_{12} + 1_{w>1}(2 \times 1_{z\leq 0} - 1). \end{aligned} \tag{38}$$

We place probability $\overline{p}(1-|z|)$ on (m_{11}, m_{12}) and $\overline{p}|z|$ on (m_{21}, m_{22}), and define the timestep to be $\Delta_t = \overline{\Delta}_t \overline{p}$, where

$$\overline{\Delta}_t = 1_{w>1} \times 2(m_{11})^2/\overline{\sigma}_1^2 + 1_{w\leq 1} \times 2(m_{12})^2/\overline{\sigma}_2^2. \tag{39}$$

The description of this five-point approximation is summarized as follows.

Algorithm 4 (Five-point approximation)**.** Given a (possibly state-dependent) transition probability $\overline{p} \in (0, 1/2)$, define an approximation to the process $(dx_{1t}, dx_{2t}) = (\sigma_1(x_t)dZ_t, \sigma_2(x_t)dZ_t)$ by declaring, for $x \in S_h$,

$$\begin{aligned} p(x_1 \pm m_{11}\Delta_1, x_2 \pm m_{12}\Delta_2) &= \overline{p}|z| \\ p(x_1 \pm m_{12}\Delta_1, x_2 \pm m_{22}\Delta_2) &= \overline{p}(1 - |z|) \end{aligned}$$

and $p(x_1, x_2) = 1 - 2\overline{p}$, where z and m are chosen according to (37) and (38), and the

timestep satisfies $\Delta_t = \overline{\Delta}_t \overline{p}$ for $\overline{\Delta_t}$ given by (39). Two of the three local consistency requirements may be satisfied exactly, and the remaining requirement has error equal to $2\overline{p}\Delta(w)^2|z|(1-|z|)$, where $\Delta(w) = 1_{w>1} \times \Delta_2 + 1_{w\leq 1} \times \Delta_1$.

The proof of the claims within Algorithm 4 requires only elementary algebra. Further details are given in Appendix B.1, and the accuracy of the above method is illustrated in the context of linear-quadratic problems in Appendix A.2. In what follows we apply Algorithm 4 in which the timestep is chosen to be independent of the state.

We now outline two simplifications of the problem in Definition 10 before describing our algorithm for competitive equilibria. Using Lemma 9 and denoting partial derivatives with subscripts, the value function of each agent satisfies $aW_{12}(a,x,\sigma) = (1-\gamma)W_2(a,x,\sigma)$ and $aW_{13}(a,x,\sigma) = (1-\gamma)W_3(a,x,\sigma)$. Substituting into the Hamilton-Jacobi-Bellman equation of the agent gives

$$\begin{aligned}\rho W = \max_{\bar{c},\bar{k}\geq 0} \; & \frac{\rho(\bar{c}a)^{1-\gamma}}{1-\gamma} + \big[r - \bar{c} + (\Pi - r)\overline{k}\big] aW_1 + \frac{\sigma^2 \overline{k}^2}{2} a^2 W_{11} \\ & + [\mu_x x + \sigma_x x \sigma \overline{k}(1-\gamma)] W_2 + \frac{\sigma_x^2 x^2}{2} W_{22} \\ & + [\theta(\overline{\sigma} - \sigma) + \sigma \overline{k} \sigma_\sigma (1-\gamma)] W_3 + \sigma_x x \sigma_\sigma W_{23} + \frac{\sigma_\sigma^2}{2} W_{33}.\end{aligned}$$

We now define $y_t := \ln a_t$ and note that by Ito's lemma we obtain a control problem with state (y,x,σ), controls (c,k), flow payoffs $\rho \bar{c}^{1-\gamma} e^{(1-\gamma)y}/(1-\gamma)$ and law of motion

$$\begin{aligned} dy_t &= \Big(r - \bar{c}_t + (\Pi - r)\overline{k}_t - \sigma_t^2 \overline{k}_t^2/2\Big) dt + \sigma_t \overline{k}_t dZ_t^{(1)}, \\ dx_t &= \big(\mu_x + \sigma_t \sigma_x (1-\gamma)\overline{k}_t\big) x_t dt + \sigma_x x_t dZ_t^{(2)}, \\ d\sigma_t &= \big(\theta(\overline{\sigma} - \sigma_t) + \sigma_t \sigma_\sigma (1-\gamma)\overline{k}_t\big) dt + \sigma_\sigma dZ_t^{(2)}, \end{aligned} \tag{40}$$

where $(Z^{(1)}, Z^{(2)})$ are now independent. Although this system is not the original one faced by the agent, the above homogeneity arguments show that it leads to the same value function. This does not eliminate all correlation between the state variables and so we still require non-local transitions. However, crucially, the diffusion terms of the states (x,σ) exhibiting high correlation are not controlled by the agent, and so the non-local transitions in Algorithm 4 depend only on prices and not the agent's choices.

To solve (40), first define the infinite grid $S_y = \{\dots, -\Delta_y, 0, \Delta_y, \dots\}$ for log wealth, and the finite grids $S_x = \{0, \Delta_x, \dots, 1-\Delta_x, 1\}$ and $S_\sigma = \{\underline{\Sigma}, \underline{\Sigma}+\Delta_\sigma, \dots, \overline{\Sigma}-\Delta_\sigma, \overline{\Sigma}\}$ for x and σ, where $\Delta_x = (1-0)/N_x$ and $\Delta_\sigma = (\overline{\Sigma} - \underline{\Sigma})/N_\sigma$, and define $S = S_y \times S_x \times S_\sigma$. To ensure that the process remains on the grid, we alter (30) so that σ_σ vanishes at $\underline{\Sigma}$ and $\overline{\Sigma}$. For clarity, we write the transition probabilities as the sum of the transitions for individual wealth, and the

drift and diffusion for the aggregate state, respectively, $p(X') = p^{(1)}(X') + p^{(2)}(X') + p^{(3)}(X')$ for any $X' = (y', x', \sigma') \neq X = (y, x, \sigma)$, with $p(X)$ chosen such that probabilities sum to unity. The transition probabilities are then

$$\begin{aligned}
p^{(1)}(y + \Delta_y, x, \sigma) &= \frac{\Delta_t}{\Delta_y^2}\Big(\sigma^2\overline{k}^2/2 + \Delta_y\big[r + (\Pi - r)\overline{k}\big]\Big) \\
p^{(1)}(y - \Delta_y, x, \sigma) &= \frac{\Delta_t}{\Delta_y^2}\Big(\sigma^2\overline{k}^2/2 + \Delta_y\Big[\overline{c} + \sigma^2\overline{k}^2/2\Big]\Big) \\
p^{(2)}(y, x \pm \Delta_x, \sigma) &= \frac{\Delta_t}{\Delta_x}\big([\mu_x x]^{\pm} + [\sigma_x x \sigma(1-\gamma)\overline{k}_t]^{\pm}\big) \\
p^{(2)}(y, x, \sigma \pm \Delta_\sigma) &= \frac{\Delta_t}{\Delta_\sigma}\big([\theta(\overline{\sigma} - \sigma)]^{\pm} + [\sigma_\sigma \sigma(1-\gamma)\overline{k}_t]^{\pm}\big) \\
p^{(3)}(y, x \pm m_{12}\Delta_x, \sigma \pm m_{13}\Delta_\sigma) &= \overline{p}|z| \\
p^{(3)}(y, x \pm m_{22}\Delta_x, \sigma \pm m_{23}\Delta_\sigma) &= \overline{p}(1 - |z|)
\end{aligned} \tag{41}$$

where the non-local transitions and timestep are given by Algorithm 4, and for simplicity we assume that Δ_t is independent of the state and adjust the variable $\overline{p}$ accordingly. Using Lemma 9, the Bellman operator $B_i(\Delta_t, \Delta_y)$ for type $i \in \{E, H\}$ may be viewed as operating on functions $V : S_x \times S_\sigma \to \mathbb{R}$ and given by

$$B_i(\Delta_t, \Delta_y)[V](x, \sigma) = (1 - \gamma) \max_{\overline{c}, \overline{k} \geq 0} \Delta_t \rho_i u(\overline{c}) + e^{-\rho_i \Delta_t} \mathbb{E}\left[\frac{V(x', \sigma')}{1 - \gamma} e^{(1-\gamma)(y' - y)}\right] \tag{42}$$

where the expectation operator on the right-hand side of (42) is with respect to the probabilities defined in (41).[11] For ease of notation, we write $V_{EjD} = 1_{\gamma<1} V_{EjF} + 1_{\gamma \geq 1} V_{EjB}$, where V_{EjF} and V_{EjB} denote forward and backward differences for the function V_E in the jth component, respectively.

