CESifo WORKING PAPERS

10600
2023

August 2023

A Probabilistic Solution to High-Dimensional Continuous-Time Macro and Finance Models

Ji Huang

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 10600*

A Probabilistic Solution to High-Dimensional Continuous-Time Macro and Finance Models

Abstract

This paper introduces the probabilistic formulation of continuous-time economic models: forward stochastic differential equations (SDE) govern the dynamics of backward-looking variables, and backward SDEs capture that of forward-looking variables. Deep learning streamlines the search for the probabilistic solution, which is less sensitive to the "curse of dimensionality." The paper proposes a straightforward algorithm and assesses its accuracy by considering a multiple-country model with an explicit solution under symmetric states. Combining with the finite volume method, the algorithm can obtain global dynamics of heterogeneous-agent models with aggregate shocks, in which agents consider the distribution of individual states as a state variable.

JEL-Codes: C630, G210, E440.

Keywords: backward stochastic differential equation, deep reinforcement learning, the curse of dimensionality, heterogeneous-agent continuous-time model, finite volume method.

Ji Huang
The Chinese University of Hong Kong
9/F Esther Lee Building
Shatin / Hong Kong / China
jihuang@cuhk.edu.hk

May 24, 2023
I am incredibly grateful to Han Jiequn for answering numerous questions on their algorithms and BSDEs. My CUHK colleague Vinci Chow provided tremendous technical support on Tensorow. Yu Jinghai offered superb research assistance.

Introduction

The curse of dimensionality is an unavoidable obstacle for anyone who attempts to build and solve a large-scale dynamic economic model. The dimensionality constraint is particularly tight for continuous-time macro or finance models because their equilibrium dynamics are characterized by systems of partial differential equations (PDEs), which is extremely difficult to solve if the number of state variables is more than three.

With the recent success of Machine Learning, applied mathematicians have solved high-dimensional PDEs by searching for their probabilistic solutions with deep reinforcement learning (E, Han and Jentzen, 2017). The theoretical foundation of the probabilistic solutions of PDEs is the development in the theory of nonlinear Backward Stochastic Differential Equations (BSDEs) since the celebrated paper by Pardoux and Peng (1990).

This paper demonstrates that a continuous-time macro or finance model can be formulated as a system of coupled forward-backward stochastic differential equations (FBSDEs). As in E, Han and Jentzen (2017), I construct a reinforcement learning problem, which gives rise to the solution of the FBSDE system. To combine the probabilistic approach with the *finite volume method*, my algorithm can solve for the global solution of heterogeneous-agent models with aggregate shocks, in which individual agents' dynamic optimization takes into account the law of motion for the distribution of individual states (e.g., wealth distribution) rather than its moments.

Since most economists are unfamiliar with BSDEs, I will use an intuitive example to illustrate key features of the probabilistic approach and also compare with the analytic (PDE) approach. Suppose X_t is an uncontrolled stochastic process. In economics, we're interested in a forward-looking process $V(X_t)$ given by the fixed point $V(\cdot)$ of the following equation.

$$V(X_t) = u(X_t)\,\Delta + E\left[V(X_{t+\Delta})\middle|\, X_t\right], \tag{1}$$

in which $u(\cdot)$ could be utility or discounted cash flow per unit of time and Δ denotes the time length of a period. In the introduction, I will focus on the key step of the fixed point searching process, that is, how to compute the value of $V(\cdot)$ at a point x. The calculation involved in this step demonstrates the main difference between the discrete-time and continuous-time settings and that between analytic and probabilistic approaches.

In a discrete-time setting where Δ is fixed, we need the information of the function $V(\cdot)$ over its entire domain and the transition probability $P\left(X_{t+\Delta}\middle|\, X_t = x\right)$ to compute $V(x)$ via the mapping given by equation (1). The computation work becomes increasingly heavier when X_t has higher dimension, which is the so-called curse of dimensionality.

In a continuous-time setting where Δ goes to the limit zero, we can take advantage of the powerful analysis tool, *calculus* or more precisely *stochastic calculus*. Suppose X_t is driven by an Ito process

$$\mathrm{d}X_t = \mu(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t, \tag{2}$$

where X_t is one-dimensional to simplify the notation and W_t is a one-dimension standard

Brownian motion. In this setting, the transition probability is captured by the Ito process. We rearrange equation (1) and take the limit $\Delta \to 0$

$$\begin{aligned} 0 &= u(x) + E\left[\left. \frac{V(X_{t+\Delta}) - V(X_t)}{\Delta} \right| X_t = x \right] \\ &= u(x) + V'(x)\,\mu(x) + \frac{1}{2} V''(x)\,\sigma^2(x) \end{aligned} \tag{3}$$

The second equation above, which utilizes Taylor expansion (namely, Ito's lemma in stochastic calculus), yields a differential equation of function $V(\cdot)$. The differential equation implies we only need the local information of $V(\cdot)$ around x to compute $V(x)$. In particular, with the finite difference method we can transform equation (3) into

$$0 = u(x) + \frac{V(x+k) - V(x-k)}{2k} \mu(x) + \frac{V(x+k) + V(x-k) - 2V(x)}{2k^2} \sigma^2(x), k > 0$$

which shows how to compute $V(x)$ given the values of $V(\cdot)$ at two points $x + k$ and $x - k$. When X_t is high-dimensional, the computation work becomes heavier because even the local space around x contains many directions. The dimensionality problem is particularly troublesome in heterogeneous-agent models where the infinite-dimensional conditional distribution of idiosyncratic states governs the state of an economy.

Next, I discuss the probabilistic approach to solve for $V(x)$. First, I rewrite equation (1) as

$$V(x) = E\left[u(x)\Delta + V(X_{t+\Delta}) \middle| X_t = x \right].$$

Given the Ito process (2), if Δ is sufficiently close to zero we can represent the random variable within the expectation as

$$V(X_{t+\Delta}) + u(x)\Delta = V(x) + z(W_{t+\Delta} - W_t), \tag{4}$$

where z is an unknown constant. The equation (4) is known as the Martingale Representation Theorem in probability theory. Here, we have a pair of unknowns: $V(x)$ and z. To solve for the pair, we simulate two sample paths of $W_{t+\Delta}$: $W^1_{t+\Delta}$ and $W^2_{t+\Delta}$, and formulate a 2-by-2 linear system

$$V\left(x + \mu(x)\Delta + \sigma(x)\left(W^i_{t+\Delta} - W_t\right)\right) + u(x)\Delta = V(x) + z\left(W^i_{t+\Delta} - W_t\right), i = 1, 2.$$

Both the analytic (differential equation) and probabilistic approach require two evaluation points of $V(\cdot)$. Their difference is that the analytic approach involves a single equation and the probabilistic approach solves a system of two equations. Hence, the analytic approach has the advantage in the low-dimensional case, which is why most economists are more familiar with this approach. Nevertheless, when X_t has higher dimensions, the analytic approach demands increasingly more evaluation points, and the probabilistic approach still needs two.

To show the intrinsic connection between the analytic and probabilistic approaches, I rearrange equation (4)

$$V(X_{t+\Delta}) - V(x) + u(x)\Delta = z(W_{t+\Delta} - W_t),$$

and to apply Taylor expansion and drop higher order terms the equation above gives rise to

$$\left[u\left(x\right)+V'\left(x\right)\mu\left(x\right)+\frac{1}{2}V''\left(x\right)\sigma^{2}\left(x\right)\right]\Delta+V'\left(x\right)\sigma\left(x\right)\left(W_{t+\Delta}-W_{t}\right)=z\left(W_{t+\Delta}-W_{t}\right),$$

which is equivalent to two equations: the differential equation (3) and $z = V'(x)\,\sigma(x)$. It is straightforward to see the probabilistic approach has embedded the analytic solution in the limit when $\Delta \to 0$.

To implement the probabilistic approach, I search for the parametric approximations of $V(\cdot)$ and $z(\cdot)$, denoted as $\tilde{V}(\cdot;\Theta)$ and $\tilde{z}(\cdot;\Theta)$, respectively. Equation (4) suggests that the set of parameters should solve the loss minimization problem:

$$\begin{aligned}\min_{\Theta}: \quad & \frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}\left(\tilde{V}(\hat{x}^{i,j};\Theta)+u(x^{i})\Delta-\tilde{V}(x^{i};\Theta)-\tilde{z}(\cdot;\Theta)w^{i,j}\right)^{2}\\ \text{s.t.} \quad & \hat{x}^{i,j}=x^{i}+\mu(x^{i})\Delta+\sigma(x^{i})w^{i,j}\\ & w^{i,j}\text{ is sampled independently from }N(0,\Delta)\\ & x^{i}\text{ is sampled from a prior distribution}\end{aligned}$$

This formulation makes the difficulty of solving the fixed-point problem (1) much less sensitive to the dimensionality of the state variable X_t. One only needs to increase the size of the sample, i.e., N and M, to ensure the accuracy of the probabilistic approach when the dimensionality of X_t increases. Moreover, this formulation can make use of parallel computing as the evaluation of each sample point is independent of others.

There are two major principles of dynamic optimization: the principle of optimality (i.e., dynamic programming) and the Pontryagin-type maximum principle. The dynamic programming principle emphasizes the optimization over the entire state space, which is often implemented via the analytic approach. In contrast, the maximum principle focuses on the optimization along a particular trajectory, which follows the probabilistic approach. A typical heterogeneous-agent model has a low-dimensional individual state variable (e.g., asset holding) and a high-dimensional aggregate state variable (e.g., the distribution of asset holdings) that is uncontrolled. To apply the maximum principle, we only need to trace the dynamics of the state and the (low-dimensional) co-state variables along a trajectory, which is a relatively easy problem to solve. In order to uncover the value functions and policy functions over the entire state space, we can simulate a large number of sample paths and employ techniques of deep learning to approximate these functions. Notice that it adds relatively small computational costs to have more sample trajectories as the paths are independent and parallel computing applies.

To leverage on the probabilistic approach, I employ the finite volume method to approximate the flow of the distribution of all agents' states with a system of finite number of SDEs. The idea is to discretize the state space into a finite number of intervals or cells. The law of motion for a cell's probability follows an SDE given the transition dynamics of agents on the boundaries of the cell. Hence, I can establish a system of FBSDEs by combining with BSDEs that govern dynamics of forward-looking variables. As an example, I solve for the global solution of the model considered by Fernández-Villaverde, Hurtado and Nuno (2022). The key feature of the global solution is that households explicitly treat the wealth distribution as a state variable (up to the discretization of the state space).

Literature. As a powerful tool of analyzing forward-looking phenomena, BSDEs have been tightly connected to economics and finance problems ever since the seminal work Pardoux and Peng (1990). In fact, Duffie and Epstein (1992) developed a result related to BSDE independently when they establish the theoretical work on recursive utility. Along with a mathematician, Larry Epstein later generalized the recursive utility to include ambiguity aversion using the BSDE framework (Chen and Epstein, 2002).

It was not easy to solve BSDEs numerically prior to the era of deep learning, which is probably why economists pay little attention to the probabilistic formulation of forward-looking problems. Recently, an influential paper in applied mathematics has demonstrated the power of deep learning in solving high-dimensional BSDEs (E, Han and Jentzen, 2017). In fact, the idea of formulating the solvability of FBSDEs as an optimal control problem dates back to Ma and Yong (1995). The mere contribution of my paper is to bring the probabilistic numerical approach to economists' attention.

Broadly speaking, this paper is related to recent papers that apply deep learning to solve high-dimensional macro and finance models. Examples in the discrete-time setting include Azinovic, Gaegauf and Scheidegger (2022), Maliar, Maliar and Winant (2021), and Han, Yang and E (2021); in the continuous-time setting, there are Fernández-Villaverde, Hurtado and Nuno (2022), Duarte (2018), Gopalakrishna (2021), and Sauzet (2021). All the continuous-time setting papers follow the analytic approach to solve PDEs with deep learning. There are detailed comparison between the analytic and probabilistic approaches later in the paper.

My paper contributes to the continuous-time heterogeneous-agent literature (e.g., Kaplan, Moll and Violante (2018), Achdou, Han, Lasry, Lions and Moll (2022)). I provide a numerical method that could solve for the global solution of the equilibrium while characterizing the flow of the distribution of heterogeneous agents in the presence of aggregate shocks. The novelty of my numerical method is the combination of the probabilistic approach, deep learning, and the finite volume method, which is new even in the mean-field game literature to the best of my knowledge.

The organization of the paper follows. Section 1 presents the BSDEs characterization of forward-looking variables over time and also discusses the finite volume method. In Section 2, I define the mathematical framework that characterizes the economic dynamics, explain the algorithm in detail, and discuss other related algorithms. Section 3 considers a multiple-country macro-finance model and exemplifies how to establish the FBSDE system and implement the algorithm. And Section 4 focuses on the global solution of a heterogeneous-agent model with aggregate shocks. In Section 5, I employ the Malliavin derivative to capture the propagation of diffusion shocks in a continuous-time model and illustrate its probabilistic numerical scheme.

1 Probabilistic Formulation

The purpose of this section is to formulate forward-looking elements commonly seen in economic models as solutions to BSDEs, e.g., recursive utility, dynamic optimization, and asset pricing. To handle heterogeneous-agent models with aggregate shocks, I approximate the flow of conditional distribution with a finite but large number of forward SDEs. The technique I use is the finite element method, which is widely used in the computational fluid dynamics. To simplify notation, all variables in this section have one dimension. Since the contribution of my paper is not on the technical side, I intentionally omit mathematical terms like measurability, filtration, adaptedness, etc, which would make definitions more rigorous.

1.1 Introduction to BSDE

A (forward) stochastic differential equation is

$$
\begin{aligned}
\mathrm{d}S_t &= b\left(t, S_t\right)\mathrm{d}t + \sigma\left(t, S_t\right)\mathrm{d}W_t \\
&\text{given initial } S_0,
\end{aligned}
$$

where $\{W_t\}_{t\geq 0}$ is a standard Brownian process (see Chapter 5 of Karatzas and Shreve (1998) for rigorous treatment). The integral version is

$$
S_t = S_0 + \int_0^t b(s, S_s)\mathrm{d}s + \int_0^t \sigma(s, S_s)\mathrm{d}W_s.
$$

A BSDE, however, has a terminal condition instead of an initial one. In particular, a BSDE is

$$
\begin{aligned}
-\mathrm{d}Y_t &= h(t, Y_t, Z_t)\mathrm{d}t - Z_t\mathrm{d}W_t \\
Y_T &= \xi,
\end{aligned}
$$

where T is the terminal date, and ξ is a random variable whose realized value depends on information available up to the terminal date (see the textbook by Zhang (2017) for the rigorous treatment on the topic). In an integral form, the BSDE is written as

$$
Y_t = \xi + \int_t^T h(s, Y_s, Z_s)\mathrm{d}s - \int_t^T Z_s\mathrm{d}W_s.
$$

Unlike the solution of the SDE, which is a stochastic process $\{S_t\}_{t\geq 0}$, the solution of a BSDE is a pair of stochastic processes $\{Y_t, Z_t\}_{0\leq t\leq T}$, where the values of Y_t and Z_t depend on the information realized up to time t.