4.3 Competitive equilibria

In this section we describe a pair of algorithms for computing competitive equilibria. Each algorithm consists of two distinct parts:

(i) Updating equilibrium quantities: take the value functions as given, and find the market-clearing interest rate and law of motion for the wealth share consistent with continuation values.

(ii) Updating value functions: take prices, law of motion of wealth share, and current guesses of the value functions as given, and update the value functions. For this step we consider two possibilities:

[11] Also note that $y' - y$ on the right-hand side of (42) is independent of y.

(a) Ignore the current guess of the value function and solve the stationary problem given the interest rate and law of motion of the state variable.

(b) Compute the value function by assuming that next period the payoff will be given by the current guess of the value function.

Following (ii)b leads to an algorithm similar to the iterative method employed in Brunnermeier and Sannikov (2016), while following (ii)a is our novel algorithm that exploits the stationarity of the individual problem. We first describe (i). In deriving policy functions we will assume that $\sigma_x \geq 0$, and verify ex-post that this is indeed the case. For the transition probabilities in Section 4.2, the policy functions are given in the following lemma, a proof of which is contained in Appendix B.3, along with explicit expressions for the constants E_1, E_2 and E_c.

Lemma 12. *For $i \in \{E, H\}$ consumption is given by $\bar{c}_i = \rho_i^{1/\gamma} V_i^{-1/\gamma} e^{\rho_i \Delta_t/\gamma} E_c$ and the capital policy function for the expert is*

$$\overline{k} = \frac{1}{\gamma\sigma^2}\left(E_1 E_2^{-1}[\Pi - r] + E_2^{-1}\sigma(\sigma_x x V_{E1D} + \sigma_\sigma V_{E2D})/V_E\right)^+$$

where E_1, E_2 and E_c depend only upon Δ_y and tend to unity as $\Delta_y \to 0$.

We now impose two requirements: the market for bonds clears and the law of motion of the aggregate state is consistent with individual policy functions. Using Lemma 12, the bond market-clearing condition becomes

$$\frac{1}{\gamma\sigma^2}\left(E_1 E_2^{-1}(\Pi - r) + E_2^{-1}\sigma(\sigma_x x V_{E1D} + \sigma_\sigma V_{E2D})/V_E\right)^+ x \leq 1 \tag{43}$$

with equality if $r > \underline{r}$. Note the inequality in (43) may be strict if the storage technology is utilized in equilibrium. The left-hand side of (43) is decreasing in r so there are two cases: if $\overline{k}(\underline{r})x \leq 1$ then $r = \underline{r}$; otherwise r solves $\overline{k}(r)x = 1$. Rearranging gives

$$r = \max\left\{\underline{r}, \Pi + E_1^{-1}\sigma(\sigma_x x V_{E1D} + \sigma_\sigma V_{E2D})/V_E - (E_1/E_2)^{-1}\gamma\sigma^2/x\right\}. \tag{44}$$

Substituting (44) into the expression for capital in Lemma 12 gives

$$\overline{k} = \min\left\{\frac{1}{\gamma\sigma^2}\left(E_1 E_2^{-1}(\Pi - \underline{r}) + E_2^{-1}\sigma(\sigma_x x V_{E1D} + \sigma_\sigma V_{E2D})/V_E\right), \frac{1}{x}\right\}^+. \tag{45}$$

Expressions (44) and (45) give the risk-free rate and capital policy function consistent with bond market-clearing. Imposing consistency between individual and aggregate laws of motion then gives the following.

Lemma 13. *If $V_{E1D} \leq 0$ everywhere, then the volatility of the wealth share consistent with individual optimization is given by*

$$\sigma_x x = \sigma x(1-x) \min\left\{ \frac{(E_1/E_2)[\Pi - \underline{r}]/\sigma^2 + \sigma_\sigma V_{E2D}/[\sigma E_2 V_E]}{\gamma - x(1-x)V_{E1D}/[E_2 V_E]}, 1/x \right\}^+. \tag{46}$$

For this volatility, the interest rate is (44) and the drift in the wealth share is

$$\mu_x x = x(1-x)\left[\left((\rho_H/V_H)^{1/\gamma} e^{\rho_H \Delta_t/\gamma} - (\rho_E/V_E)^{1/\gamma} e^{\rho_E \Delta_t/\gamma}\right) E_c + (\Pi - r)\overline{k} - \sigma^2 \overline{k}^2 x\right]$$

where $\overline{k}$ is given by (45) and Δ_t is constructed using σ_x following Algorithm 4.

Rearranging the discrete Bellman equation (42) for type $i \in \{E, H\}$ and dividing by the timestep gives an equation of the form $0 = \max_{\overline{c},\overline{k}} \rho_i u(\overline{c}) + T_i(\overline{c}, \overline{k}; \Delta_t, \Delta_y)V$ for the operator T_i given by

$$T_i(\overline{c}, \overline{k}; \Delta_t, \Delta_y)V(x, \sigma) = \frac{1}{\Delta_t}\left(e^{-\rho_i \Delta_t} \mathbb{E}\left[\frac{V(x', \sigma')}{1-\gamma} e^{(1-\gamma)(y'-y)} \right] - \frac{V(x, \sigma)}{1-\gamma} \right), \tag{47}$$

As in Section 2, it is convenient to note that $\lim_{\Delta_t, \Delta_y \to 0} T(\overline{c}, \overline{k}; \Delta_t, \Delta_y) \to \overline{T}(\overline{c}, \overline{k})$ for some well-defined operator $\overline{T}$, given explicitly in Lemma 17. We may now describe the two equilibrium algorithms.

Algorithm 5 (False transient approach)**.** Given a guess $\{r, \mu_x, \sigma_x\}$ for the interest rate and law of motion of the wealth share, and a guess for the value functions $\{V_E, V_H\}$:

(i) update the value functions using the Bellman operators in (42);

(ii) calculate the risk-free rate and law of motion of the wealth share using Lemma 13;

(iii) return to Step (i) with the new interest rate, law of motion of the wealth share, and value functions and repeat until convergence.

Algorithm 6 (Policy iteration approach)**.** Given a guess $\{r, \mu_x, \sigma_x\}$ for the interest rate and law of motion of the wealth share:

(i) solve $0 = \max_{\overline{c},\overline{k}} \rho_i u(\overline{c}) + \overline{T}_i(\overline{c}, \overline{k}; \Delta_t, \Delta_y)V$ using policy function iteration;

(ii) calculate the interest rate and law of motion of the wealth share using Lemma 13;

(iii) return to Step (i) with the new interest rate and law of motion of the wealth share, and repeat until convergence.

Before turning to a numerical illustration and comparison of Algorithm 5 and Algorithm 6 we note that in this environment there exists a convenient initial guess of equilibrium

quantities. As $\gamma \to 1$ one may check that $V_E, V_H \to 1$ everywhere, with policy functions satisfying $(\bar{c}_E, \bar{c}_H) \to (\rho_E, \rho_H)$ and $\overline{k} \to \min\left\{[\Pi - \underline{r}]/\sigma^2, 1/x\right\}^+$, respectively. The equilibrium interest rate is then $r = \max\left\{\underline{r}, \Pi - \sigma^2/x\right\}$, and the law of motion of the wealth share satisfies

$$(\mu_x x, \sigma_x x) = \left(\left[\rho_H - \rho_E + (\Pi - r)\overline{k} - \sigma^2 x \overline{k}^2\right] x(1-x), \sigma x(1-x)\overline{k}\right).$$

For $\underline{r} = -\infty$ we obtain $r = \Pi - \sigma^2/x$, $\mu_x x = \left[(\rho_H - \rho_E)x + \sigma^2(1/x - 1)\right](1-x)$ and $\sigma_x x = \sigma(1-x)$. In this case, as the wealth share of experts vanishes, the interest rate and drift in the wealth share diverge to negative and positive infinity, respectively. Further, the volatility of the wealth share is everywhere increasing with the exogenous volatility. This will not be true when there is a lower bound on the interest rate, since whenever the storage technology is utilized in equilibrium, reductions in exogenous uncertainty will *increase* the volatility in the wealth share.[12] These equilibrium quantities for the logarithmic case therefore serve as a useful initial guess for the general case.