To show the forward-looking nature of the BSDE solution, let us consider the valuation of a European call option with an expiration date of T and strike price of K.[1] The price of the underlying stock is S_t. By the no-arbitrage principle, the option's price is the value of a self-financing portfolio of risk-free debt and the stock, replicating the option's payoff. Let Y_t denote the time-t value of the portfolio. Then, Y_t is supposed to satisfy the terminal condition

[1] This example is taken from Chapter 7 of the textbook Yong and Zhou (1999).

$Y_T = \max\{0, S_T - K\}$. Given the risk-free rate process $\{r_t\}_{t\geq 0}$, the law of motion for Y_t is

$$\mathrm{d}Y_t = \left(r_t Y_t + \left(\frac{b(t,S_t)}{S_t} - r_t\right)\pi_t\right)\mathrm{d}t + \frac{\sigma(t,S_t)}{S_t}\pi_t \mathrm{d}W_t,$$

where π_t is the position in the stock. Let Z_t denote $\frac{\sigma(t,S_t)}{S_t}\pi_t$. Then, the option value and the replicating strategy $\{Y_t, Z_t\}_{0\leq t\leq T}$ is the solution of the BSDE

$$\begin{aligned}\mathrm{d}Y_t &= \left(r_t Y_t + (b(t,S_t) - r_t S_t)\frac{Z_t}{\sigma(t,S_t)}\right)\mathrm{d}t + Z_t \mathrm{d}W_t\\ Y_T &= \max\{0, S_T - K\}.\end{aligned}$$

More precisely, the pair of the option value and the trading strategy $\{Y_t, Z_t\}_{0\leq t\leq T}$ is the solution of a decoupled Forward-Backward SDE because the drift of Y_t depends on S_t, which, in turn, is the solution of a forward SDE. The term "decoupled" refers to the fact that the dynamics of X_t are independent of the solution of the BSDE.

1.2 Recursive Utility

Given a consumption process $\{c_t\}_{t\geq 0}$, Duffie and Epstein (1992) show that an agent's utility at time t can be defined as

$$V_t = E_t\left[\int_t^\infty f(c_s, V_s)\,\mathrm{d}s\right],$$

where f is an aggregator, and $E_t[\cdot]$ denotes the expectation conditional on information available up to time t.[2] If $f(c,v) = u(c) - \beta v$, then V_t is the standard expected discounted utility, i.e.,

$$V_t = E_t\left[\int_t^\infty e^{-\beta(s-t)} u(c_s)\,\mathrm{d}s\right].$$

Note that

$$\int_0^t f(c_s, V_s)\,\mathrm{d}s + V_t = E_t\left[\int_0^\infty f(c_s, V_s)\,\mathrm{d}s\right]$$

is a martingale. By the Martingale Representation Theorem,[3] there exists a stochastic process $(Z_t)_{t\geq 0}$ such that

$$\int_0^t f(c_s, V_s)\,\mathrm{d}s + V_t = V_0 + \int_0^t Z_s \mathrm{d}W_s.$$

Given a time $T > t$,

$$V_t = V_T + \int_t^T f(c_s, V_s)\,\mathrm{d}s - \int_t^T Z_s \mathrm{d}W_s.$$

Taking V_T as given, the above equation yields a BSDE with solution $(V_t, Z_t)_{0\leq t<T}$.

In a Markov equilibrium, the consumption policy $c(\cdot)$ is a function of a state variable X_t,

[2] More generally, the aggregator f could include Z_t as an argument to accommodate ambiguity aversion (Chen and Epstein, 2002).

[3] See Karatzas and Shreve (1998), p. 182.

which follows

$$\begin{aligned} \mathrm{d}X_t &= \mu\left(X_t\right)\mathrm{d}t + \sigma\left(X_t\right)\mathrm{d}W_t \\ X_0 &= x. \end{aligned} \tag{5}$$

Along with the SDE (5), the following BSDE

$$\begin{aligned} -\mathrm{d}V_t &= f\left(c\left(X_t\right), V_t\right)\mathrm{d}t - Z_t\mathrm{d}W_t \\ V_T &= V\left(X_T\right) \end{aligned}$$

gives rise to the fixed point mapping $V(\cdot)$ from X_t to V_t at any time t.

1.3 Stochastic Maximum Principle and the Principle of Optimality

Consider an agent who maximizes the objective function

$$V\left(X_0, X_0^0\right) = E_0\left[\int_0^\infty e^{-\rho s} f\left(X_s, X_s^0, \alpha_s\right)\mathrm{d}s\right]$$

where ρ is the discount factor, the control variable α_t takes values in a convex set, and X_t and X_t^0 are individual and aggregate state variables, respectively. The individual state variable is driven by a controlled process

$$\mathrm{d}X_t = \mu\left(X_t, X_t^0, \alpha_t\right)\mathrm{d}t + \sigma\left(X_t, X_t^0, \alpha_t\right)\mathrm{d}W_t + \sigma^0\left(X_t, X_t^0, \alpha_t\right)\mathrm{d}W_t^0,$$

where W_t and W_t^0 are independent standard Brownian motions. W_t and W_t^0 represent idiosyncratic and aggregate shocks, respectively. The aggregate state variable X_t^0 follows an uncontrolled process from an individual agent's perspective

$$\mathrm{d}X_t^0 = b\left(X_t^0\right)\mathrm{d}t + \Sigma\left(X_t^0\right)\mathrm{d}W_t^0.$$

Given the optimal control process $(\hat{\alpha}_t)_{t\geq 0}$, the principle of optimality indicates that the value function takes a recursive form

$$V\left(X_t, X_t^0\right) = E_t\left[\int_t^T e^{-\rho(s-t)} f\left(X_s, X_s^0, \hat{\alpha}_s\right)\mathrm{d}s + e^{-\rho(T-t)} V\left(X_T, X_T^0\right)\right].$$

One can solve the value function based on the BSDE formulation specified in Section 1.2. To derive the optimal control process, I apply the stochastic maximum principle and define the generalized current-value Hamiltonian

$$H\left(x, x^0, \alpha, y, z, z^0\right) = \mu\left(x, x^0, \alpha\right) y + \sigma\left(x, x^0, \alpha\right) z + \sigma^0\left(x, x^0, \alpha\right) z^0 + f\left(x, x^0, \alpha\right).$$

The optimal control $\hat{\alpha}_t$ satisfies

$$\hat{\alpha}_t = \arg\max_{\alpha} : H\left(X_t, X_t^0, \alpha, Y_t, Z_t, Z_t^0\right),$$

where $\left(Y_t, Z_t, Z_t^0\right)$ follows the BSDE

$$-\mathrm{d}Y_t = \left(\frac{\partial}{\partial x} H\left(X_t, X_t^0, \hat{\alpha}_t, Y_t, Z_t, Z_t^0\right) - \rho Y_t\right) \mathrm{d}t - Z_t \mathrm{d}W_t - Z_t^0 \mathrm{d}W_t^0$$
$$Y_T = \frac{\partial}{\partial x} V\left(X_T, X_T^0\right)$$

supposing the value function is differentiable.

Remarks. There are two approaches in solving dynamic optimization: dynamic programming and Pontryagin's maximum principle. The former yields Hamilton-Jacobi-Bellman (HJB) equations, which involve the partial derivatives with respect to all state variables: individual and systematic. In a heterogeneous-agent model, one has to find the derivative of a value function with respect to an infinite-dimension variable, which poses an extremely difficult challenge. To apply the maximum principle, we only trace dynamics of the costate or adjoint variable Y_t, whose dimensionality equals that of the individual state variable.

Since most dynamic models are time homogeneous, I apply the principle of optimality and impose the terminal condition to Y_T. In fact, $Y_t = \frac{\partial}{\partial x} V(X_t, X_t^0)$ at any time t. See Chapter 5 of Yong and Zhou (1999) for the rigorous analysis of the connection between the dynamic programming approach and the stochastic maximum principle approach.

If the control variable takes values in a noncovex set, one needs to establish a more general Hamiltonian and trace adjoint variables of the second order. For detailed treatment on this, see Chapter 3 of Yong and Zhou (1999). In economics, we often encounter constraints on states such as the wealth of an agent cannot be negative. Nevertheless, there is still lack of general treatment in the stochastic control literature on state constraints. The solution would be on a case-by-case basis.

1.4 Homothetic Preference with a Linear Budget Constraint

In the asset pricing literature, investors are often assumed to have homothetic preferences with linear budget constraints. In this case, the characterization of investors' optimal choices are simpler. Typically, investors have the Epstein-Zin preference

$$V_t = \mathbb{E}_t\left[\int_t^\infty f\left(c_u, V_u\right) \mathrm{d}u\right],$$

$$f\left(c, V\right) = \frac{1}{1-\psi}\left\{\frac{\rho c^{1-\psi}}{\left[\left(1-\gamma\right)V\right]^{(\gamma-\psi)/(1-\gamma)}} - \rho\left(1-\gamma\right)V\right\}$$

with the budget constraint

$$\mathrm{d}N_t = \left(\mu\left(X_t, \alpha_t\right) N_t - c_t\right) \mathrm{d}t + \sigma\left(X_t, \alpha_t\right) N_t \mathrm{d}W_t.$$

The convention is to conjecture that V_t has the functional form

$$V_t = \frac{\left(\xi_t N_t\right)^{1-\gamma}}{1-\gamma},$$

and ξ_t follows a BSDE

$$\mathrm{d}\xi_t = \xi_t \left(\mu_t^\xi \mathrm{d}t + \sigma_t^\xi \mathrm{d}W_t \right)$$

since V_t does so. Next, I will derive μ_t^ξ. Recall that V_t follows

$$-\mathrm{d}V_t = f\left(c\left(X_t\right), V_t\right)\mathrm{d}t - Z_t \mathrm{d}W_t,$$

and to apply Ito's formula to V_t as a function of ξ_t and N_t, we have

$$\begin{aligned} \frac{\mathrm{d}V_t}{\xi_t^{1-\gamma} N_t^{1-\gamma}} &= \left(\mu\left(X_t, \alpha_t\right) - \hat{c}_t\right)\mathrm{d}t + \sigma\left(X_t, \alpha_t\right)\mathrm{d}W_t + \mu_t^\xi \mathrm{d}t + \sigma_t^\xi \mathrm{d}W_t \\ &\quad + \left(-\frac{1}{2}\gamma \left(\sigma_t^\xi\right)^2 + (1-\gamma)\sigma_t^\xi \sigma\left(X_t, \alpha_t\right) - \frac{1}{2}\gamma \sigma^2\left(X_t, \alpha_t\right) \right)\mathrm{d}t, \end{aligned}$$

where $\hat{c}_t$ denotes $\frac{c_t}{N_t}$. Then

$$\begin{aligned} -\mu_t^\xi = \frac{1}{1-\psi} &\left\{ \rho \left(\frac{\hat{c}_t}{\xi_t} \right)^{1-\alpha} - \rho \right\} + \mu\left(X_t, \alpha_t\right) - \hat{c}_t + \sigma_t^\xi \sigma\left(X_t, \alpha_t\right) \\ &- \frac{1}{2}\gamma \left(\sigma_t^\xi\right)^2 + (1-\gamma)\sigma_t^\xi \sigma\left(X_t, \alpha_t\right) - \frac{1}{2}\gamma \sigma^2\left(X_t, \alpha_t\right), \end{aligned}$$

which is exactly the HJB equation when $\hat{c}_t$ and α_t are chosen optimally. Hence, as long as we can solve the BSDE with respect to ξ_t, it is straightforward to find investors' optimal consumption and portfolio choices.

1.5 Asset Pricing

Consider an asset in a Markov equilibrium, in which the state variable X_t follows

$$\frac{\mathrm{d}X_t}{X_t} = \mu\left(X_t\right)\mathrm{d}t + \sigma\left(X_t\right)\mathrm{d}W_t.$$

Let $m\left(X_t\right)$ denote the stochastic discount factor, which follows

$$\frac{\mathrm{d}m_t}{m_t} = \mu^m\left(X_t\right)\mathrm{d}t + \sigma^m\left(X_t\right)\mathrm{d}W_t.$$

Suppose the dividend flow of the asset is a function of the state variable $D(\cdot)$ with a payoff $g\left(X_T\right)$ at the terminal date T. The asset pricing equation implies

$$m\left(X_t\right) p\left(t, X_t\right) = E\left[\int_t^T m\left(X_u\right) D\left(X_u\right)\mathrm{d}u + m\left(X_T\right) g\left(X_T\right) | X_t \right] \tag{6}$$

at any $t \leq T$, where $p\left(t, X_t\right)$ denotes the asset price as a function of time t and the state variable X_t.

Next, I will link the asset price p_t and its volatility $p_t\sigma_t^p$ with the solution of a BSDE. Note

$$\begin{aligned} & E\left[\int_0^T m\left(X_u\right) D\left(X_u\right) \mathrm{d}u + m\left(X_T\right) g\left(X_T\right) | X_t\right] \\ = & \int_0^t m\left(X_u\right) D\left(X_u\right) \mathrm{d}u + m\left(X_t\right) p\left(t, X_t\right) \end{aligned} \tag{7}$$

is a martingale, which gives rise to the Euler equation

$$\frac{D\left(X_t\right)}{p\left(t, X_t\right)} + \mu_t^p + \mu^m\left(X_t\right) = -\sigma^m\left(X_t\right)\sigma_t^p, \tag{8}$$

where μ_t^p denotes the drift of p_t. Apply the Martingale Representation Theorem

$$\int_0^t m\left(X_u\right) D\left(X_u\right) \mathrm{d}u + m\left(X_t\right) p\left(t, X_t\right) = m\left(X_0\right) p\left(X_0\right) + \int_0^t m\left(X_u\right) p\left(X_u\right)\left(\sigma_t^m + \sigma_t^p\right) \mathrm{d}W_t.$$

Given the terminal condition $p\left(T, X_T\right) = g\left(X_T\right)$

$$\int_0^T m\left(X_u\right) D\left(X_u\right) \mathrm{d}u + m\left(X_T\right) g\left(X_T\right) = m\left(X_0\right) p\left(X_0\right) + \int_0^T m\left(X_u\right) p\left(X_u\right)\left(\sigma_t^m + \sigma_t^p\right) \mathrm{d}W_t,$$

the difference of the two equations above yields a BSDE.

$$m\left(X_t\right) p\left(t, X_t\right) = m\left(X_T\right) g\left(X_T\right) + \int_t^T m\left(X_u\right) D\left(X_u\right) \mathrm{d}u - \int_t^T m\left(X_u\right) p\left(X_u\right)\left(\sigma_u^m + \sigma_u^p\right) \mathrm{d}W_u.$$

In the differential form,

$$\begin{aligned} -\mathrm{d}\left(m_t p_t\right) &= m_t D_t \mathrm{d}t - m_t p_t\left(\sigma_t^m + \sigma_t^p\right) \mathrm{d}W_t, \\ m_T p_T &= m\left(X_T\right) g\left(X_T\right), \end{aligned} \tag{9}$$

which is a BSDE with the solution pair $\{m_t p_t, m_t p_t\left(\sigma_t^m + \sigma_t^p\right)\}_{t\geq 0}$. And, Equation (6) is the well-known Feynman-Kac formula that gives the solution of the BSDE above.

To apply Ito's formula

$$-\mathrm{d}\left(m_t p_t\right) = -m_t \mathrm{d}p_t - p_t m_t \mu_t^m \mathrm{d}t - p_t m_t \sigma_t^m \mathrm{d}W_t - m_t p_t \sigma_t^m \sigma_t^p \mathrm{d}t.$$

Then,

$$\begin{aligned} -m_t \mathrm{d}p_t - p_t m_t \mu_t^m \mathrm{d}t - p_t m_t \sigma_t^m \mathrm{d}W_t - m_t p_t \sigma_t^m \sigma_t^p \mathrm{d}t &= m_t D_t \mathrm{d}t - m_t p_t\left(\sigma_t^m + \sigma_t^p\right) \mathrm{d}W_t \\ -\frac{\mathrm{d}p_t}{p_t} &= \left(\frac{D_t}{p_t} + \mu_t^m + \sigma_t^m \sigma_t^p\right) \mathrm{d}t - \sigma_t^p \mathrm{d}W_t. \end{aligned}$$

The stochastic process of the asset price and its volatility $\{p_t, \sigma_t^p\}$ is the adapted solution of the following decoupled FBSDE

$$\begin{aligned}
\frac{\mathrm{d}X_t}{X_t} &= \mu\left(X_t\right)\mathrm{d}t + \sigma\left(X_t\right)\mathrm{d}W_t \\
-\frac{\mathrm{d}p_t}{p_t} &= \left(\frac{D_t\left(X_t\right)}{p_t} + \mu^m\left(X_t\right) + \sigma^m\left(X_t\right)\sigma_t^p\right)\mathrm{d}t - \sigma_t^p\mathrm{d}W_t \\
X_0 &= x_0 \\
p_T &= g\left(X_T\right)
\end{aligned} \tag{10}$$

It is decoupled because the law of motion of X_t is independent of $\{p_t, \sigma_t^p\}_{t\geq 0}$, which is not the case for most macro-finance models. In the continuous-time macro-finance context, the shortcut is to postulate the law of motion for the asset price $\{p_t\}_{t\geq 0}$

$$\frac{\mathrm{d}p_t}{p_t} = \mu_t^p\mathrm{d}t + \sigma_t^p\mathrm{d}W_t,$$

and use the Euler equation (8) to derive BSDE (10).

1.6 Kolmogorov Forward Equation and Finite Volume Method

Consider an economy with a continuum of heterogeneous agents. Let X_t denote an agent's individual state variable and $g\left(\cdot,t\right)$ the density of the conditional distribution of X_t at time t. Suppose X_t follows

$$\mathrm{d}X_t = \mu\left(X_t, X_t^0, g\left(\cdot,t\right)\right)\mathrm{d}t + \sigma\left(X_t, X_t^0, g\left(\cdot,t\right)\right)\mathrm{d}W_t + \sigma^0\left(X_t, X_t^0, g\left(\cdot,t\right)\right)\mathrm{d}W_t^0,$$

where W_t and W_t^0 are independent Brownian motions, and the former captures the idiosyncratic risk and the latter represents the aggregate risk. Note the drift and volatility terms above depend on density function $g(\cdot,t)$ instead of its value at a particular point. To highlight the difference, I use notation $g(\cdot,t)$ rather than $g(x,t)$.

The law of motion for $g\left(\cdot,t\right)$ is governed by the stochastic Kolmogorov forward equation (KFE), also known as Fokker-Planck equation[4]

$$\begin{aligned}
\mathrm{d}g\left(x,t\right) = &-\frac{\partial}{\partial x}\left(\mu\left(x, X_t^0, g\left(\cdot,t\right)\right)g\left(x,t\right)\right)\mathrm{d}t - \frac{\partial}{\partial x}\left(\left(\sigma^0\left(x, X_t^0, g\left(\cdot,t\right)\right)\mathrm{d}W_t^0\right)g\left(x,t\right)\right) \\
&+ \frac{1}{2}\frac{\partial^2}{\partial x^2}\left\{\left[\left(\sigma\left(x, X_t^0, g\left(\cdot,t\right)\right)\right)^2 + \left(\sigma^0\left(x, X_t^0, g\left(\cdot,t\right)\right)\right)^2\right]g\left(x,t\right)\right\}\mathrm{d}t.
\end{aligned}$$

The idea of the finite volume method is to discretize the space domain of $g\left(\cdot,t\right)$ into a finite number of intervals, e.g., $\left(x_0,x_1\right),\cdots,\left(x_{N-1},x_N\right)$, and to approximate the distribution $g\left(\cdot,t\right)$ with a finite number of probabilities over intervals, i.e.,

$$G_t^i = \int_{x_{i-1}}^{x_i} g\left(x,t\right)\mathrm{d}t.$$

[4]Readers can find an intuitive derivation of the stochastic KFE on Page 111 in the Volume II of Carmona and Delarue (2018).

Taking integration on both sides of the KFE over an interval (x_{i-1}, x_i),

$$\begin{aligned} \mathrm{d}G_t^i = &-\left(\mu\left(x_i, X_t^0, g\left(\cdot, t\right)\right) g\left(x_i, t\right) - \mu\left(x_{i-1}, X_t^0, g\left(\cdot, t\right)\right) g\left(x_{i-1}, t\right)\right) \mathrm{d}t \\ &-\left(\sigma^0\left(x_i, X_t^0, g\left(\cdot, t\right)\right) g\left(x_i, t\right) - \sigma^0\left(x_{i-1}, X_t^0, g\left(\cdot, t\right)\right) g\left(x_{i-1}, t\right)\right) \mathrm{d}W_t^0 \\ &+\frac{1}{2}\frac{\partial}{\partial x}\left\{\left[\left(\sigma\left(x_i, X_t^0, g\left(\cdot, t\right)\right)\right)^2 + \left(\sigma^0\left(x_i, X_t^0, g\left(\cdot, t\right)\right)\right)^2\right] g\left(x_i, t\right)\right\} \mathrm{d}t \\ &-\frac{1}{2}\frac{\partial}{\partial x}\left\{\left[\left(\sigma\left(x_{i-1}, X_t^0, g\left(\cdot, t\right)\right)\right)^2 + \left(\sigma^0\left(x_{i-1}, X_t^0, g\left(\cdot, t\right)\right)\right)^2\right] g\left(x_{i-1}, t\right)\right\} \mathrm{d}t, \end{aligned}$$

which yields a forward SDE with respect to G_t^i. Therefore, we can approximate the flow of the distribution $g\left(\cdot, t\right)$ with a system of finite forward SDEs, and the approximation is more accurate if the intervals are finer.

Remarks. If the individual state variable has multiple dimension, the format of the space partition is quite flexible, e.g., uneven cells, quadrilateral cells, triangular cells, etc. In the multi-dimensional case, one must resort to the divergence theorem to simplify the right-hand side of the stochastic KFE.

2 Deep Learning-Based Probabilistic Approach

In this section, I will define the FBSDE system that characterizes a model's dynamics and illustrate the algorithm in detail as well as how to apply deep learning.

2.1 Forward-Backward SDE

Consider the following system of infinite horizon coupled forward-backward stochastic differential equations:

$$\begin{aligned} \mathrm{d}X_t &= b(X_t, Y_t, Z_t)\mathrm{d}t + \sigma(X_t, Y_t, Z_t)\mathrm{d}W_t \\ -\mathrm{d}Y_t &= h(X_t, Y_t, Z_t)\mathrm{d}t - Z_t\mathrm{d}W_t \\ X_0 &= x, \end{aligned}$$

where $x \in \mathbb{R}^n$,

$$\begin{aligned} b &: \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^n, \\ \sigma &: \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^{n\times d}, \\ h &: \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^m. \end{aligned}$$

Here, W_t is a standard d-dimensional Brownian motion defined on a complete probability space $\left(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_t \geq 0, P\right)$ satisfying the *usual conditions*.[5] The solution of the FBSDEs $\{X_t, Y_t, Z_t\}_{t\geq 0}$ is required to be $\{\mathscr{F}_t\}_{t\geq 0}-$adapted. The well-posedness of the FBSDE above is beyond the scope of the paper.[6] I assume that the system permits a unique solution for the

[5]See Karatzas and Shreve (1998) for details of technical terms.

[6]See Chapter 8 of Zhang (2017) for a survey of papers on the well-posedness of FBSDEs.

rest of the paper.

Markov Property. Given the infinite horizon setting and the time homogeneity of $b(\cdot), \sigma(\cdot)$, and $h(\cdot)$, it is straightforward to see that $\{X_t, Y_t, Z_t\}_{t\geq 0}$ the solution of the FBSDE system satisfies the Markov property. That is, there exist functions $y(\cdot)$ and $z(\cdot)$ that map X_t to Y_t and Z_t, respectively.

2.2 Algorithm

My algorithm takes advantage of the Markov property of the FBSDE system. The basic idea follows. I first use the feedforward neural network denoted by $\hat{y}(\cdot;\Theta)$ and $\hat{z}(\cdot;\Theta)$ to approximate $y(\cdot)$ and $z(\cdot)$, respectively.[7] Θ denotes the set of parameters that identifies the network. Second, Θ is optimized such that $\{X_t^i, Y_t^i, Z_t^i\}_{t\geq 0}$ sample paths simulated according the FBSDE satisfy the Markov property, that is, $Y_t^i = \hat{y}(X_t^i;\Theta)$.

Fixing an arbitrary date T, the solution of the original infinite horizon FBSDE must satisfy the following finite horizon system

$$
\begin{aligned}
\mathrm{d}X_t &= b(X_t, Y_t, z(X_t))\mathrm{d}t + \sigma(X_t, Y_t, z(X_t))\mathrm{d}W_t, \\
X_0 &= x, \\
-\mathrm{d}Y_t &= h(X_t, Y_t, z(X_t))\mathrm{d}t - z(X_t)\mathrm{d}W_t, \\
Y_T &= y(X_T).
\end{aligned}
$$

I discretize time $[0, T]$ as $0 = t_0 < t_1 < t_2 < \cdots < t_n = T$, simulate M sample paths of Brownian motions $\{W_{i,m}, i = 0, \cdots, n-1\}_{m=1}^M$ as well as random initial states $\{x_{0,m}\}_{m=1}^M$. The path index m is dropped later for brevity. Compute $y_0 = \hat{y}(x_0;\Theta)$. Let $\Delta_i = t_{i+1} - t_i$ and $w_i = W_{i+1} - W_i$. Next, repeat the following procedure from $i = 0$ to $i = n-1$.

1. Compute $z_i = \hat{z}(x_i;\Theta)$;
2. Calculate x_{i+1} and y_{i+1} according to

$$
\begin{aligned}
x_{i+1} &= x_i + b(x_i, y_i, z_i)\Delta + \sigma(x_i, y_i, z_i)w_i, \\
y_{i+1} &= y_i - h(x_i, y_i, z_i)\Delta + z_i w_i;
\end{aligned}
$$

3. Compute $\tilde{y}_{i+1} = \hat{y}(x_{i+1};\Theta)$.

The simulated sample paths yield a loss function

$$
\text{Loss}\Big(\Theta; \{x_{0,m}, W_{i,m}, i = 0, \cdots, n-1\}_{m=1}^M\Big) = \frac{1}{Mn}\sum_{m=1}^{M}\sum_{i=1}^{n}\|y_{i,m} - \tilde{y}_{i,m}\|^2,
$$

where $\|\cdot\|$ denotes the square norm. Popular Machine Learning libraries like PyTorch and TensorFlow in Python have a rich set of algorithms to find the optimal Θ that minimizes the loss function.

[7] See Chapter 6 of Goodfellow et al. (2016) for details of feedforward neural network.

2.3 Connections with Related Papers

The idea of combining the probabilistic approach with deep learning comes from E et al. (2017). In E et al. (2017), the BSDE is formulated to solve a PDE. However, it is more intuitive to explicitly characterize the dynamics of a continuous-time macro or finance model with an FBSDE system. The PDE approach is the mainstream in the literature only because of its advantage in low-dimensional cases (e.g., 2 or 3 dimensions).

Regarding the algorithm, E et al. (2017) solve for y_0 of a BSDE given the initial state x_0 and the terminal condition $g(X_T)$. This approach does not apply to the infinite horizon problem since the hypothetical terminal condition is also part of the solution. My algorithm is closer to the one used in Raissi (2018), which approximates the solution function with the deep neural network. The construction of the loss function follows Scheme 2 proposed by Zhang and Cai (2020) and Algorithm 2 by Ji, Peng, Peng and Zhang (2020), which appears more robust according to my early numerical experiments.

There is a subtle difference between the FBSDE I solve and those treated in E et al. (2017); Raissi (2018); Zhang and Cai (2020): Z_t is not an argument of $\sigma(\cdot)$ in their examples. In other words, the volatility terms of the BSDEs do not enter the volatility terms of forward SDEs in their examples. However, a critical feature of macro-finance models is that the volatility of asset price influences the volatility of the aggregate economy, namely endogenous risk. Another difference between examples in their papers and the one covered in the following section is that I consider a more complicated model, which involves a system of coupled BSDEs rather than a single BSDE, i.e., $m > 1$.