Numerical illustration: We have solved an example of the competitive equilibrium for the above economy with the parameters:

$$\begin{aligned}&(\gamma = 2.0, \rho = (0.1, 0.075), \theta = 0.5, \sigma_\sigma = 0.15, \overline{\sigma} = 0.2, \Pi_E = 0.1, \underline{r} = 0,\\ &(\underline{\Sigma}, \overline{\Sigma}) = (0.1, 0.3), N = (120, 60), \overline{m} = 4, \Delta_t = 10^{-9}, \Delta_y = 10^{-4}).\end{aligned} \tag{48}$$

With a tolerance of 10^{-6} for both the individual problem in the supremum norm for $V^{\frac{1}{1-\gamma}}$ and the iterations of the equilibrium quantities in the norm $||(r, x\mu_x, x\sigma_x)|| := ||r||_\infty + ||x\mu_x||_\infty + ||x\sigma_x||_\infty$, our policy iteration algorithm converges in around five seconds beginning from an initial guess in which r, μ_x and σ_x are fixed at their logarithmic values, where we again use the standard scipy sparse solver (scipy.sparse.linalg.spsolve) in Python and an Intel Core i7-8650U processor. Figure 5 gives the interest rate and the investment function of experts, Figure 6 gives the drift and diffusion of the wealth share and Figure 7 gives the value functions.[13] Note that the domain for the investment function has been truncated to $[0.25, 1] \times [\underline{\Sigma} \times \overline{\Sigma}]$ as investment appears to diverge to negative infinity as the wealth share vanishes.

[12] This is reminiscent of (but distinct from) the the so-called "volatility paradox" highlighted by Brunnermeier and Sannikov (2016), in which illiquidity of capital leads to the non-mononicity of endogenous risk with respect to exogenous risk, even in the absence of a lower bound on the interest rate.

[13] These figures appear different from those in the original FRBC working paper due in part to a coding error for the updating of the law of motion for the wealth share that was corrected upon revision.

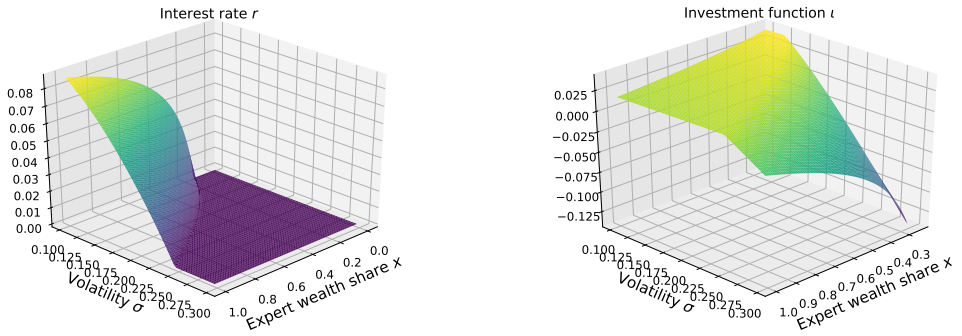

Figure 5: Interest rate and investment function

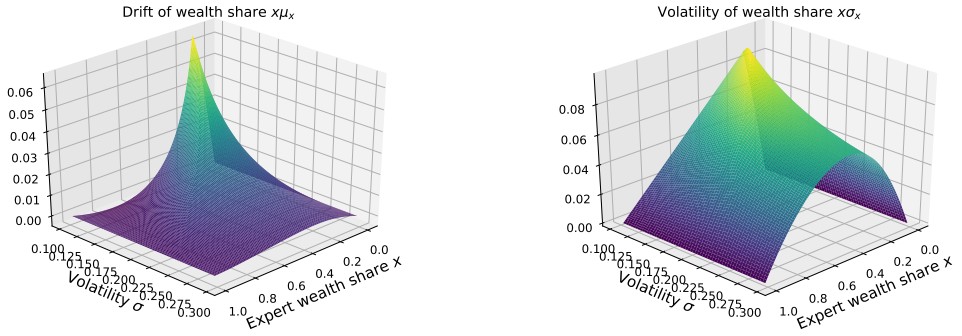

Figure 6: Drift and diffusion of wealth share

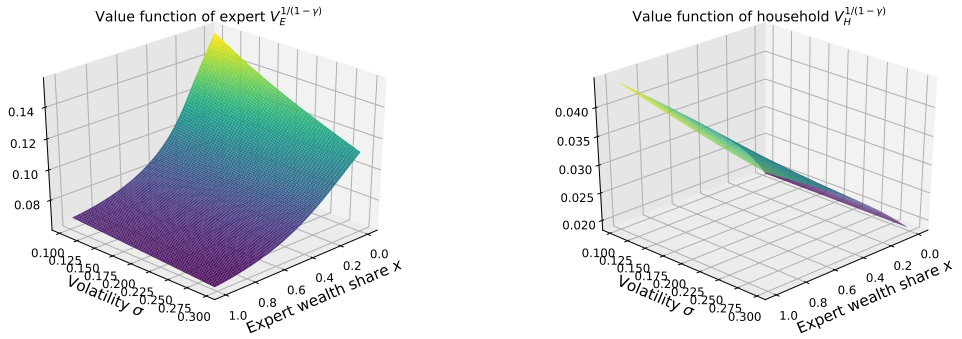

Figure 7: Value functions of expert and household

The lack of general closed-form solutions makes it difficult to systematically compare Algorithm 5 with Algorithm 6 for general preferences. Nonetheless, we believe that the policy

function approach in Algorithm 6 has several advantages over the false transient approach in Algorithm 5. First note that the chain construction and updating of all equilibrium quantities are common to both algorithms, and so they possess essentially the same level of programming difficulty, since we are only replacing a value function iteration step with a policy iteration step. The important differences between the two concern the stability of each algorithm and the speed of convergence as the timestep becomes small. In the absence of any changes in prices or the law of motion of the expert wealth share, the Bellman operator appearing in the false transient approach is a contraction with modulus equal to the rate of discount between successive periods. Consequently, as the grid size increases and an increasingly small timestep is necessary to ensure that probabilities remain within the unit interval, the value functions and equilibrium quantities update very slowly, and small changes between successive iterations provide no assurance of convergence.

We now illustrate the behavior over time of the false transient approach for various choices of the timestep and two separate initial conditions. We fix parameters at (48), except that we make the grid coarser and increase the timestep in order to give the false transient approach the best chance for rapid convergence. We now choose $N = (40, 20)$ and $\Delta_y = 10^{-1}$, and vary Δ_t between 10^{-3} and 3×10^{-3}.[14] We denote the interest rate and drift and diffusion of the wealth share found using the policy iteration approach above by $r_{\text{PI}}, \mu_{x\text{PI}}$ and $\sigma_{x\text{PI}}$, and the corresponding quantities found in the ith stage of the false transient approach by $r^i_{\text{FT}}, \mu^i_{x\text{FT}}$ and $\sigma^i_{x\text{FT}}$. We then define

$$E_i = ||r^i_{\text{FT}} - r_{\text{PI}}||_\infty + ||x\mu^i_{x\text{FT}} - x\mu_{x\text{PI}}||_\infty + ||x\sigma^i_{x\text{FT}} - x\sigma_{x\text{PI}}||_\infty.$$

Figure 8 plots $\log_{10}(E_i)$ against i for various values of Δ_t. The figure on the left begins at prices and the law of motion associated with logarithmic utility (and hence the value functions are identically unity), while the figure on the right begins at the values found with the policy iteration algorithm. The left figure shows that as we increase the number of iterations, the quantities found using the false transient algorithm appear to be converging to those found using the policy iteration approach, but the convergence is very slow, even for a fairly coarse grid, taking several minutes to compute. The figure on the right shows that when beginning at the value found by the policy iteration approach, the values found with the false transient approach remain close to the initial condition, providing confidence that the computed solution does indeed constitute a competitive equilibrium.

We emphasize that we do not claim guaranteed convergence of our algorithm to the true competitive equilibrium. Indeed, for some parameters we have observed a failure of convergence of prices and the law of motion within the above tolerance, often near regions in

[14] For $\Delta_t = 4 \times 10^{-3}$, the false transient approach broke down, as expressions for probabilities failed to remain in the unit interval.