3 Implementation I: Multiple-Country Macro-Finance

I consider a multiple-country model to illustrate that the solution of coupled FBSDEs can characterize the equilibrium dynamics of a continuous-time macro-finance model. Each country is populated by two groups of agents: experts and savers. All agents have the logarithm preference with the time discount factor ρ. Only experts in a country can hold the physical assets residing in the same country, and physical assets cannot move across the border. Experts use the physical assets to produce final goods and transform outputs into new physical assets as capital production. The final goods produced in all countries are homogeneous. The critical financial friction is that experts cannot issue outside equity. The international risk-free bond market is frictionless, and so is the international trade market. All countries are identical in terms of their technologies and agents' preferences. The symmetric setting allows for an analytic solution under symmetric states across different countries, which helps assess the algorithm's performance. I consider a 5-country case as a numerical exercise.

3.1 Model and Equilibrium

Given the logarithm preference, the consumption flow of an agent c_t equals the time discount factor ρ times her wealth level at time t. The law of motion for a saver's wealth W_t^s is

$$\frac{dW_t^s}{W_t^s} = (r_t - \rho)\,\mathrm{d}t,$$

where r_t is the risk-free rate. Without the loss of generality, I consider the optimal decisions of a representative expert in each country. Experts across different countries are equipped with the same linear production technology $y_t = ak_t$, where k_t is the efficiency units of physical assets that an expert manages. Given the capital production function $\Phi(\iota)$ and the depreciation rate δ, the efficiency units of physical assets evolve according to

$$\frac{\mathrm{d}k_t^i}{k_t^i} = \left(\Phi\left(\iota_t^i\right) - \delta\right)\mathrm{d}t + \sigma \mathrm{d}Z_t^i, i = 1, \cdots, J, \text{ where } \Phi(\iota) = \frac{1}{\psi}\ln(\psi\iota + 1)$$

and $\left\{Z_t^1, \cdots, Z_t^J\right\}_{t\geq 0}$ is a standard J-dimensional Brownian motion. Note that Z_t^j and Z_t^i are independent for any $i \neq j$. I conjecture that the price of assets in country i follows

$$\frac{\mathrm{d}q_t^i}{q_t^i} = \mu_t^{q,i}\mathrm{d}t + \sum_{j=1}^{J}\sigma_t^{q,i,j}\mathrm{d}Z_t^j.$$

The foreign shock influences the domestic asset price via the international bond market and trade.

Let W_t^i denote the wealth of the representative expert in country i. The law of motion for W_t^i is

$$\frac{\mathrm{d}W_t^i}{W_t^i} = \left(r_t + \varphi_t^i\left(R_t^i - r_t\right) - \rho\right)\mathrm{d}t + \sum_{j=1}^{J}\underbrace{\varphi_t^i\left(\mathbf{1}\{j=i\}\,\sigma + \sigma_t^{q,i,j}\right)}_{\equiv\sigma_t^{W,i,j}}\mathrm{d}Z_t^j,$$

$$\text{where } R_t^i \equiv \frac{a - \iota_t^i}{q_t^i} + \Phi\left(\iota_t^i\right) - \delta + \mu_t^{q,i} + \sigma\sigma_t^{q,i,i}$$

and φ_t^i is the asset-to-equity ratio of the expert.

Given the aggregate stock of physical assets K_t^i and the aggregate consumption c_t^i, the market clearing condition of final goods is

$$a\sum_{j=1}^{J}K_t^i = \sum_{j=1}^{J}c_t^j + \iota_t^j K_t^j.$$

The market clearing of physical assets in each country implies

$$\varphi_t^i W_t^i = q_t^i K_t^i.$$

I will focus on the Markov equilibrium with $2J-1$ endogenous state variables: $\omega_t^1, \cdots, \omega_t^J$, and $\zeta_t^1, \cdots, \zeta_t^{J-1}$, where ω_t^i is the ratio of the expert's net worth to the value of physical assets in country i, i.e.,

$$\omega_t^i \equiv \frac{W_t^i}{q_t^i K_t^i},$$

and

$$\zeta_t^j \equiv \frac{q_t^j K_t^j}{\sum_{h=1}^J q_t^h K_t^h}$$

is the proportion of country j's physical asset value in the world economy. Given the model setting that only experts can manage physical assets residing in their own countries, experts' leverage choices in equilibrium has a one-to-one mapping with state variables ω_t^i

$$\varphi_t^i = \frac{1}{\omega_t^i}.$$

I defer the discussion of the Euler equations to the next subsection on BSDEs.

Under the symmetric states, i.e., $\omega_t^1 = \cdots = \omega_t^J$ and $\zeta_t^j = 1/J, j = 1, \cdots, J-1$, the final-good market clearing condition gives rise to the analytic solution of q_t^j. Given the consumption rule, I rewrite the market clearing condition as

$$\rho = \sum_{j=1}^J \frac{a - \iota_t^j}{q_t^j} \frac{q_t^j K_t^j}{\sum_{i=1}^J q_t^i K_t^i} = \frac{a - \iota_t^J}{q_t^J} \left(1 - \sum_{i=1}^{J-1} \zeta_t^i\right) + \sum_{i=1}^{J-1} \frac{a - \iota_t^i}{q_t^i} \zeta_t^i.$$

Note that the capital production decision is static for experts, that is, ι_t^i maximizes

$$\Phi\left(\iota_t^i\right) - \frac{\iota_t^i}{q_t^i},$$

which yields

$$\iota_t^i = \frac{q_t^i - 1}{\psi}.$$

Therefore,

$$q_t^j = \frac{a\psi + 1}{\rho\psi + 1}, j = 1, \cdots, J,$$

since $q_t^i = q_t^j, i \neq j$ under symmetric states. The analytic result helps gauge the performance of the probabilistic solution.

3.2 Backward SDEs

Euler equations for experts are

$$R_t^i - r_t = \varphi_t^i \sum_{j=1}^J \left(\mathbf{1}\{j = i\}\sigma + \sigma_t^{q,i,j}\right)^2, i = 1, \cdots, J,$$

which give rise to BSDEs w.r.t. q_t^i, that is,

$$-\mathrm{d}q_t^i = h^i\left(\omega_t^i, q_t^i, \{\sigma_t^{q,1,j}\}_{j=1}^J\right)\mathrm{d}t - \sum_{j=1}^J q^j \sigma_t^{q,1,j}\mathrm{d}Z_t^j, \tag{11}$$

$$h^i(\cdot) \equiv \frac{a\psi+1}{\psi} + \frac{q_t^i}{\psi}\ln\left(q_t^i\right) - q_t^i\left(\frac{1}{\psi}+\delta\right) + \sigma q_t^i \sigma_t^{q,i,i} - \frac{q_t^i}{\omega_t^i}\sum_{j=1}^J\left(\mathbf{1}\{j=i\}\sigma + \sigma_t^{q,i,j}\right)^2 - q_t^i r_t.$$

Note that the risk-free rate is also part of the solution. Although there are no BSDE specifically for $\{r_t\}_{t\geq 0}$, the final-good market clearing condition fixes the risk-free rate.

3.3 Forward SDEs

Given the optimal portfolio and capital production decisions, the law of motion for the physical asset value in country i is

$$\frac{\mathrm{d}q_t^i K_t^i}{q_t^i K_t^i} = \underbrace{\left(-\frac{a\psi+1}{\psi q_t^i} + \frac{1}{\psi} + \frac{1}{\omega_t^i}\sum_{j=1}^J\left(1\{j=i\}\sigma + \sigma_t^{q,i,j}\right)^2 + r_t\right)}_{\equiv \mu_t^{qK,i}}\mathrm{d}t + \sum_{j=1}^J \underbrace{\left(\mathbf{1}\{j=i\}\sigma + \sigma_t^{q,i,j}\right)}_{\equiv \sigma_t^{qK,i,j}}\mathrm{d}Z_t^j.$$

To apply Ito's formula, I derive the laws of motion for $\omega_t^i = \frac{W^i}{q_t^i K_t^i}$

$$\mathrm{d}\omega_t^i = b^i\left(\omega_t^i, q_t^i, \{\sigma_t^{q,1,j}\}_{j=1}^J\right)\mathrm{d}t + \sum_{j=1}^J\left(1-\omega_t^i\right)\left(\mathbf{1}\{j=i\}\sigma + \sigma_t^{q,i,j}\right)\mathrm{d}Z_t^j \tag{12}$$

$$b^i(\cdot) \equiv \left(\frac{a\psi+1}{\psi q_t^i} - \frac{1}{\psi} - \rho\right)\omega_t^i + \left(\frac{1}{\omega_t^i} - 1\right)^2 \omega_t^i \sum_{j=1}^J\left(\mathbf{1}\{j=i\}\sigma + \sigma_t^{q,i,j}\right)^2$$

The key feature of the dynamics of ω_t^i is that the solutions of BSDEs (11) enter both its drift and volatility terms, which is a critical feature of many macro-finance models. Next, I consider the law of motion for $\zeta_t^i, i = 1, \cdots, J-1$

$$\mathrm{d}\zeta_t^i = \zeta_t^i \mu_t^{\zeta,i}\mathrm{d}t + \sum_{l=1}^J \zeta_t^i\left(\mathbf{1}\{l=i\}\sigma + \sigma_t^{q,i,l} - \sigma_t^{H,l}\right)\mathrm{d}Z_t^l, \tag{13}$$

$$\mu_t^{\zeta,i} \equiv \mu_t^{qK,i} - \sum_{j=1}^{J-1}\zeta_t^j \mu_t^{qK,j} - \left(1 - \sum_{j=1}^{J-1}\zeta_t^j\right)\mu_t^{qK,J} - \sum_{l=1}^J \sigma_t^{H,l}\left(\mathbf{1}\{l=i\}\sigma + \sigma^{q,i,l} - \sigma_t^{H,l}\right)$$

$$\sigma_t^{H,l} \equiv \sum_{k=1}^{J-1}\zeta_t^k\left(\mathbf{1}\{l=k\}\sigma + \sigma_t^{q,k,l}\right) + \left(1 - \sum_{k=1}^{J-1}\zeta_t^k\right)\left(\mathbf{1}\{l=J\}\sigma + \sigma_t^{q,J,l}\right)$$

Again, the dynamics of ζ_t^i depend on the solutions of BSDEs (11), i.e., $\left(q_t^i, \sigma_t^{q,i,j}\right)$.

3.4 Numerical Example

The solution of the Markov equilibrium is a set of functions of state variables

$$q^i\left(\omega_t^1,\cdots,\omega_t^J,\zeta_t^1,\cdots,\zeta_t^{J-1}\right);r\left(\omega_t^1,\cdots,\omega_t^J,\zeta_t^1,\cdots,\zeta_t^{J-1}\right);$$
$$\sigma^{q,i,j}\left(\omega_t^1,\cdots,\omega_t^J,\zeta_t^1,\cdots,\zeta_t^{J-1}\right),i=1,\cdots,J,j=1,\cdots,J.$$

Notice that the final-good market clearing condition implies

$$\left(a+\frac{1}{\psi}\right)\left(1-\sum_{i=1}^{J-1}\zeta_t^i\right)=q_t^J\left(\rho+\frac{1}{\psi}-\sum_{i=1}^{J-1}\frac{a\psi+1}{\psi q_t^i}\zeta_t^i\right),$$

which yields a mapping from $\{q^1,\cdots,q_t^{J-1},\zeta_t^1,\cdots,\zeta_t^{J-1}\}$ to q_t^J. Thanks to the mapping and Ito's formula, I can take $q_t^J,\sigma^{q,J,j},j=1,\cdots,J$ away from the set of functions for which I need to solve. Detailed derivations are in Appendix A.

As a numerical experiment, I implement the algorithm illustrated in Section 2.2 to solve for the equilibrium of a 5-country version with parameter values: $a=0.1,\delta=0.05,\sigma=0.023,\psi=5$, and $\rho=0.03$. In the 5-country model, there are 9 endogenous state variables: $\omega_t^i,i=1,\cdots,5$, $\zeta_t^i,i=1,\cdots,4$, and 25 functions:

$$q^i\left(\cdot\right),i=1,\cdots,4;r\left(\cdot\right);$$
$$\sigma^{q,i,j}\left(\cdot\right),i=1,\cdots,4,j=1,\cdots,5.$$

Therefore, the neural network I construct involves nine inputs and 25 outputs. The network has three hidden layers, each with 256 inputs and 256 outputs. The activation function is $f(x)=\sin(x)$. I utilize the TensorFlow library in Python to solve the deep reinforcement learning problem.

The initial values of $\omega_0^i,i=1,\cdots,5$ are randomly selected from [0.2 0.8] according to a uniform distribution; those of $\zeta_0^i,i=1,\cdots,4$ are chosen from [0.03 0.26] in the same manner. $\Delta_i=0.001$, $T=0.2$, and the number of sample paths $M=20,000$. Notice that if I use the finite difference method and pick only four grid points along each dimension, the total number of grid points would be $4^9=262,144$.

Test Sample. I generate a separate 500 test sample paths to assess the performance of the learning outcome. These 500 sample paths are NOT used to train the network in the search for optimal Θ. To plug in the test sample, the value of the loss function indicates the average discrepancy between q_i and $\tilde{q}_i$ is around 5×10^{-5} in absolute value.[8]

Symmetric Cases. Note that all countries are identical regarding their exogenous technologies and agents' preferences. Therefore, when the state variables across different countries are the

[8] q_i is the value updated according to the FBSDE, and $\tilde{q}_i$ is the corresponding network output given the same state variables.

Table 1: Symmetric State

This table presents the equilibrium solution under five symmetric states.

$\omega_t^i = 0.3, \zeta_t^j = 0.2$

	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$
q^i	1.3035	1.3036	1.3039	1.3038	1.3070
$\sigma^{q,i,1}$	0.00093	−0.00025	−0.00027	−0.00027	−0.00013
$\sigma^{q,i,2}$	−0.00017	0.00101	−0.00017	−0.00018	−0.00048
$\sigma^{q,i,3}$	−0.00020	−0.00019	0.00097	−0.00018	−0.00039
$\sigma^{q,i,4}$	−0.00026	−0.00026	−0.00027	0.00086	−0.00007
$\sigma^{q,i,5}$	−0.00014	−0.00015	−0.00015	−0.00013	0.00053

$\omega_t^i = 0.4, \zeta_t^j = 0.2$

	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$
q^i	1.3045	1.3048	1.3046	1.3045	1.3034
$\sigma^{q,i,1}$	0.00079	−0.00021	−0.00023	−0.00023	−0.00013
$\sigma^{q,i,2}$	−0.00014	0.00086	−0.00014	−0.00015	−0.00045
$\sigma^{q,i,3}$	−0.00017	−0.00016	0.00083	−0.00015	−0.00035
$\sigma^{q,i,4}$	−0.00022	−0.00022	−0.00023	0.00073	−0.00006
$\sigma^{q,i,5}$	−0.00013	−0.00014	−0.00013	−0.00012	0.00053

$\omega_t^i = 0.5, \zeta_t^j = 0.2$

	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$
q^i	1.3056	1.3059	1.3053	1.3052	1.2998
$\sigma^{q,i,1}$	0.00066	−0.00018	−0.00019	−0.00018	−0.00014
$\sigma^{q,i,2}$	−0.00011	0.00072	−0.00011	−0.00012	−0.00042
$\sigma^{q,i,3}$	−0.00014	−0.00013	0.00069	−0.00012	−0.00031
$\sigma^{q,i,4}$	−0.00018	−0.00019	−0.00019	0.00060	−0.00005
$\sigma^{q,i,5}$	−0.00011	−0.00012	−0.00012	−0.00010	0.00052

$\omega_t^i = 0.6, \zeta_t^j = 0.2$

	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$
q^i	1.3067	1.3070	1.3060	1.3058	1.2962
$\sigma^{q,i,1}$	0.00053	−0.00014	−0.00015	−0.00014	−0.00014
$\sigma^{q,i,2}$	−0.00008	0.00058	−0.00008	−0.00008	−0.00039
$\sigma^{q,i,3}$	−0.00011	−0.00011	0.00055	−0.00009	−0.00027
$\sigma^{q,i,4}$	−0.00015	−0.00015	−0.00015	0.00047	−0.00005
$\sigma^{q,i,5}$	−0.00009	−0.00010	−0.00010	−0.00008	0.00051

$\omega_t^i = 0.7, \zeta_t^j = 0.2$

	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$
q^i	1.3078	1.3081	1.3067	1.3065	1.2927
$\sigma^{q,i,1}$	0.00039	−0.00010	−0.00010	−0.00010	−0.00015
$\sigma^{q,i,2}$	−0.00005	0.00044	−0.00005	−0.00005	−0.00035
$\sigma^{q,i,3}$	−0.00008	−0.00008	0.00041	−0.00006	−0.00023
$\sigma^{q,i,4}$	−0.00011	−0.00012	−0.00012	0.00035	−0.00004
$\sigma^{q,i,5}$	−0.00007	−0.00008	−0.00008	−0.00006	0.00051

same, i.e.,

$$\omega_t^1 = \omega_t^2 = \omega_t^3 = \omega_t^4 = \omega_t^5,$$
$$\zeta_t^1 = \zeta_t^2 = \zeta_t^3 = \zeta_t^4 = 0.2,$$

the endogenous variables, such as asset prices and their volatility terms, are supposed to be identical. In particular, the asset price under any symmetric cases is 1.3043 across all countries, given the chosen parameter values. At symmetric states, positive domestic shocks increase the net worth of domestic experts and also the price of domestic assets. The market clearing condition implies that the price of foreign assets must decline. Hence, at symmetric states

$$\sigma_t^{q,i,j} > 0 \text{ if } i = j;$$
$$\sigma_t^{q,i,j} < 0 \text{ if } i \neq j.$$

Table 1 displays the solutions of the FBSDE under five symmetric states: $\omega_t^i = 0.3$, $\omega_t^i = 0.4$, $\omega_t^i = 0.5$, $\omega_t^i = 0.6$, and $\omega_t^i = 0.7$. Note that the training process is not guided by any information about the symmetric property of the model. The solution appears fairly reasonable. The accuracy declines when there are fewer data samples to learn during the training process.

The approximation of the volatility terms is challenging, as the symmetric solution is a knife-edge case. Table 1 show that the signs of the domestic risk exposure, $\sigma_t^{q,i,i}, i = 1, \cdots, 5$ are positive. The signs of all foreign shock exposure terms $\sigma_t^{q,i,j}, i \neq j$, are negative, and their magnitudes are close.

Dynamics. Figure 1 displays how endogenous variables of country 1 respond to the variation in ω_t^1 its expert's wealth share while other state variables stay unchanged ($\omega_t^i = 0.45, i = 2, \cdots, 5, \zeta_t^i = 0.2, i = 1, \cdots, 4$). When the expert's net worth increases, the domestic asset price rises, and the endogenous risk due to financial amplification declines. In the economy of five countries, the change in the financial condition of one country has limited influence on the world's risk-free rate. When the country-1 expert's net worth is low, she demands a high risk premium, which comes with a low relatively low risk-free rate. Due to the financial friction, the aggregate risk in country 1

$$\left(\sigma + \sigma^{q,1,1}\right)^2 + \left(\sigma^{q,1,2}\right)^2 + \left(\sigma^{q,1,3}\right)^2 + \left(\sigma^{q,1,4}\right)^2 + \left(\sigma^{q,1,5}\right)^2$$

is decreasing in the wealth share of its productive agent.

3.5 Comparison with the PDE approach for Macro-Finance Models

If one solves the corresponding PDEs of the multiple-country model, the computation of endogenous risk will be expensive. The standard procedure is to employ Ito's Lemma

$$q_t^i \sigma_t^{q,i,j} = \sum_{h=1}^{J} \frac{\partial q_t^i}{\partial \omega_t^h} (1 - \omega_t^h) \left(\mathbf{1}\{h = j\} + \sigma_t^{q,h,j} \right) + \sum_{h=1}^{J-1} \frac{\partial q_t^i}{\partial \zeta_t^h} \zeta_t^h \left(\mathbf{1}\{h = j\} + \sigma_t^{q,h,j} - \sigma_t^{H,j} \right) \quad (14)$$

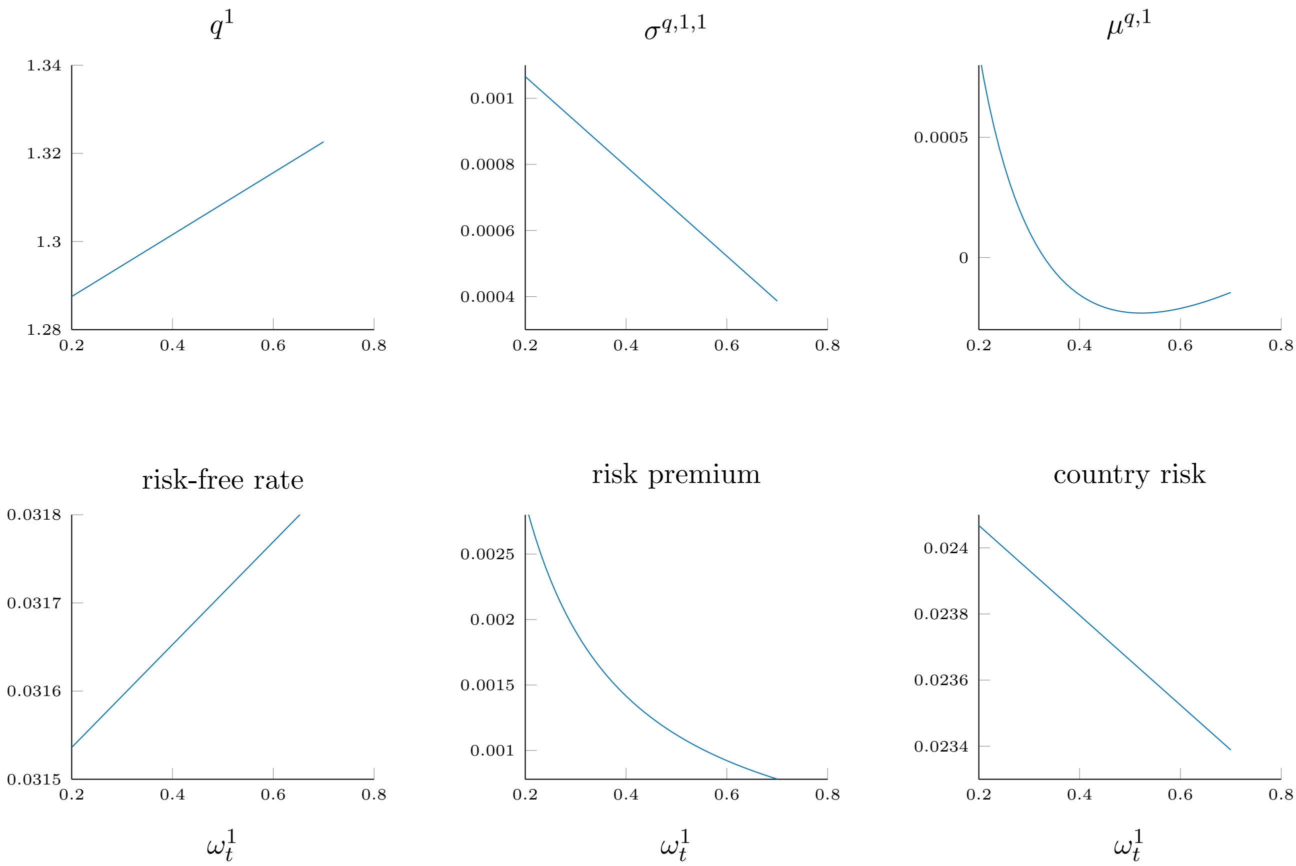

Figure 1: Dynamics
This figure shows asset price, asset price volatility driven by domestic shocks, the drift of asset price, risk-free rate, risk premium, and country-wide risk in country 1 as functions of its expert wealth share when other state variables stay constant ($\omega_t^i = 0.45, i = 2, \cdots, 5, \zeta_t^i = 0.2, i = 1, \cdots, 4$).

to solve for $\sigma_t^{q,i,j}$ where $i = 1, \cdots, J; j = 1, \cdots, J$. Notice that $\sigma_t^{H,j}$ is a linear function of $\sigma_t^{q,i,j}, i = 1, \cdots, J$ (see equation 13). To sum up, there are J systems of linear equations, and each system identified by the j'th shock is composed of J linear equations of the same form as (14).

Solving PDEs directly with deep learning is costly for two reasons. First, given the partial derivatives of $q(\cdot)$, it requires the inverses of J matrices to derive $\sigma_t^{q,i,j}$. Second, it slows down the updating process if the loss function of deep learning requires the differentiation of the network with respect to input variables. Solving for endogenous risk requires the evaluation of $\hat{q}(\cdot;\Theta)$'s Jacobian matrix, and the second order terms of the PDEs themselves involve $\hat{q}(\cdot;\Theta)$'s Hessian matrix. Note that the updating process of deep learning relies on the gradient regarding parameters Θ. With a slight abuse of notation, solving PDE directly with deep learning requires the evaluations of two types of additional terms:

$$\frac{\partial^2 \hat{q}}{\partial \omega^i \partial \Theta} \text{ and } \frac{\partial^3 \hat{q}}{\partial \omega^i \partial \omega^j \partial \Theta}.$$

The probabilistic approach is not subject to the two computational costs. This benefit comes with the cost of larger networks (i.e., a larger scale of Θ) to learn since endogenous risks terms $\sigma_t^{q,i,j}$ are also outputs of the neural network. Nevertheless, the strength of deep learning is to handle large-scale problems with enormous amounts of parameters.

The two types of computational burdens are simplified in examples considered by papers that solve high-dimensional PDEs directly with deep learning, e.g., Duarte (2018), Gopalakrishna (2021), and Sauzet (2021). Their high-dimensional models typically have either low-dimensional or no endogenous state variables, which reduces the dependence of the right-hand side of equation (14) on endogenous risks $\sigma_t^{q,i,j}$. In their multiple-asset models (i.e., Lucas Orchard), the absence of endogenous risk implies that no inverse of a matrix is needed, and they solve decoupled PDEs where the solution of one PDE does not depend on other PDEs' solutions. Although the example in Gopalakrishna (2021) contains three coupled PDEs, there is no need to calculate the inverse matrix since the model has only one risky asset. The numerical example in my paper gives rise to five coupled PDEs ($q^i(\cdot), i = 1, \cdots, 5$) with 9 endogenous variables and an algebraic equation (the final-good market clearing condition). Table 2 compares the differences between the example in this paper and those solving PDEs with deep learning.

Table 2: PDEs and State Variables

Models	PDEs		state variables	
	coupled	decoupled	endogenous	exogenous
this paper	5		9	
growth model in Duarte (2018)		1	1	10
Lucas Orchard[1]		10		9
"222" model in Sauzet (2021)	2		1	1
the example in Gopalakrishna (2021)	3		1	3

[1] Both Duarte (2018) and Sauzet (2021) consider the 10-dimensional Lucas Orchard.

Smoothness. Applying the PDE approach requires that the solution $q(\cdot)$ is twice differentiable everywhere in its domain. If the solution $q(\cdot)$ has kinky points, the auto-differentiation technique that deep learning relies on seems to generate numerical errors to a certain degree. The non-differentiable points are not rare in macro-finance models due to the effects of asset fire-sales, e.g., the logarithmic preference version of Brunnermeier and Sannikov (2014). To my knowledge, there is no conclusive result in the applied mathematics literature about whether solving PDE directly with deep learning would yield viscosity solutions, which are less demanding for smoothness. Economists, however, do not have to bother with the differentiability condition because FBSDE systems do not require their solutions be differentiable.

4 Implementation II: A Continuum of Heterogeneous Agents

In this section, I combine the deep learning-based probabilistic algorithm and the finite volume method to solve the model considered by Fernández-Villaverde et al. (2022), in which I allow all agents to take the conditional distribution of individual characteristics as a state variable. The economy is composed of a representative expert and a continuum of heterogeneous households. Only the expert can hold and produce physical asset without technological illiquidity. Hence, the price of physical asset is always one. Households only accept the risk-free debt issued by the expert or each other.

4.1 Model

Firm. A representative firm has the production function $Y_t = K_t^{\alpha} L_t^{1-\alpha}$, where L_t denotes the effeciency units of labor from households and K_t the physical asset from the expert. Labor wage is

$$w_t = (1-\alpha)\frac{Y_t}{L_t} = (1-\alpha)\left(\frac{K_t}{L_t}\right)^{\alpha},$$

and the rent of physical assets is

$$r_t^k = \alpha\frac{Y_t}{K_t} = \alpha\left(\frac{K_t}{L_t}\right)^{\alpha-1}.$$

The physical asset follows

$$\frac{\mathrm{d}K_t}{K_t} = (\iota_t - \delta)\,\mathrm{d}t + \sigma \mathrm{d}W_t,$$

where ι_t is the investment rate per unit of physical asset, δ the depreciation rate, and σ the volatility of the Brownian shocks to the efficiency units of physical asset. The asset quality shock is the only aggregate shock of the model. The investment is passively determined by the market clearing condition of the final goods

$$Y_t = C_t^e + C_t^h + \iota_t K_t,$$

where C_t^e and C_t^h are the consumption of the expert and the household sector, respectively.