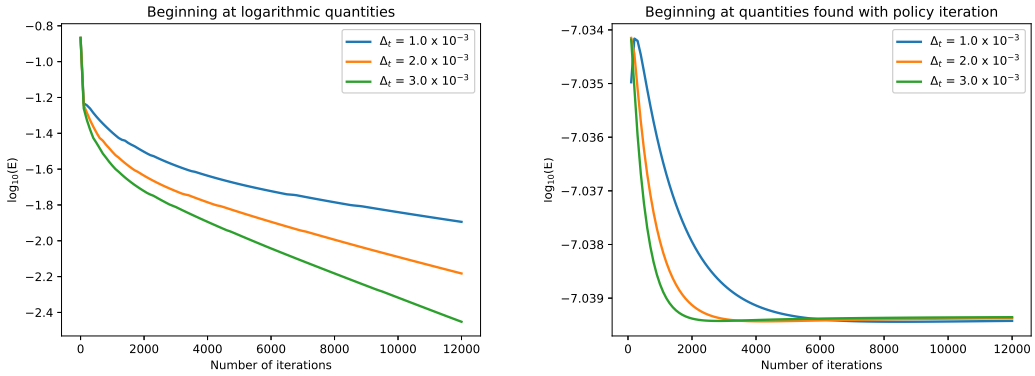

Figure 8: Log_{10} difference between policy iteration and false transient quantities

which the interest rate hits the lower bound $\underline{r}$, at which point the policy iteration algorithm appears to oscillate. However, in light of the preceding analysis and the absence of general existence and convergence results in the literature, we believe the policy iteration algorithm will prove useful to practitioners in the burgeoning macrofinance literature surveyed in Brunnermeier and Sannikov (2016), and so we leave further exploration of competitive equilibria to future work.

5 Conclusion

In this paper we explore several applications of the Markov chain approximation (MCA) method of Harold Kushner and Paul Dupuis to optimal control problems in economics, illustrating some unutilized benefits. We first show that for certain choices of the approximating chain, the MCA method with policy function iteration coincides with a limiting version of the widely-used implicit finite-difference scheme of Achdou et al. (2022). We then demonstrate the benefits of a more general specification by means of two classes of examples. In the first, we use variations of modified policy function iteration to solve income fluctuation problems, both with and without discrete choices. In the second, we show how the MCA method may be applied to problems with high correlation among state variables, and illustrate the benefit of exploiting stationarity wherever applicable. In both cases, the MCA is robust, easy to apply and can result in an increase in speed of more than an order of magnitude over more commonly-applied methods.

References

Achdou, Y., Han, J., Lasry, J.-M., Lions, P.-L., and Moll, B. Income and wealth distribution in macroeconomics: A continuous-time approach. *The Review of Economic Studies*, 89

(1):4–86, January 2022. ISSN 0034-6527. doi:10.1093/restud/rdab002.

Barczyk, D. and Kredler, M. Altruistically motivated transfers under uncertainty. *Quantitative Economics*, 5(3):705–749, 2014a. doi:10.3982/QE353.

Barczyk, D. and Kredler, M. A Dynamic Model of Altruistically-Motivated Transfers. *Review of Economic Dynamics*, 17(2):303–328, 2014b. doi:10.1016/j.red.2013.07.002.

Bonnans, J. F. and Zidani, H. Consistency of Generalized Finite Difference Schemes for the Stochastic HJB Equation. *SIAM Journal on Numerical Analysis*, 41(3):1008–1021, January 2003. ISSN 0036-1429. doi:10.1137/S0036142901387336.

Bonnans, J. F., Ottenwaelter, É., and Zidani, H. A fast algorithm for the two dimensional HJB equation of stochastic control. *ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique*, 38(4):723–735, 2004. doi:10.1051/m2an:2004034.

Brunnermeier, M. K. and Sannikov, Y. A Macroeconomic Model with a Financial Sector. *American Economic Review*, 104(2):379–421, February 2014. ISSN 0002-8282. doi:10.1257/aer.104.2.379.

Brunnermeier, M. K. and Sannikov, Y. Macro, money, and finance: A continuous-time approach. In *Handbook of Macroeconomics*, volume 2, pages 1497–1545. Elsevier, 2016. doi:10.1016/bs.hesmac.2016.06.002.

Candler, G. V. Finite-Difference Methods for Continuous-Time Dynamic Programming. In Marimon, R. and Scott, A., editors, *Computational Methods for the Study of Dynamic Economies*, pages 172–194. Oxford University Press, October 2001. ISBN 978-0-19-924827-8. doi:10.1093/0199248273.003.0008.

Carroll, C. D. The method of endogenous gridpoints for solving dynamic stochastic optimization problems. *Economics Letters*, 91(3):312–320, 2006. doi:10.1016/j.econlet.2005.09.013.

d'Avernas, A. and Vandeweyer, Q. A solution method for continuous-time general equilibrium models. 2019.

Fella, G. A generalized endogenous grid method for non-smooth and non-concave problems. *Review of Economic Dynamics*, 17(2):329–344, 2014. doi:10.1016/j.red.2013.07.001.

Golosov, M. and Lucas Jr, R. E. Menu costs and Phillips curves. *Journal of Political Economy*, 115(2):171–199, 2007. doi:10.1086/512625.

Kushner, H. J. and Dupuis, P. *Numerical Methods for Stochastic Control Problems in Continuous Time*, volume 24. Springer Science & Business Media, 2001. doi:10.1007/978-1-4684-0441-8.

Puterman, M. L. and Brumelle, S. L. On the convergence of policy iteration in stationary dynamic programming. *Mathematics of Operations Research*, 4(1):60–69, 1979. doi:10.1287/moor.4.1.60.

Puterman, M. L. and Shin, M. C. Modified policy iteration algorithms for discounted Markov decision problems. *Management Science*, 24(11):1127–1137, 1978. doi:10.1287/mnsc.24.11.1127.

Stokey, N. L. *Recursive Methods in Economic Dynamics*. Harvard University Press, October 1989. ISBN 978-0-674-73518-7. doi:10.2307/j.ctvjnrt76.

Tourin, A. Optimal stochastic control, stochastic target problems, and backward SDE. In *Optimal Stochastic Control, Stochastic Target Problems, and Backward SDE*, volume 29 of *Fields Institute Monographs*, chapter Introduction to Finite Differences Methods, pages 201–212. Springer, New York, NY, 2013. ISBN 978-1-4614-4286-8. doi:10.1007/978-1-4614-4286-8_13.

A Linear-quadratic regulator problem

To verify the accuracy of the algorithms, this section records closed-form expressions for a class of linear-quadratic control problems. Lemma 14 treats the standard case of an infinite-horizon linear-quadratic regulator problem, while Lemma 15 treats the (slightly non-standard) case in which volatility is linear in the state variable. The former will be useful for illustrating the benefits associated with modified policy function iteration, while the latter will illustrate the applications of non-local transitions. Suppose that the objective to be maximized is

$$\mathbb{E}\left[\int_0^\infty e^{-\rho t}F(x_t,u_t)dt\right]$$

where for some symmetric positive definite matrices Q and R the flow payoffs are given by

$$F(x,u)=-\frac{1}{2}x^TQx-\frac{1}{2}u^TRu \tag{49}$$

where $x\in\mathbb{R}^n$ and $u\in\mathbb{R}^q$ are the state and control vectors, respectively, for some integers $n,q\geq 1$. Now suppose that for some $A\in\mathbb{R}^{n\times n}, B\in\mathbb{R}^{n\times q}$ and $\sigma:\mathbb{R}^n\to\mathbb{R}^{n\times m}$ the law of motion for the state is

$$dx_t=[Ax_t+Bu_t]dt+\sigma(x_t)dZ_t$$

where $Z := (Z_t)_{t\geq 0}$ is m-dimensional Brownian motion. Write $\mu(x,u) = Ax + Bu$ for the drift as a function of the state and controls, and note that the Hamilton-Jacobi-Bellman equation is

$$\rho V(x) = \max_{u\in\mathbb{R}^q} F(x,u) + \sum_{i=1}^{n} \mu_i(x,u) V_i(x) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} V_{ij}(x) \tag{50}$$

where $a_{ij}(x) = (\sigma(x)\sigma(x)^T)_{ij}$. If σ is constant then we obtain the following.