Expert. The expert has the logarithmic preference with time discount factor ρ. Given the risk-free rate r_t, its law of motion for the expert's net worth N_t is

$$\frac{\mathrm{d}N_t}{N_t} = \left(r_t + x_t\left(r_t^k - \delta - r_t\right) - \frac{c_t}{N_t}\right)\mathrm{d}t + \sigma x_t \mathrm{d}W_t,$$

where x_t is the physical asset to net worth ratio. Hence, $(x_t - 1)N_t$ is the expert's demand for risk-free debt. Optimality conditions are

$$c_t = \rho N_t$$
$$r_t^k - \delta - r_t = x_t \sigma^2$$

Household. A household's expected life-time discounted utility is

$$E_0\left[\int_0^\infty e^{-\rho t}\frac{c_t^{1-\gamma}-1}{1-\gamma}\mathrm{d}t\right].$$

Households supply labor inelastically and the efficiency units of labor is captured by a two-state Markov process: $z_t \in \{z_1, z_2\}$, where $0 < z_1 < z_2$. The transition intensity from state 1 to state 2 is λ_1, and that from state 2 to state 1 is λ_2. The wealth of a household follows

$$\mathrm{d}a_t = \underbrace{(w_t z_t + r_t a_t - c_t)}_{\equiv s(a_t, z_t)}\mathrm{d}t$$

with the borrowing constraint $a_t \geq 0$.

To characterize households' consumption choices, I trace dynamics of their expected life-time utility denoted as V_t^i and the costate variable denoted as Y_t^i, in which superscript $i = 1$, or 2 indicates the status of a household's labor efficiency. V_t^i follows the BSDE

$$\mathrm{d}V_t^i = -\left(\frac{c_t^{1-\gamma}-1}{1-\gamma} - \rho V_t^i\right)\mathrm{d}t + U_t^{v,i}\mathrm{d}\Lambda_t^i + Z_t^{v,i}\mathrm{d}W_t,$$

where Λ_t^i is the Poisson shock at state i. Note that in the presence of the jump risk, the solution of the BSDE is $\left(V_t^i, U_t^{v,i}, Z_t^{v,i}, i = 1, 2\right)$. And, Y_t^i follows the BSDE

$$\mathrm{d}Y_t^i = -\left(r_t Y_t^i - \rho Y_t^i\right)\mathrm{d}t + U_t^{y,i}\mathrm{d}\Lambda_t^i + Z_t^{y,i}\mathrm{d}W_t$$

with the solution $\left(Y_t^i, U_t^{y,i}, Z_t^{y,i}, i = 1, 2\right)$. The Hamiltonian is

$$H^i(a, c, y) = \frac{c^{1-\gamma}-1}{1-\gamma} + (w_t z_i + r_t a - c)\, y$$

when the state constraint is not binding. The optimal consumption c_t^i of a household in state i satisfies

$$\left(c_t^i\right)^{-\gamma} = Y_t^i.$$

The key condition my algorithm will employ is the relationship between the costate variable Y_t^i

and the value function $V^i(a_t, \cdot)$, i.e.,

$$Y_t^i = \frac{\partial V^i}{\partial a}(a_t, \cdot). \tag{15}$$

The borrowing constraint implies that if $a_t = 0$,

$$w_t z^i - c_t^i \geq 0, i = 1, 2.$$

In addition, Proposition 1 in Achdou, Han, Lasry, Lions and Moll (2022) shows that $w_t z^1 = c_t^1$.

Market Clearing Conditions and State Variables. Assume that the conditional distribution of households' labor efficiency is always over time, i.e.,

$$P(z_t = z_1) = \frac{\lambda_2}{\lambda_1 + \lambda_2}, \text{ and } P(z_t = z_2) = \frac{\lambda_1}{\lambda_1 + \lambda_2};$$

and the aggregate labor supply is normalized to one

$$z_1 P(z_t = z_1) + z_2 P(z_t = z_2) = 1 = L_t.$$

Let $G(t, a, z)$ denote the conditional distribution of (a, z) over households at time t, where

$$\int_0^\infty \mathrm{d}G(t, a, z_1) = \frac{\lambda_2}{\lambda_1 + \lambda_2} \text{ and } \int_0^\infty \mathrm{d}G(t, a, z_2) = \frac{\lambda_1}{\lambda_1 + \lambda_2}.$$

The risk-free debt market clearing condition is

$$(x_t - 1) N_t = \sum_{i=1}^{2} \int a \mathrm{d}G(t, a, z_i).$$

The physical asset market clears when

$$x_t N_t = K_t = N_t + \sum_{i=1}^{2} \int_0^\infty a \mathrm{d}G(t, a, z_i).$$

The aggregate state variables are N_t and the flow of the conditional density $g(t, a, z)$. The stochastic KFE is

$$\frac{\partial g(t, a, z_i)}{\partial t} = -\frac{\partial}{\partial a}\left((w_t z_i + r_t a_t - c_t) g(t, a, z_i)\right) - \lambda_i g(t, a, z_i) + \lambda_j g(t, a, z_j), i \neq j,$$

where w_t and r_t are stochastic due to the aggregate Brownian shock that hits N_t directly. The law of motion for N_t is

$$\frac{\mathrm{d}N_t}{N_t} = \left(r_t + x_t\left(r_t^k - \delta - r_t\right) - \rho\right)\mathrm{d}t + \sigma K_t \mathrm{d}W_t.$$

4.2 Finite Volume Method

Since it is rarely used in the economics literature, the implementation of the finite volume method deserves detailed exposition.[9] I divide the individual state space into N intervals or cells according to $N+1$ points: $0 = x_0 < x_1 < x_2 < \cdots < x_N$ with the assumption that $G(t, x_N, z_i) = 1, i = 1, 2$. The literature shows that $G(t, \cdot, z_i)$ has a mass point at x_0 conditional on state 1 (Achdou et al., 2022; Fernández-Villaverde et al., 2022). Interval n is $(x_{n-1}, x_n]$. I approximate $G(t, a, z_1)$ with $(N+1)$ scalars, $G(t, a, z_2)$ with N scalars

$$G^n(t, z_i) = \int_{x_{n-1}}^{x_n} \mathrm{d}G(t, a, z_i) = \int_{x_{n-1}}^{x_n} g(t, a, z_i)\, \mathrm{d}a.$$

Given the KFE, the law of motion for $G^n(t, z_i)$ with $0 < n < N$ follows

$$\frac{\partial G^n(t, z_i)}{\partial t} = -\left(s(x_n, z_i)\, g(t, x_n, z_i) - s(x_{n-1}, z_i)\, g(t, x_{n-1}, z_i)\right) - \lambda_i G^n(t, z_i) + \lambda_j G^n(t, z_j).$$

Note that as the wage w_t and the risk-free rate r_t depends on the aggregate state N_t, the KFE above is stochastic.

Next, I discuss how to update $G^n(t, z_i)$ over time for each cases. Assuming that the probability density within an interval is constant and apply the upwind scheme.

$$\begin{aligned}
&\frac{G^n(t+\Delta, z_i) - G^n(t, z_i)}{\Delta} \\
&= -\left[s(x_n, z_i)\left(\frac{G^{n+1}(t, z_i)}{x_{n+1} - x_n}\mathbf{1}\{s(x_n, z_i) < 0\} + \frac{G^n(t, z_i)}{x_n - x_{n-1}}\mathbf{1}\{s(x_n, z_i) \geq 0\}\right)\right] \\
&\quad + \left[s(x_{n-1}, z_i)\left(\frac{G^n(t, z_i)}{x_n - x_{n-1}}\mathbf{1}\{s(x_{n-1}, z_i) < 0\} + \frac{G^{n-1}(t, z_i)}{x_{n-1} - x_{n-2}}\mathbf{1}\{s(x_{n-1}, z_i) \geq 0\}\right)\right] \\
&\quad - \lambda_i G^n(t, z_i) + \lambda_j G^n(t, z_j)
\end{aligned}$$

for $1 < n < N$. For $n = 0$, recall that $G(t, \cdot, z_1)$ has a mass point at $x = x_0$ and

$$\begin{aligned}
&\frac{G^0(t+\Delta, z_1) - G^0(t, z_1)}{\Delta} \\
&= -s(x_0, z_1)\left(\frac{G^1(t, z_1)}{x_1 - x_0}\mathbf{1}\{s(x_0, z_1) < 0\} + G^0(t, z_1)\,\mathbf{1}\{s(x_0, z_1) \geq 0\}\right) - \lambda_1 G^0(t, z_1).
\end{aligned}$$

For $n = 1$,

$$\begin{aligned}
&\frac{G^1(t+\Delta, z_1) - G^1(t, z_1)}{\Delta} \\
&= -\left[s(x_1, z_1)\left(\frac{G^2(t, z_1)}{x_2 - x_1}\mathbf{1}\{s(x_1, z_1) < 0\} + \frac{G^1(t, z_1)}{x_1 - x_0}\mathbf{1}\{s(x_1, z_1) \geq 0\}\right)\right] \\
&\quad + \left[s(x_0, z_1)\left(\frac{G^1(t, z_1)}{x_1 - x_0}\mathbf{1}\{s(x_0, z_1) < 0\} + G^0(t, z_1)\,\mathbf{1}\{s(x_0, z_1) \geq 0\}\right)\right] \\
&\quad - \lambda_1 G^1(t, z_1) + \lambda_2 G^1(t, z_2);
\end{aligned}$$

[9] Another application of FVM in economics is Ahn (2019).

assuming that $g(t, x_0, z_2) = 0$,

$$
\begin{aligned}
&\frac{G^1(t+\Delta, z_2) - G^1(t, z_2)}{\Delta} \\
&= -\left[s(x_1, z_2)\left(\frac{G^2(t, z_2)}{x_2 - x_1}\mathbf{1}\{s(x_1, z_2) < 0\} + \frac{G^1(t, z_2)}{x_1 - x_0}\mathbf{1}\{s(x_1, z_2) \geq 0\}\right)\right] \\
&\quad - \lambda_2 G^1(t, z_2) + \lambda_1\left(G^0(t, z_1) + G^1(t, z_1)\right).
\end{aligned}
$$

For $n = N$,

$$
\begin{aligned}
\int_{x_{N-1}}^{x_N} \frac{\partial g(t, a, z_i)}{\partial t} &= \int_{x_{N-1}}^{x_N} -\frac{\partial}{\partial a}\left((w_t z_i + r a_t - c_t) g(t, a, z_i)\right) \mathrm{d}a \\
&\quad - \lambda_i \int_{x_{N-1}}^{x_N} g(t, a, z_i)\, \mathrm{d}a + \lambda_j \int_{x_{N-1}}^{x_N} g(t, a, z_j)\, \mathrm{d}a \\
\frac{\partial G^N(t, z_i)}{\partial t} &= s(x_{N-1}, z_i) g(t, x_{N-1}, z_i) - \lambda_i G^N(t, z_i) + \lambda_j G^N(t, z_j)
\end{aligned}
$$

Then,

$$
\begin{aligned}
&\frac{G^N(t+\Delta, z_i) - G^N(t, z_i)}{\Delta} \\
&= s(x_{N-1}, z_i)\left(\frac{G^N(t, z_i)}{x_N - x_{N-1}}\mathbf{1}\{s(x_{N-1}, z_i) < 0\} + \frac{G^{N-1}(t, z_i)}{x_{N-1} - x_{N-2}}\mathbf{1}\{s(x_{N-1}, z_i) \geq 0\}\right) \\
&\quad - \lambda_i G^N(t, z_i) + \lambda_j G^N(t, z_j)
\end{aligned}
$$

4.3 Numerical Solution

When solving the model numerically, I use a neural network with five hidden layers and each has 256 nodes. Given $N = 50$, the neural network's inputs are $a_t, G_t^{1,0}, G_t^{1,1}, \cdots, G_t^{1,50}, G_t^{2,1}, \cdots, G_t^{2,50}$, N_t, and outputs are $V_t^i, Y_t^i, U_t^{v,i}, U_t^{y,i}, Z_t^{v,i}, Z_t^{v,i}, i = 1, 2$. I simulate 10,000 economies, and in each economy I initialize 150 households for each initial labor efficiency status, i.e., z_1 or z_2. Along a simulated path, the stock of physical asset is given by

$$
K_t = N_t + \sum_{i=1}^{2}\sum_{j=1}^{N} 0.5(x_{j-1} + x_j) G_t^{i,j},
$$

which, in turn, gives rise to r_t and w_t. Given the distribution G_t and N_t, the network produces $c^{1,j}$ and $c^{2,j}$ for each $x_t^j, j = 0, \cdots, N$. Then,

$$
s_t^{i,j} = w_t z_i + r_t x_j - c_t^{i,j}, i = 1, 2; j = 0, \cdots, N
$$

which are used to update G_t in the next period according to the finite volume method detailed in Section 4.2. The dynamics of each household include its wealth status a_t, labor efficiency z_t, recursive utility V_t, and costate variable Y_t. The later two follow BSDEs. The loss function designed for the reinforcement learning includes the Markov property of the FBSDE solution, equation (15), $w_t z^1 = c_t^1$, $w_t z^2 \geq c_t^2$, and the asymptotic feature of households' consumption

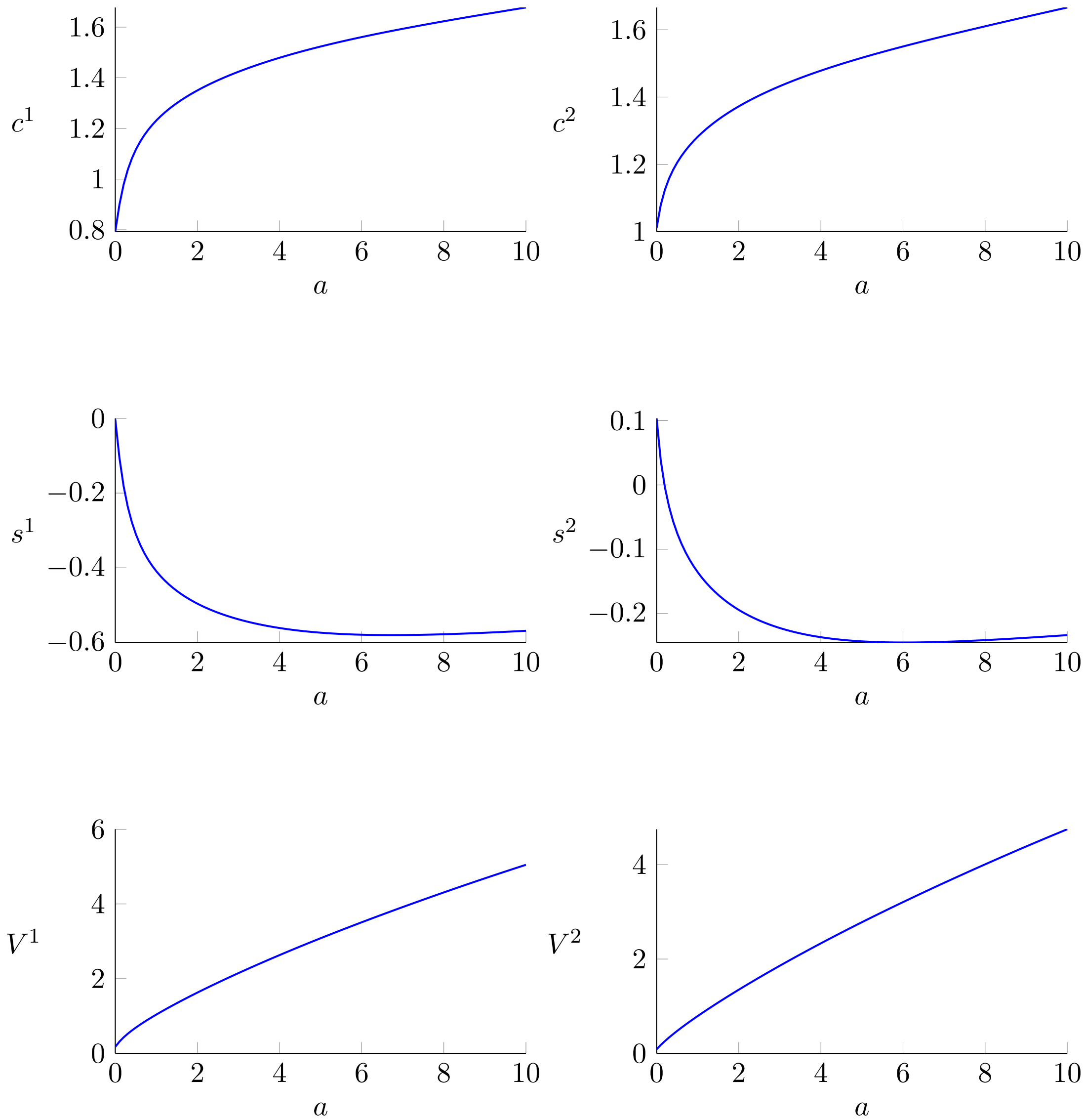

Figure 2: Policy and Value Functions

for large a_t.[10] G_t is initialized based on the Beta distribution with a modification to generate a mass point at x_0 of type-1 households.

Figure 2 displays the policy and value functions of both types of households given an arbitrary G_t and N_t. The policy function displays very high marginal propensity to consumption when the borrowing constraint is close to be binding. When their wealth level is higher, households' consumption tends to be linear in their net worth.

The key idea in Fernández-Villaverde et al. (2022) is to approximate the law of motion for the expert's net worth N_t and the aggregate debt defined as

$$B_t \equiv \sum_{i=1}^{2} \int_0^{\infty} a \mathrm{d}G\left(t, a, z_i\right),$$

and assume that households consider the approximation of the law of motion as the "reality,"

[10]Imposing boundary conditions will lower the chance the algorithm ends at a local optimizer. See Proposition 2 of Achdou et al. (2022) for the asymptotic behavior of households' consumption.

namely, the perceived law of motion. Given the global solution of the model with the distribution G_t as a state variable, I can assess the accuracy of the assumption used in Fernández-Villaverde et al. (2022). My numerical experiment indicates that the consumption policy functions under two different distributions, who yield the same aggregate debt, tend to be very close with the maximum difference at the magnitude of 10^{-4}.

5 Propagation of Diffusion Shocks

In the continuous-time setting with diffusion shocks, a mathematical tool—Malliavin derivative—has to be employed to capture the impacts of diffusion shocks on economic time series (e.g. Borovička and Hansen (2016)). This section will start with the brief introduction of the tool and the highlight of its properties relevant to the characterization of economic dynamics. Then, I will illustrate the computation of Malliavin derivatives that also relies on the FBSDE framework and deep learning.

5.1 Malliavin Derivative and Chain Rule

Recall the d-dimensional Brownian motion $\{W_t\}_{t\geq 0}$ and the probability space $\big(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_t \geq 0, P\big)$ introduced in Section 2. Let $\mathbf{S}$ denote the set of random variables of the form

$$F = f\bigg(\int_0^\infty \langle h_t^1, \mathrm{d}W_t\rangle, \cdots, \int_0^\infty \langle h_t^n, \mathrm{d}W_t\rangle\bigg),$$

where f is an infinitely differentiable function with bounded partial derivatives of any order and $h_t^1, \cdots, h_t^n$ are square-integrable function mapping from $\Omega \times [0,\infty)$ to $\mathbf{R}^d$. The Malliavin derivative of F is defined as the d-dimensional process

$$\mathscr{D}_t F = \sum_{i=1}^n \frac{\partial f}{\partial x_i}\bigg(\int_0^\infty \langle h_t^1, \mathrm{d}W_t\rangle, \cdots, \int_0^\infty \langle h_t^n, \mathrm{d}W_t\rangle\bigg) h_t^i.$$

$\mathbb{D}^{1,2}$ denotes the closure of $\mathbf{S}$ with respect to the norm

$$\|F\|_{1,2} = \left[E|F|^2 + E\left(\int_0^\infty (\mathscr{D}_t F)^2 \mathrm{d}t\right)\right]^{\frac{1}{2}}.$$

It can be shown that the Malliavin derivative has a closed extension to the space $\mathbb{D}^{1,2}$. A relatively accessible reference for Malliavin derivatives is Nualart and Nualart (2018). Intuitively, Malliavin derivative $\mathscr{D}_t F$ captures the impact of the Brownian shock $W_{t+\Delta} - W_t$ on the random variable F when the length of the time interval Δ is arbitrarily small.

Similar to the ordinary derivative, the chain rule also applies to Malliavin derivative. Suppose $\phi : \mathbb{R}^d \to R$ is a continuously differentiable function with bounded partial derivatives, and $F_1, \cdots, F_n \in \mathbb{D}^{1,2}$. Then $\phi(F_1, \cdots, F_n) \in \mathbb{D}^{1,2}$ and

$$\mathscr{D}_t \phi(F_1, \cdots, F_n) = \sum_{i=1}^n \frac{\partial \phi}{\partial x_t}(F_1, \cdots, F_n)\mathscr{D}_t F_i.$$

5.2 Malliavin Derivatives of FBSDE solutions

Following the notation in Section 2, the Malliavin derivative of the equilibrium dynamics $(\mathscr{D}_u^i X_t, \mathscr{D}_u^i Y_t, \mathscr{D}_u^i Z_t, u \le t \le T)$ with respect to the i'th dimension of W_t satisfies a linear FBSDE system

$$
\begin{aligned}
\mathscr{D}_u^i X_t = \mathscr{D}_u^i X_u &+ \int_u^t \frac{\partial b}{\partial x}(X_s, Y_s, Z_s)\mathscr{D}_u^i X_s + \frac{\partial b}{\partial y}(X_s, Y_s, Z_s)\mathscr{D}_u^i Y_s + \frac{\partial b}{\partial z}(X_s, Y_s, Z_s)\mathscr{D}_u^i Z_s \,\mathrm{d}s \\
&+ \int_u^t \frac{\partial \sigma}{\partial x}(X_s, Y_s, Z_s)\mathscr{D}_u^i X_s + \frac{\partial \sigma}{\partial y}(X_s, Y_s, Z_s)\mathscr{D}_u^i Y_s + \frac{\partial \sigma}{\partial z}(X_s, Y_s, Z_s)\mathscr{D}_u^i Z_s \,\mathrm{d}W_s, \\
\mathscr{D}_u^i Y_t = \frac{\partial y}{\partial x}(X_T)\mathscr{D}_u^i X_T &+ \int_t^T \frac{\partial h}{\partial x}(X_s, Y_s, Z_s)\mathscr{D}_u^i X_s + \frac{\partial h}{\partial y}(X_s, Y_s, Z_s)\mathscr{D}_u^i Y_s + \frac{\partial h}{\partial z}(X_s, Y_s, Z_s)\mathscr{D}_u^i Z_s \,\mathrm{d}s \\
&- \int_t^T \mathscr{D}_u^i Z_s \,\mathrm{d}W_s,
\end{aligned}
$$

assuming that $b(\cdot), \sigma(\cdot), h(\cdot)$, and $y(\cdot)$ are continuously differentiable with uniformly bounded derivatives. Moreover,

$$
\left[\mathscr{D}_u^1 X_u, \cdots, \mathscr{D}_u^d X_u\right] = \sigma(X_u, Y_u, Z_u), \tag{16}
$$

$$
\left[\mathscr{D}_u^1 Y_u, \cdots, \mathscr{D}_u^d Y_u\right] = Z_u. \tag{17}
$$

Equation (16) is a general result for forward SDE (see Chapter 7.5 of Nualart and Nualart (2018)); Equation (17) is the result of Proposition 5.9 in El Karoui et al. (1997). The two equations indicate that Malliavin derivatives coincide with the volatility terms regarding the immediate impacts of diffusion shocks. The solution of the linear FBSDE system above $(\mathscr{D}_u X_t, \mathscr{D}_u Y_t, \mathscr{D}_u Z_t, u \le t \le T)$ characterizes the propagation of the Brownian shock $(W_{u+\Delta}^i - W_u^i)$ in the dynamic system from u to T.

Malliavin derivatives are no longer time-homogeneous, and the time t will serve as a state variable. The Malliavin derivatives of equilibrium state variables $\mathscr{D}_u^i X_t$ would also serve as state variables. As the FBSDEs that $(\mathscr{D}_u^i X_t, \mathscr{D}_u^i Y_t, \mathscr{D}_u^i Z_t)$ follow are independent of those $(\mathscr{D}_u^j X_t, \mathscr{D}_u^j Y_t, \mathscr{D}_u^j Z_t)$ follow given that $i \ne j$, the state variables of $(\mathscr{D}_u^i Y_t, \mathscr{D}_u^i Z_t)$ are time t, the equilibrium state variable X_t, and its Malliavin derivative $\mathscr{D}_u^i X_t$.

5.3 Computation

While numerically solving for $(\mathscr{D}_u^j X_t, \mathscr{D}_u^j Y_t, \mathscr{D}_u^j Z_t, u \le t \le T)$, I focus on the mapping denoted by $\Sigma : \left(t, X_t, \mathscr{D}_u^j X_t\right) \to \mathscr{D}_u^j Z_t$, which is approximated by a feedforward neural network denoted as $\hat{\Sigma}(\cdot;\Theta)$. Note that when $t = u$, the initial state X_u is fixed, and the parametric approximation $\hat{\Sigma}$ reduces to $m \times d$ variables $\mathscr{D}_u^j Z_u$.

As in Section 2.2, I discretize time $[u, T]$ as $u = t_0 < t_1 < t_2 < \cdots < t_n = T$ and simulate M sample paths of Brownian motions $\{W_{i,m}, i = 0, \cdots, n-1\}_{m=1}^M$. According to the equilibrium FBSDE, I first generate M sample paths of $\{x_{i,m}, y_{i,m}, z_{i,m}\}$. The sample path index m will be dropped later. Equations (16) and (17) yield $\mathscr{D}_u^j x_0$ and $\mathscr{D}_u^j y_0$, respectively. Next, repeat the following procedure from $i = 0$ to $i = n-1$.

1. compute $\mathscr{D}_u^j z_i$ given the network $\hat{\Sigma}(\cdot;\Theta)$, x_i, and $\mathscr{D}_u^j x_i$;

2. calculate $\mathscr{D}_u^j x_{i+1}$ and $\mathscr{D}_u^j y_{i+1}$ according to

$$\begin{aligned}\mathscr{D}_u^j x_{i+1} &= \mathscr{D}_u^j x_i + \left.\frac{\partial b}{\partial x}\right|_i \mathscr{D}_u^j x_i \Delta_i + \left.\frac{\partial b}{\partial y}\right|_i \mathscr{D}_u^j y_i \Delta_i + \left.\frac{\partial b}{\partial z}\right|_i \mathscr{D}_u^j z_i \Delta_i \\ &\quad + \left.\frac{\partial \sigma}{\partial x}\right|_i \mathscr{D}_u^j x_i w_i + \left.\frac{\partial \sigma}{\partial y}\right|_i \mathscr{D}_u^j y_i w_i + \left.\frac{\partial \sigma}{\partial z}\right|_i \mathscr{D}_u^j z_i w_i \\ \mathscr{D}_u^j y_{i+1} &= \mathscr{D}_u^j y_i - \left.\frac{\partial h}{\partial x}\right|_i \mathscr{D}_u^j x_i \Delta_i - \left.\frac{\partial h}{\partial y}\right|_i \mathscr{D}_u^j y_i \Delta_i - \left.\frac{\partial h}{\partial z}\right|_i \mathscr{D}_u^j z_i \Delta_i + \mathscr{D}_u^j z_i w_i\end{aligned}$$

The loss function is

$$\text{Loss}\left(\Theta; \{W_{i,m}, i = 0, \cdots, n-1\}_{m=1}^{M}\right) = \frac{1}{Mn} \sum_{m=1}^{M} \sum_{i=1}^{N} \|\mathscr{D}_u^j y_i - \frac{\partial y}{\partial x}(x_i) \mathscr{D}_u^j x_i\|^2,$$

which is motivated by the chain rule as well as the terminal condition.