Lemma 14. *The solution to (50) is $V(x) = -x^T Px/2 - d$, where P is a symmetric matrix that solves*

$$\rho P = Q + PA + A^T P - P^T B R^{-1} B^T P \tag{51}$$

the constant term is $d = [2\rho]^{-1}\text{trace}(\sigma\sigma^T P)$, and the policy function is

$$u = -R^{-1}B^T Px. \tag{52}$$

Proof of Lemma 14. For an arbitrary symmetric positive definite matrix M we have

$$\frac{d}{dx_k} x^T M x = \sum_{i=1}^{n}\sum_{j=1}^{n} M_{ij}\frac{d}{dx_k}(x_i x_j) = \sum_{j=1}^{n} M_{kj}x_j + \sum_{i=1}^{n} M_{ik}x_i.$$

Using the symmetry of M, we have $\nabla x^T Mx/2 = Mx$ and $Hx^T Mx/2 = M$. Assuming a solution of the form $V(x) = -x^T Px/2 + D$, substitution into the right-hand side of (50) gives

$$\max_{u\in\mathbb{R}^q} -\frac{1}{2}x^T Qx - \frac{1}{2}u^T Ru - \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\sum_{k=1}^{n} A_{ik}x_k + \sum_{k=1}^{q} B_{ik}u_k\right)P_{ij}x_j - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}P_{ij}.$$

For any $i = 1,\ldots,n$ we have

$$-\frac{d}{du_i}\sum_{h=1}^{n}\sum_{g=1}^{n} P_{hg}\left(\sum_{k=1}^{n} A_{hk}x_k + \sum_{k=1}^{q} B_{hk}u_k\right)x_g = -\sum_{h=1}^{n}\sum_{g=1}^{n} P_{hg}B_{hi}x_g.$$

The first-order conditions become $\sum_{k=1}^{n} R_{ik}u_k = -\sum_{h=1}^{n}\sum_{g=1}^{n} P_{hg}B_{hi}x_g$, which is exactly (52). Substitution then gives

$$\sum_{i=1}^{n}\mu_i(x,u)V_i(x) = -\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\sum_{k=1}^{n} A_{ik}x_k + \sum_{k=1}^{q} B_{ik}u_k\right)P_{ij}x_j = -\frac{1}{2}x^T(A^T P + PA)x - u^T B^T Px$$

where we used the fact that $x^T(E^T + E)x/2 = x^T Ex$ for all x and any matrix E. Using this and substituting into the Hamilton-Jacobi-Bellman equation then gives

$$-\frac{\rho}{2}x^T Px - \rho d = -\frac{1}{2}x^T Qx - \frac{1}{2}x^T(A^T P + PA)x - \frac{1}{2}u^T Ru - u^T B^T Px - \frac{1}{2}\text{trace}(\Sigma\Sigma^T P) \tag{53}$$

where we used the fact that $\sum_{i=1}^n \sum_{j=1}^n a_{ij} P_{ij} = \text{trace}(\Sigma\Sigma^T P)$, since P is symmetric. Finally,

$$\begin{aligned} -\frac{1}{2} u^T R u - u^T B^T P x &= -\frac{1}{2}(R^{-1}B^T P x)^T R R^{-1} B^T P x + (B R^{-1} B^T P x)^T P x \\ &= \frac{1}{2} x P^T B R^{-1} B^T P x \end{aligned}$$

so substituting into (56) and equating quadratic and constant parts gives the result. □

Note that P solves (51) if and only if it solves the undiscounted problem

$$0 = Q + P(A - \rho I/2) + (A - \rho I/2)^T P - P^T B R^{-1} B^T P$$

and so the analysis for the discounted case follows from results from the undiscounted case, with A replaced by $A - \rho I/2$. The standard linear-quadratic regulator problem analyzed in Lemma 14 assumes constant volatility. To illustrate the flexibility of the method in treating some cases that appear with high correlation, we consider here an extension of the standard linear-quadratic framework in which the volatility is linear in the state variables. The flow payoff remains (49) for some symmetric positive definite matrices Q and R, but the law of motion for the state variables is now $dx_t = [Ax_t + Bu_t]dt + \sigma x_t dZ_t$ where $\sigma \in \mathbb{R}^{n \times n}$ and Z is scalar Brownian motion.

Lemma 15. *If σ is a multiple of the identity, then the value function is $V(x) = -x^T P x/2$ where P solves*

$$0 = Q + P(A + [\sigma^2 - \rho]I/2) + (A + [\sigma^2 - \rho]I/2)^T P - P^T B R^{-1} B^T P. \tag{54}$$

Proof. The Hamilton-Jacobi-Bellman equation remains

$$\rho V(x) = \max_{u \in \mathbb{R}^q} F(x,u) + \sum_{i=1}^n \mu_i(x,u) V_i(x) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n a_{ij}(x) V_{ij}(x) \tag{55}$$

where $a_{ij}(x) = (\sigma x x^T \sigma^T)_{ij}$. If we again assume a solution of the form $V(x) = -x^T P x/2$ for some symmetric positive definite matrix P then the expressions for both the derivatives and optimal policy remain unchanged and the Hamilton-Jacobi-Bellman equation reduces to

$$-\frac{\rho}{2} x^T P x = -\frac{1}{2} x^T Q x - \frac{1}{2} x^T (A^T P + P A) x + \frac{1}{2} x P^T B R^{-1} B^T P x - \frac{1}{2} x^T \big(\sigma^T P \sigma\big) x \tag{56}$$

where we used

$$\frac{1}{2} x^T \big(\sigma^T P \sigma\big) x = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n P_{ij} \sum_{k=1}^n \sum_{h=1}^n \sigma_{ik} \sigma_{jh} x_k x_h$$

The equation we must solve is then

$$0 = Q + P(A - \rho I/2) + (A - \rho I/2)^T P - P^T B R^{-1} B^T P + \sigma^T P \sigma. \tag{57}$$

If σ is a multiple of the identity then $\sigma^T P \sigma = (P\sigma^2 I + \sigma^2 I P)/2$ and (59) reduces to (54). □

Lemma 15 will be useful to illustrate how the FSMC method may be used to deal with problems for which the noise terms are perfectly correlated. We now use the closed-form expressions for value

functions and policy functions in Lemma 14 and Lemma 15 to verify the accuracy of the Markov chain approximation method. There are many tests that we could conduct. We choose only a select few to illustrate the points highlighted in the main text and to provide confidence in the results recorded in the main text.

A.1 Independent noise

We first consider a three-dimensional problem with independent noise terms. This corresponds to the above situation in which σ is a diagonal matrix with main diagonal written $[\sigma_0\ \sigma_1\ \sigma_2]$. We consider domains in $\mathbb{R}^3$ of the form $[0, M_0] \times [0, M_1] \times [0, M_2]$ for constants M_0, M_1 and M_2. Now define $\Delta_i = M_i/N_i$ and $S_i = \{\Delta_i, \ldots, M_i - \Delta_i\}$ for $i = 0, 1, 2$ and we let our grids be $S = S_0 \times S_1 \times S_2$ and adopt the transition probabilities

$$\begin{aligned} p(x_0 \pm \Delta_0, x_1, x_2) &= \frac{\Delta_t}{\Delta_0^2}\left(\frac{\sigma_0^2}{2} + \Delta_0 (Ax + Bu)_0^{\pm}\right) \\ p(x_0, x_1 \pm \Delta_1, x_2) &= \frac{\Delta_t}{\Delta_1^2}\left(\frac{\sigma_1^2}{2} + \Delta_1 (Ax + Bu)_1^{\pm}\right) \\ p(x_0, x_1, x_2 \pm \Delta_2) &= \frac{\Delta_t}{\Delta_2^2}\left(\frac{\sigma_2^2}{2} + \Delta_2 (Ax + Bu)_2^{\pm}\right). \end{aligned} \tag{58}$$

We wish to avoid case-by-case technicalities and so consider problems with a single control, so that R may be interpreted as a scalar that we normalize as $R = 1$, and we write $B = [b_0, b_1, b_2]^T$. The Riccati equation and policy function from Lemma 14 become

$$\begin{aligned} \rho P &= Q + PA + A^T P - P^T B B^T P \\ u &= -B^T P x. \end{aligned} \tag{59}$$

We choose parameters for which the drift is always negative. The maximization becomes

$$\max_{u \le 0} \ -\frac{1}{2}u^2 + e^{-\rho \Delta_t}\left(b_0 V_B^0 + b_1 V_B^1 + b_2 V_B^2\right) u$$

and so the optimal control is obviously $u = \min\{e^{-\rho\Delta_t}\left(b_0 V_B^0 + b_1 V_B^1 + b_2 V_B^2\right), 0\}$. We will choose our timestep to be as large as possible while ensuring that the expressions for probabilities lie in the unit interval. If we restrict attention to controls such that $Ax + Bu \le 0$ in each component the above probabilities will lie in the unit interval provided

$$1 \ge \Delta_t \left(\frac{\sigma_0^2}{\Delta_0^2} + \frac{\sigma_1^2}{\Delta_1^2} + \frac{\sigma_2^2}{\Delta_2^2} + \frac{(Ax + Bu)_0^-}{\Delta_0} + \frac{(Ax + Bu)_1^-}{\Delta_1} + \frac{(Ax + Bu)_2^-}{\Delta_2}\right).$$