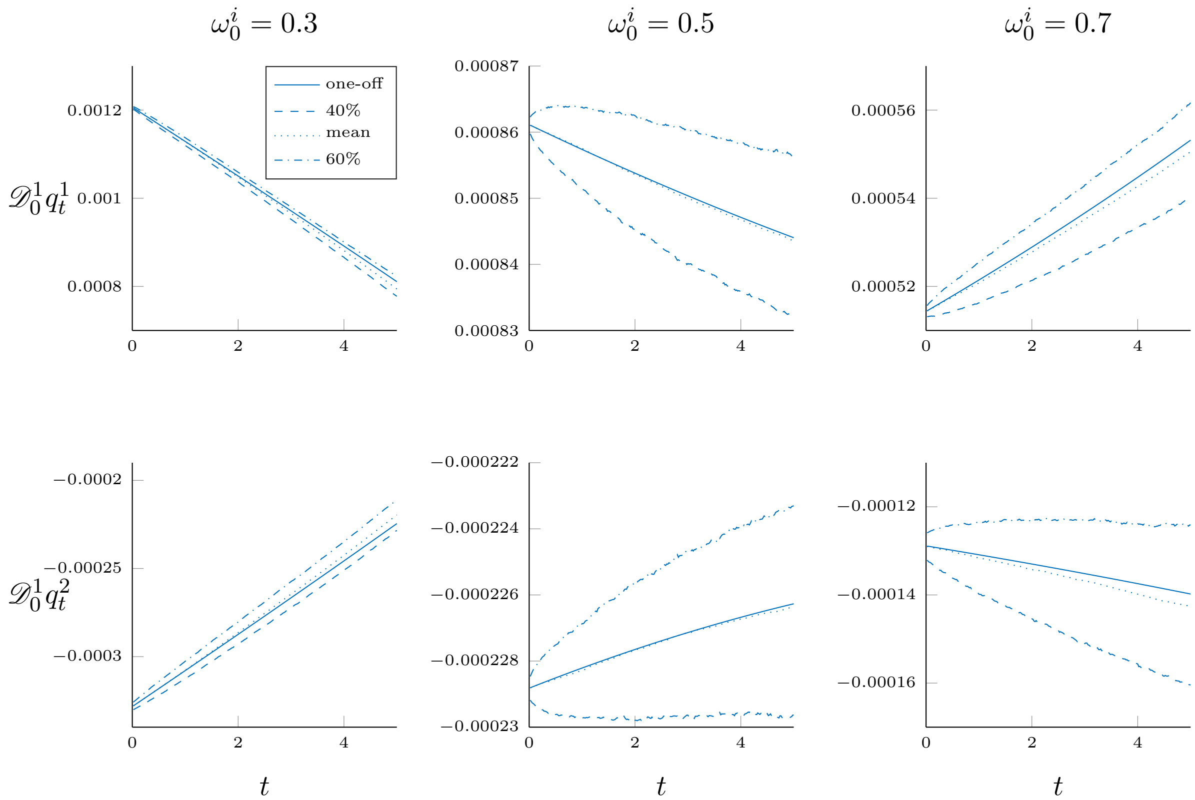

Figure 3: Propagation of Diffusion Shock
This figure displays the impacts of the Brownian shock to country 1 at $t = 0$ on the asset prices in countries 1 and 2 from $t = 0$ to $t = 5$ at three initial states: $(\omega_0^i = 0.3, \zeta_0^j = 0.2)$, $(\omega_0^i = 0.5, \zeta_0^j = 0.2)$, and $(\omega_0^i = 0.7, \zeta_0^j = 0.2)$

Example. I use Malliavin derivatives to characterize the propagation of Brownian shocks in the economy introduced in Section 3. Figure 3 depicts the impacts of country 1's Brownian shock at $t = 0$ on the domestic asset price q_t^1 and country 2's asset price q_t^2 from $t = 0$ to $t = 5$. In the figure, solid curves correspond to the zero-probability scenario that the economy only experiences the one-off shock, i.e., no further shocks between $t = 0$ and $t = 5$; dotted, dashed, dash-dotted lines are the mean, 40 and 60 percentiles of Malliavin derivatives in a $20,000$ simulated sample. Figure 3 shows that the short-term impacts of the Brownian shock are relatively large due to the significant financial amplification effect when the experts' wealth levels are low. Nevertheless, the impacts gradually decay in a relatively long term. If experts have good financial conditions ($\omega_0^i = 0.7$), the Brownian shock's effects are small but still persistent.

Impulse Response. The impulse response function (IRF) characterizes the propagation of an exogenous shock along the sample path of the one-off shock displayed in Figure 3. By contrast, Malliavin derivatives capture the propagation along arbitrary sample paths. It is easier to solve for the impulse response function since the one-off shock sample path bypasses the computation of $\mathscr{D}_0^i Z_t$. Nevertheless, Figure 3 shows that IRF does not capture the expected impacts of the Brownian shock at $t = 0$ on endogenous variables from $t = 0$ to $t = 5$. The numerical example indicates that IRF's deviation from the expected impact increases over time up to 2 percent at time $t = 5$ given initial state ($\omega_0^i = 0.3, \zeta_0^j = 0.2$). This is a natural result for a model with nonlinear dynamics.

6 Final Remarks

This paper shows that the equilibrium of a continuous-time macro or finance model is characterized by a system of coupled infinite-horizon FBSDEs. Unlike the conventional analytic approach that transforms the FBSDEs into PDEs, I highlight that it is straightforward to solve high-dimensional FBSDEs with deep reinforcement learning directly. The idea comes from the applied mathematics literature where the problem of solving high-dimensional PDEs is reformulated as (F)BSDEs, which in turn are solved with reinforcement learning. The algorithm I provide is easy to implement, whose difficulty level is equivalent to simulating economic dynamics given the equilibrium solution. Combining with the finite volume method, my algorithm can solve the global dynamics of heterogeneous-agent model with aggregate shocks. In addition, the paper employs Malliavin derivatives to capture the propagation of Brownian shocks.

References

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll (2022) "Income and wealth distribution in macroeconomics: A continuous-time approach," *The review of economic studies*, Vol. 89, pp. 45–86.

Ahn, SeHyoun (2019) *Computing the distribution: Adaptive finite volume methods for economic models with heterogeneous agents*, No. 10/2019: Working Paper.

Azinovic, Marlon, Luca Gaegauf, and Simon Scheidegger (2022) "Deep equilibrium nets," *International Economic Review*.

Borovička, Jaroslav and Lars Peter Hansen (2016) "Term structure of uncertainty in the macroeconomy," in *Handbook of Macroeconomics*, Vol. 2: Elsevier, pp. 1641–1696.

Brunnermeier, Markus K and Yuliy Sannikov (2014) "A Macroeconomic Model with a Financial Sector," *The American Economic Review*, Vol. 104, pp. 379–421.

Carmona, René and François Delarue (2018) *Probabilistic theory of mean field games with applications I-II*: Springer.

Chen, Zengjing and Larry Epstein (2002) "Ambiguity, risk, and asset returns in continuous time," *Econometrica*, Vol. 70, pp. 1403–1443.

Duarte, Victor (2018) "Machine learning for continuous-time finance,"Technical report, Working paper.

Duffie, Darrell and Larry G Epstein (1992) "Stochastic differential utility," *Econometrica: Journal of the Econometric Society*, pp. 353–394.

E, Weinan, Jiequn Han, and Arnulf Jentzen (2017) "Deep Learning-Based Numerical Methods for High-Dimensional Parabolic Partial Differential Equations and Backward Stochastic Differential Equations," *Communications in Mathematics and Statistics*, Vol. 5, pp. 349–380.

El Karoui, Nicole, Shige Peng, and Marie Claire Quenez (1997) "Backward stochastic differential equations in finance," *Mathematical finance*, Vol. 7, pp. 1–71.

Fernández-Villaverde, Jesús, Samuel Hurtado, and Galo Nuno (2019) "Financial frictions and the wealth distribution,"Technical report, National Bureau of Economic Research.

——— (2022) "Financial frictions and the wealth distribution,"Technical report, National Bureau of Economic Research.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville (2016) *Deep learning*: MIT press.

Gopalakrishna, Goutham (2021) "Aliens and continuous time economies," *Swiss Finance Institute Research Paper*.

Han, Jiequn, Yucheng Yang, and Weinan E (2021) "Deepham: A global solution method for heterogeneous agent models with aggregate shocks," *arXiv preprint arXiv:2112.14377*.

Ji, Shaolin, Shige Peng, Ying Peng, and Xichuan Zhang (2020) "Three algorithms for solving high-dimensional fully coupled FBSDEs through deep learning," *IEEE Intelligent Systems*, Vol. 35, pp. 71–84.

Kaplan, Greg, Benjamin Moll, and Giovanni L Violante (2018) "Monetary policy according to HANK," *American Economic Review*, Vol. 108, pp. 697–743.

Karatzas, Ioannis and Steven Shreve (1998) *Brownian motion and stochastic calculus*, Vol. 113: Springer Science & Business Media.

Ma, Jin and Jiongmin Yong (1995) "Solvability of Forward-Backward SDES and the Nodal Set of Hamilton-Jacobi-Bellman Equations.," in *Chinese Annals of Mathematics, Series B*, Vol. 16, pp. 279–298.

Maliar, Lilia, Serguei Maliar, and Pablo Winant (2021) "Deep learning for solving dynamic economic models.," *Journal of Monetary Economics*, Vol. 122, pp. 76–101.

Nualart, David and Eulalia Nualart (2018) *Introduction to Malliavin calculus*, Vol. 9: Cambridge University Press.

Pardoux, Etienne and Shige Peng (1990) "Adapted solution of a backward stochastic differential equation," *Systems & Control Letters*, Vol. 14, pp. 55–61.

Raissi, Maziar (2018) "Forward-backward stochastic neural networks: Deep learning of high-dimensional partial differential equations," *arXiv preprint arXiv:1804.07010*.

Sauzet, Maxime (2021) "Projection methods via neural networks for continuous-time models," *Available at SSRN*.

Yong, Jiongmin and Xun Yu Zhou (1999) *Stochastic controls: Hamiltonian systems and HJB equations*, Vol. 43: Springer Science & Business Media.

Zhang, Jianfeng (2017) *Backward stochastic differential equations: from linear to fully nonlinear theory*: Springer.

Zhang, Wenzhong and Wei Cai (2020) "FBSDE based neural network algorithms for high-dimensional quasilinear parabolic PDEs," *arXiv preprint arXiv:2012.07924*.

Appendix

A Derivations Related to Market Clearing Condition

I will derive the relationship between q_t^J and $q_t^j, j = 1, \cdots, J-1$, based on the clearing condition of the final-good market.

$$\rho \sum_{j=1}^{J} q_t^j K_t^i = \sum_{j=1}^{J} \left(a - \iota_t^j\right) K_t^j$$

$$\begin{aligned}
\rho &= \sum_{j=1}^{J} \frac{a-\iota_t^j}{q_t^j} \frac{q_t^j K_t^j}{\sum_{i=1}^{J} q_t^i K_t^i} = \frac{a-\iota_t^J}{q_t^J}\left(1-\sum_{i=1}^{J-1}\zeta_t^i\right) + \sum_{i=1}^{J-1}\frac{a-\iota_t^i}{q_t^i}\zeta_t^i \\
\left(a+\frac{1}{\psi}\right)\left(1-\sum_{i=1}^{J-1}\zeta_t^i\right) &= q_t^J\left(\rho+\frac{1}{\psi}\left(1-\sum_{i=1}^{J-1}\zeta_t^i\right)-\sum_{i=1}^{J-1}\left(\frac{a\psi+1}{\psi q_t^i}-\frac{1}{\psi}\right)\zeta_t^i\right) \\
&= q_t^J\left(\rho+\frac{1}{\psi}-\sum_{i=1}^{J-1}\frac{a\psi+1}{\psi q_t^i}\zeta_t^i\right)
\end{aligned}$$

Let

$$k\left(q^1, q^2, \cdots, q^{J-1}, \zeta^1, \zeta^2, \cdots, \zeta^{J-1}\right) \equiv \rho + \frac{1}{\psi} - \sum_{i=1}^{J-1} \frac{a\psi+1}{\psi q^i}\zeta_t^i$$

$$\frac{\partial k}{\partial q^i} = \frac{a\psi + 1}{\psi}\frac{\zeta_t^i}{\left(q^i\right)^2}$$
$$\frac{\partial k}{\partial \zeta^i} = -\frac{a\psi + 1}{\psi q^i}$$

Given

$$q_t^J = k^{-1}\left(1 - \sum_{i=1}^{J-1}\zeta_t^i\right)\left(a + \frac{1}{\psi}\right)$$

$$\frac{\partial q^J}{\partial q^i} = -k^{-2}\left(1 - \sum_{i=1}^{J-1}\zeta_t^i\right)\frac{\partial k}{\partial q^i}\left(a + \frac{1}{\psi}\right)$$
$$\frac{\partial q^J}{\partial \zeta^i} = k^{-2}\left(-k - \left(1 - \sum_{i=1}^{J-1}\zeta_t^i\right)\frac{\partial k}{\partial \zeta^i}\right)\left(a + \frac{1}{\psi}\right)$$

Given $\sigma^{q,i,j}, i = 1, \cdots, J-1$, we can derive $\sigma^{q,J,j}$ by Ito's formula

$$\begin{aligned}
q_t^J \sigma_t^{q,J,j} &= \sum_{h=1}^{J-1}\frac{\partial q^J}{\partial q^h}q_t^h\sigma_t^{q,h,j} + \sum_{h=1}^{J-1}\frac{\partial q^J}{\partial \zeta^h}\zeta_t^h\sigma_t^{\zeta,h,j} \\
&= \sum_{h=1}^{J-1}\frac{\partial q^J}{\partial q^h}q_t^h\sigma_t^{q,h,j} + \sum_{h=1}^{J-1}\frac{\partial q^J}{\partial \zeta^h}\zeta_t^h\sum_{u=1}^{J}a_u^{\zeta,h,j}\left(\mathbf{1}\left\{j = u\right\}\sigma + \sigma_t^{q,u,j}\right) \\
&= \sum_{h=1}^{J-1}\frac{\partial q^J}{\partial q^h}q_t^h\sigma_t^{q,h,j} + \sum_{h=1}^{J-1}\frac{\partial q^J}{\partial \zeta^h}\zeta_t^h a_j^{\zeta,h,j}\sigma + \sum_{h=1}^{J-1}\frac{\partial q^J}{\partial \zeta^h}\zeta_t^h\sum_{u=1}^{J}a_u^{\zeta,h,j}\sigma_t^{q,u,j}
\end{aligned}$$

$$\left(q^J - \sum_{h=1}^{J-1}\frac{\partial q^J}{\partial \zeta^h}\zeta^h a_J^{\zeta,h,j}\right)\sigma_t^{q,J,j} = \sum_{h=1}^{J-1}\frac{\partial q^J}{\partial q^h}q_t^h\sigma_t^{q,h,j} + \sum_{h=1}^{J-1}\frac{\partial q^J}{\partial \zeta^h}\zeta_t^h a_j^{\zeta,h,j}\sigma + \sum_{h=1}^{J-1}\frac{\partial q^J}{\partial \zeta^h}\zeta_t^h\sum_{u=1}^{J-1}a_u^{\zeta,h,j}\sigma_t^{q,u,j}$$