To use the above to obtain an appropriate state-dependent timestep we require a bound on the control u. For this we choose u to be $3 \times$ the optimal policy function. Table 6 and Table 7 document the time until convergence for modified policy function iteration and the associated average percentage error with the closed-form solution for the value function, for the parameters

$$(\rho, Q, A, B, \sigma, (M_1, M_2, M_3)) = 0.1, I_3, 0.01 \times I_3, (0.025, 0.025, 0.025), 0.4 \times I_3, (10, 10, 10))$$

Grid size	PFI	VFI	$k = 10$	$k = 50$	$k = 100$	$k = 200$
(10, 10, 10)	0.097	1.730	0.197	0.073	0.076	0.083
(20, 20, 20)	1.576	13.320	1.437	0.419	0.249	0.285
(30, 30, 30)	15.583	76.555	8.213	2.121	1.281	0.885
(40, 40, 40)	148.697	265.587	28.631	7.348	4.261	3.014

Table 6: Time until convergence in controlled LQ problem: MPFI algorithm

Grid size	PFI	VFI	$k = 10$	$k = 50$	$k = 100$	$k = 200$
(10, 10, 10)	3.685	3.685	3.685	3.685	3.685	3.685
(20, 20, 20)	1.369	1.369	1.369	1.369	1.369	1.369
(30, 30, 30)	0.764	0.764	0.764	0.764	0.764	0.764
(40, 40, 40)	0.507	0.507	0.507	0.507	0.507	0.507

Table 7: L_1 norm of percentage error in controlled LQ problem: MPFI algorithm

where I_3 denotes the 3×3 identity matrix, and a tolerance in the supremum norm between iterations of 10^{-6}. Table 8 and Table 9 perform the analogous exercise for generalized policy function iteration.

A.2 Perfectly correlated noise

Algorithm 4 is intuitive to us as it requires only a naive minimization of errors in the local consistency requirements. To improve accuracy, however, we may draw upon the geometric analysis of Bonnans et al. (2004), which in turn builds upon the more general framework of Bonnans and Zidani (2003). Note that the local consistency requirements for our drift-free process will hold exactly if the transitions are symmetric about the origin and the non-negative components are selected from

Grid size	PFI	$k = 0$	$k = 10$	$k = 50$	$k = 100$	$k = 200$
(10, 10, 10)	0.117	0.957	0.113	0.092	0.097	0.107
(20, 20, 20)	1.603	9.950	1.077	0.301	0.280	0.324
(30, 30, 30)	15.802	64.712	7.016	1.944	1.219	0.898
(40, 40, 40)	149.726	243.207	26.408	6.870	4.154	2.876

Table 8: Time until convergence in controlled LQ problem: Generalized MPFI algorithm

Grid size	PFI	$k=0$	$k=10$	$k=50$	$k=100$	$k=200$
(10, 10, 10)	0.952	0.952	0.952	0.952	0.952	0.952
(20, 20, 20)	0.472	0.472	0.472	0.472	0.472	0.472
(30, 30, 30)	0.314	0.314	0.314	0.314	0.314	0.314
(40, 40, 40)	0.236	0.236	0.236	0.236	0.236	0.236

Table 9: L_1 norm of percentage error in controlled LQ problem: Generalized MPFI algorithm

a subset $\hat{\Gamma}(i,j) \subseteq \Gamma(i,j)$ with probabilities $\{\Delta_t \eta_\xi | \xi \in \hat{\Gamma}(i,j)\}$ satisfying

$$\sum_{\xi \in \hat{\Gamma}(i,j)} \eta_\xi \xi \xi^T = \begin{bmatrix} \overline{\sigma}_1^2/2 & \overline{\sigma}_1\overline{\sigma}_2/2 \\ \overline{\sigma}_1\overline{\sigma}_2/2 & \overline{\sigma}_2^2/2 \end{bmatrix}. \tag{60}$$

The set of sums of the form on the left-hand side of (60) is a convex cone. Bonnans et al. (2004) approximate a solution to (60) by projecting the right-hand side onto this cone. To see how, denote the set of positive semi-definite 2×2 matrices by PSD_2 and define $F : PSD_2 \to \mathbb{R}^3$ and $R : \mathbb{R}^3 \to \mathbb{R}^3$ by $F(a) = (a_{11}, \sqrt{2}a_{12}, a_{22})^T$ and $R(z) = ((z_1 - z_3)/\sqrt{2}, z_2, (z_1 + z_3)/\sqrt{2})$, respectively. Note that the function R is a rotation and that the cone $C := \{R(F(a)) | a \in PSD_2\}$ points "upwards" in $\mathbb{R}^3$. We define $H : C \to \mathbb{R}^2$ by $H(x) = (x_1/x_3, x_2/x_3)$ and identify the set of covariance matrices with the disc $D := H(R(F(PSD_2)))$. The boundary of the set D corresponds to those covariance matrices associated with perfectly correlated diffusion processes, in which case $a_{11}a_{22} = a_{12}^2$. We identify integer pairs (p,q) with the fraction q/p, and given q/p and q'/p' define the *child* $q''/p'' := (q+q')/(p+p')$. Given a point in the grid, we choose transitions $\hat{\Gamma}(i,j)$ by beginning near the origin before passing from adjacent points to their children. For any q/p write $\xi_{pq} = [p,q]^T$ and $X_{pq} = \xi_{pq}\xi_{pq}^T = [p,q]^T[p,q]$, and given q/p and q'/p' define $H(q/p, q'/p')$ to be the plane generated by X_{pq} and $X_{p'q'}$ and identify it with its range under F. The associated projection operator is $P_H(q/p, q'/p') = A(A^TA)^{-1}A^T$, where

$$A^T = \begin{bmatrix} p^2 & \sqrt{2}pq & q^2 \\ (p')^2 & \sqrt{2}p'q' & (q')^2 \end{bmatrix}.$$

We will be concerned only with the case in which $\sigma_1, \sigma_2 \geq 0$. For a fixed $\overline{m} \geq 1$ we approximate a solution to (60) as follows, where we write $a_{ij}^h = (\sigma\sigma^T)_{ij}/(\Delta_i\Delta_j)$.[15]

Algorithm 7 (Bonnans et al. (2004) approximation)**.** Define $(q/p, q'/p') = (1/1, 0/1)$ if $a_{11}^h \geq a_{22}^h$ and $(q/p, q'/p') = (1/0, 1/1)$ otherwise. Notice that we have $q/p \geq a_{22}^h/a_{11}^h \geq q'/p'$. We then update as follows:

[15] We are using the fact that in this problem, noise is perfectly correlated, which allows us to simplify the algorithm. For example, our covariance matrix is never diagonally dominant, and for simplicity we omit reference to a tolerance level.

Grid size	$\overline{m}=2$	$\overline{m}=4$	$\overline{m}=6$	$\overline{m}=8$	$\overline{m}=10$
(50, 50)	5.45	2.37	2.16	2.30	2.58
(100, 100)	4.68	1.56	1.16	1.10	1.14
(150, 150)	4.43	1.30	0.88	0.77	0.74
(200, 200)	4.31	1.17	0.74	0.62	0.57

Table 10: L_1 norm of percentage error in correlated LQ problem: three-point algorithm

Grid size	$\overline{m}=2$	$\overline{m}=4$	$\overline{m}=6$	$\overline{m}=8$	$\overline{m}=10$
(50, 50)	2.50	0.56	0.36	0.43	0.54
(100, 100)	2.44	0.48	0.21	0.16	0.18
(150, 150)	2.42	0.46	0.18	0.11	0.11
(200, 200)	2.41	0.45	0.16	0.10	0.08

Table 11: L_1 norm of percentage error in correlated LQ problem: five-point algorithm

(i) If $\max\{p+p', q+q'\} > \overline{m}$ then stop. Otherwise, go to Step (ii).

(ii) If $a^h_{22}/a^h_{11} \geq (q+q')/(p+p')$, return to Step (i) with $q'/p' = (q+q')/(p+p')$. Otherwise, return to Step (i) with $q/p = (q+q')/(p+p')$.

We then choose $S_h = \{(p,q),(p',q')\}$ and $\eta_{pq}, \eta_{p'q'} \geq 0$ satisfying

$$\eta_{pq} X_{pq} + \eta_{p'q'} X_{p'q'} = P_H(q/p, q'/p') a^h$$

and for these weights define probabilities $p_i = \eta_i \Delta_t$. Finally, we choose the timestep so that $p_0 + p_1 = \overline{p}$ for some given (possibly state-dependent) $\overline{p} \in (0, 1/2)$. This amounts to choosing $\Delta_t(\eta_0 + \eta_1) = \overline{p}$, with implied probabilities $p_i = \overline{p}\eta_i/(\eta_0 + \eta_1)$.

We now verify the accuracy of the Markov chains constructed in both Section 4.2 and in Algorithm 7 for the case of perfectly correlated diffusion processes in two dimensions. In all of the following cases we set Q equal to the identity, omit drift and controls, set $\sigma = 0.3 \times I_3$, $\rho = 0.15$ and $M_1 = M_2 = 1$, and set $\overline{p} = 0.001$ everywhere (adjusting the timestep accordingly in each algorithm). Table 10 documents the average of the percent difference between the true and computed value function with three points, where $\overline{m}$ refers to the largest size of the non-local transitions. Table 11 documents the same quantities for the case with five points (Algorithm 4) and Table 12 gives the same for the method of Bonnans et al. (2004). As can be seen, for all grids considered these methods are strictly increasing in their accuracy (3-point, then 5-point, then Bonnans et al. (2004)).

Grid size	$\overline{m}=2$	$\overline{m}=4$	$\overline{m}=6$	$\overline{m}=8$	$\overline{m}=10$
(50, 50)	0.53	0.15	0.10	0.09	0.09
(100, 100)	0.49	0.10	0.04	0.03	0.03
(150, 150)	0.48	0.09	0.03	0.02	0.01
(200, 200)	0.47	0.08	0.03	0.02	0.01

Table 12: L_1 norm of percentage error in correlated LQ problem: Bonnans et al. (2004) algorithm

B Macrofinance notes

B.1 Problems with highly correlated state variables

Proof of claims in Algorithm 4. Define the candidate o terms in (36) as e_{11}, e_{12} and e_{22}. We proceed on a case-by-case basis, depending upon the sign of $w-1$. The errors are

$$\begin{aligned} e_{11} &= \mathbb{E}[(\Delta x_1)^2] - \Delta_t \sigma_1^2 \\ e_{12} &= \mathbb{E}[(\Delta x_1)(\Delta x_2)] - \Delta_t \sigma_1 \sigma_2 \\ e_{22} &= \mathbb{E}[(\Delta x_2)^2] - \Delta_t \sigma_2^2. \end{aligned} \tag{61}$$

Condition $w > 1$ is equivalent to $\sigma_1/\Delta_1 > \sigma_2/\Delta_2$ or $\sigma_1/\sigma_2 > \Delta_1/\Delta_2$. In this case $m_{11} = m_{21}$ and $\Delta_t = 2\overline{p}\Delta_1^2 m_{11}^2/\sigma_1^2$, so the first term in (61) is $e_{11} = 2\overline{p}\Delta_1^2 m_{11}^2 - \Delta_t\sigma_1^2 = 0$. We also have $m_{22} = m_{12} + 2(z \le 0) - 1$, and so the second term in (61) becomes

$$\begin{aligned} e_{12} &= 2\overline{p}\Delta_1\Delta_2 m_{11}((1-|z|)m_{12} + |z|[m_{12} + 2(z \le 0) - 1]) - \Delta_t\sigma_1\sigma_2 \\ &= 2\overline{p}\Delta_1\Delta_2 m_{11}(|z|[2(z \le 0) - 1] + z) = 0. \end{aligned}$$

Finally, the third error term in (61) simplifies to

$$\begin{aligned} e_{22} &= 2\overline{p}\Delta_2^2\big(|z|m_{12}^2 + (1-|z|)[m_{12} + 2(z \le 0) - 1]^2\big) - \Delta_t\sigma_2^2 \\ &= 2\overline{p}\Delta_2^2\big[m_{12}^2 + (1-|z|)[2(2(z \le 0) - 1)m_{12} + (2(z \le 0) - 1)^2] - \sigma_2^2(\Delta_1^2/\Delta_2^2)m_1^2/\sigma_1^2\big] \\ &= 2\overline{p}\Delta_2^2\big[|z|[2(2(z \le 0) - 1)m_2 + (2(z \le 0) - 1)^2] + m_2^2 - m_1^2/w^2\big]. \end{aligned}$$

Using $z = m_{12} - m_{11}/w$ we write $(m_{11}/w)^2 = (m_{12} - z)^2 = m_{12}^2 - 2m_{12}z + z^2$ to note

$$e_{22} = 2\overline{p}\Delta_2^2\big[|z|[(4(z \le 0) - 2)m_{12} + 1] + 2m_{12}z - z^2\big] = 2\overline{p}\Delta_2^2(|z| - |z|^2)$$

as claimed. The case with $w \le 1$ is symmetric. □

B.2 Evolution of wealth shares

We now determine the law of motion for the wealth share x, by aggregating over the choices of experts and households. Lemma 16 shows how the law of motion of the wealth share depends upon the law of motion of the wealth of the individual agents.

Lemma 16. *Suppose that $da_{it}/a_{it} = \mu_{it}dt + \sigma_{it}dZ_t$ for $i \in \{E, H\}$ and that $x := N/(qK)$, where N is the aggregate wealth of experts. Then $dx_t = x_t\mu_x dt + x_t\sigma_x dZ_t$ where μ_x and σ_x are*

$$\begin{aligned} x\mu_x &= x(1-x)(\mu_E - \mu_H - (\sigma_E x + \sigma_H(1-x))(\sigma_E - \sigma_H)) \\ x\sigma_x &= x(1-x)(\sigma_E - \sigma_H). \end{aligned}$$

Proof. Aggregating over experts gives $dN_t/N_t = \mu_E dt + \sigma_E dZ_t$ and hence

$$\begin{aligned} d(q_tK_t) &= \mu_E N_t dt + \sigma_E N_t dZ_t + \mu_H(q_tK_t - N_t)dt + \sigma_H(q_tK_t - N_t)dZ_t \\ \frac{d(q_tK_t)}{q_tK_t} &= [\mu_E x_t + \mu_H(1-x_t)]dt + [\sigma_E x_t + \sigma_H(1-x_t)]dZ_t. \end{aligned} \tag{62}$$

Note that if $da_t/a_t = \mu_a dt + \sigma_a dZ_t$ and $db_t/b_t = \mu_b dt + \sigma_b dZ_t$, then using Ito's lemma $c_t := a_t/b_t$ satisfies $dc_t/c_t = (\mu_a - \mu_b - \sigma_b(\sigma_a - \sigma_b))dt + (\sigma_a - \sigma_b)dZ_t$. Applying to (62) gives

$$\begin{aligned} \frac{dx_t}{x_t} = (\mu_E - [\mu_E x_t + \mu_H(1-x_t)] - (\sigma_E x_t + \sigma_H(1-x_t))(\sigma_E - [\sigma_E x_t + \sigma_H(1-x_t)])dt \\ + (\sigma_E - [\sigma_E x_t + \sigma_H(1-x_t)])dZ_t \end{aligned}$$

which simplifies as claimed. □

B.3 Individual problems

For an arbitrary $\Delta_y > 0$ we define

$$\begin{aligned} E_1(\Delta_y) &= \frac{e^{(1-\gamma)\Delta_y} - 1}{(1-\gamma)\Delta_y} \\ E_2(\Delta_y) &= \frac{1}{\gamma(1-\gamma)}\left(\frac{1}{\Delta_y^2}[2 - e^{-(1-\gamma)\Delta_y} - e^{(1-\gamma)\Delta_y}] + \frac{1}{\Delta_y}[1 - e^{-(1-\gamma)\Delta_y}]\right) \\ E_c(\Delta_y) &= \left(\frac{1 - e^{-(1-\gamma)\Delta_y}}{(1-\gamma)\Delta_y}\right)^{-1/\gamma}. \end{aligned} \tag{63}$$

Note that $\lim_{\Delta_y \to 0} E_1(\Delta_y) = \lim_{\Delta_y \to 0} E_2(\Delta_y) = \lim_{\Delta_y \to 0} E_c(\Delta_y) = 1$.

Proof of Lemma 12. Using homogeneity, eliminating terms independent of controls, and dividing by $\Delta_t e^{-\rho\Delta_t}V$, the maximization in the discrete Bellman equation becomes

$$\begin{aligned} &\max_{\overline{c},\overline{k} \geq 0} e^{\rho\Delta_t}\frac{\rho\overline{c}^{1-\gamma}}{V(1-\gamma)} + \frac{[e^{-(1-\gamma)\Delta_y} - 1]}{(1-\gamma)\Delta_y}\overline{c} + (\sigma_x x V_{E1D}/V + \sigma_\sigma V_{E2D}/V)\sigma\overline{k} \\ &+ \frac{[e^{(1-\gamma)\Delta_y} - 1]}{(1-\gamma)\Delta_y^2}\left(\sigma^2\overline{k}^2/2 + \Delta_y(\Pi - r)\overline{k}\right) + (1+\Delta_y)\frac{[e^{-(1-\gamma)\Delta_y} - 1]}{(1-\gamma)\Delta_y^2}\sigma^2\overline{k}^2/2. \end{aligned} \tag{64}$$

Eliminating the terms independent of capital gives

$$\frac{1}{V}(\sigma_x x V_{E1D} + \sigma_\sigma V_{E2D})\sigma\overline{k} + \frac{1}{\Delta_y(1-\gamma)}[e^{(1-\gamma)\Delta_y} - 1](\Pi - r)\overline{k}$$
$$+ \frac{1}{\Delta_y^2(1-\gamma)}\Big(e^{(1-\gamma)\Delta_y} - 1 + (1+\Delta_y)[e^{-(1-\gamma)\Delta_y} - 1]\Big)\sigma^2\overline{k}^2/2.$$

The first-order condition is then

$$0 = \frac{1}{V}(\sigma_x x V_{E1D} + \sigma_\sigma V_{E2D})\sigma + \frac{1}{\Delta_y(1-\gamma)}[e^{(1-\gamma)\Delta_y} - 1](\Pi - r)$$
$$- \frac{1}{\gamma(1-\gamma)}\left(\frac{1}{\Delta_y}[1 - e^{-(1-\gamma)\Delta_y}] - \frac{1}{\Delta_y^2}[e^{(1-\gamma)\Delta_y} - 2 + e^{-(1-\gamma)\Delta_y}]\right)\sigma^2\gamma\overline{k}$$

which may be written $0 = (\sigma_x x V_{E1D} + \sigma_\sigma V_{E2D})\sigma/V + E_1(\Delta_y)(\Pi - r) - E_2(\Delta_y)\sigma^2\gamma\overline{k}$, where E_1, E_2 and E_c are given by (63), from which rearrangement gives the result. □

Lemma 17. *For any $\overline{c}, \overline{k}$ we have $\lim_{\Delta_t,\Delta_y\to 0} T(\overline{c},\overline{k};\Delta_t,\Delta_y) = \overline{T}(\overline{c},\overline{k})$ where*

$$\overline{T}(\overline{c},\overline{k})V = -\Big(\rho - (1-\gamma)(r - \overline{c} + (\Pi - r)\overline{k} - \gamma\sigma^2\overline{k}^2/2)\Big)\frac{V}{1-\gamma}$$
$$+ \big([\mu_x x]^+ + [\sigma_x x(1-\gamma)\sigma\overline{k}]^+\big)\frac{V_{E1F}}{1-\gamma} + \big([\mu_x x]^- + [\sigma_x x(1-\gamma)\sigma\overline{k}]^-\big)\frac{[-V_{E1B}]}{1-\gamma}$$
$$+ \big([\theta(\overline{\sigma} - \sigma)]^+ + [\sigma_\sigma(1-\gamma)\sigma\overline{k}]^+\big)\frac{V_{E2F}}{1-\gamma} + \big([\theta(\overline{\sigma} - \sigma)]^- + [\sigma_\sigma(1-\gamma)\sigma\overline{k}]^-\big)\frac{[-V_{E2B}]}{1-\gamma}$$
$$+ \frac{(1-|z|)}{\overline{\Delta}_t(1-\gamma)}(V(x + m_{12}\Delta_x, \sigma + m_{13}\Delta_\sigma) + V(x - m_{12}\Delta_x, \sigma - m_{13}\Delta_\sigma) - 2V)$$
$$+ \frac{|z|}{\overline{\Delta}_t(1-\gamma)}(V(x + m_{22}\Delta_x, \sigma + m_{23}\Delta_\sigma) + V(x - m_{22}\Delta_x, \sigma - m_{23}\Delta_\sigma) - 2V)$$

where recall $\overline{\Delta}_t$ is given by (39).

Proof. Using $\Delta_t = \overline{p}\overline{\Delta}_t$ and the transition probabilities defined after (41), dividing all terms in the Bellman equation by Δ_t and rearranging gives

$$0 = \frac{1}{\Delta_t}(e^{-\rho\Delta_t}[1 - 2\overline{p}] - 1)\frac{V}{1-\gamma} + \frac{\rho\overline{c}^{1-\gamma}}{1-\gamma} + e^{-\rho\Delta_t}\frac{1}{\Delta_y}[e^{-(1-\gamma)\Delta_y} - 1]\frac{\overline{c}V}{1-\gamma}$$
$$+ \frac{e^{-\rho\Delta_t}}{\Delta_y^2}\Big[\Big(\sigma^2\overline{k}^2/2 + \Delta_y(\Pi - r)\overline{k}\Big)[e^{(1-\gamma)\Delta_y} - 1] + (1+\Delta_y)[e^{-(1-\gamma)\Delta_y} - 1]\sigma^2\overline{k}^2/2\Big]\frac{V}{1-\gamma}$$
$$+ e^{-\rho\Delta_t}(\sigma_x x V_{E1D} + \sigma_\sigma V_{E2D})\sigma\overline{k} + e^{-\rho\Delta_t}\frac{1}{\Delta_y}[e^{(1-\gamma)\Delta_y} - 1]\frac{rV}{1-\gamma}$$
$$+ \frac{e^{-\rho\Delta_t}}{1-\gamma}\big([\mu_x x]^+ V_{E1F} + [\mu_x x]^-[-V_{E1B}] + [\theta(\overline{\sigma} - \sigma)]^+ V_{E2F} + [\theta(\overline{\sigma} - \sigma)]^-[-V_{E2B}]\big)$$
$$+ e^{-\rho\Delta_t}\frac{(1-|z|)}{\overline{\Delta}_t(1-\gamma)}(V(x + m_{12}\Delta_x, \sigma + m_{13}\Delta_\sigma) + V(x - m_{12}\Delta_x, \sigma - m_{13}\Delta_\sigma))$$
$$+ e^{-\rho\Delta_t}\frac{|z|}{\overline{\Delta}_t(1-\gamma)}(V(x + m_{22}\Delta_x, \sigma + m_{23}\Delta_\sigma) + V(x - m_{22}\Delta_x, \sigma - m_{23}\Delta_\sigma)).$$

Recalling the definitions (63) and simplifying, this becomes

$$
\begin{aligned}
T(\overline{c},\overline{k};\Delta_t,\Delta_y)V &= \frac{1}{\Delta_t}(e^{-\rho\Delta_t}-1)\frac{V}{1-\gamma} \\
&+ e^{-\rho\Delta_t}(1-\gamma)\Big(E_1\Big(r-\overline{c}e^{-(1-\gamma)\Delta_y}+(\Pi-r)\overline{k}\Big)-E_2\gamma\sigma^2\overline{k}^2/2\Big)\frac{V}{1-\gamma} \\
&+ e^{-\rho\Delta_t}\left(\big([\mu_x x]^+ + [\sigma_x x(1-\gamma)\sigma\overline{k}]^+\big)\frac{V_{E1F}}{1-\gamma} + \big([\mu_x x]^- + [\sigma_x x(1-\gamma)\sigma\overline{k}]^-\big)\frac{[-V_{E1B}]}{1-\gamma}\right) \\
&+ e^{-\rho\Delta_t}\left(\big([\theta(\overline{\sigma}-\sigma)]^+ + [\sigma_\sigma(1-\gamma)\sigma\overline{k}]^+\big)\frac{V_{E2F}}{1-\gamma} + \big([\theta(\overline{\sigma}-\sigma)]^- + [\sigma_\sigma(1-\gamma)\sigma\overline{k}]^-\big)\frac{[-V_{E2B}]}{1-\gamma}\right) \\
&+ e^{-\rho\Delta_t}\frac{(1-|z|)}{\overline{\Delta}_t(1-\gamma)}(V(x+m_{12}\Delta_x,\sigma+m_{13}\Delta_\sigma)+V(x-m_{12}\Delta_x,\sigma-m_{13}\Delta_\sigma)-2V) \\
&+ e^{-\rho\Delta_t}\frac{|z|}{\overline{\Delta}_t(1-\gamma)}(V(x+m_{22}\Delta_x,\sigma+m_{23}\Delta_\sigma)+V(x-m_{22}\Delta_x,\sigma-m_{23}\Delta_\sigma)-2V)
\end{aligned}
$$

which gives the result upon simplification. □