A Selective Survey on Mathematical Programmings in Macroeconomics^{*†}

CAI Zongwu^a and HU Jingxian^{‡b}

^aDepartment of Economics, University of Kansas, Lawrence, KS 66045, USA ^bZhongtai Securities Institute for Financial Studies, Shandong University, Jinan, Shandong 250100, China

February 12, 2025

Abstract. This paper surveys the literature for the macroeconomic optimization problems in both discrete and continuous time models in macroeconomics, and provides an overview over some related computational methods in order to solve the models linearly and nonlinearly, and to compute the transition dynamics and the impulse response functions. Also, the introduction of the financial sectors, the continuous time analysis, and the advanced mathematical tools into the general equilibrium framework expands greatly the scope of the interdisciplinary research to mathematics, statistics and econometrics, and creates further space for exploration and collaboration. Finally, some future research issues related to this topic are highlighted.

Key Words: Diffusion models; Dynamic programming; General equilibrium; Finite Element methods; Nonlinear models;

MR Subject Classification: 35Q91, 62P20, 91B50, 91B51.

^{*}The authors acknowledge the financial supports, in part, from the National Natural Science Foundation of China with grant numbers 72033008 and 72133002.

[†]Contact information: Z. Cai (caiz@ku.edu) and J. Hu (jingxianhu@sdu.edu.cn) [‡]Corresponding author

§1 Introduction

Since the linear programming was introduced from mathematics to economics to help economists decide the allocation of limited resource, both mathematical programming methods and macroeconomics theories have made much more progresses. In the frontier of these progresses, the mathematical theories, the computational methods, and the macroeconomic models promote each other and make the boundaries among the three different subjects melt. This survey provides a selective overview on how the relation between mathematical programming and macroeconomics evolve from the past and about what the latest trend in this interdisciplinary subject is nowadays.

Our review on this relation starts from the dynamic stochastic general equilibrium (DSGE) models in discrete time dated back in 1970's. The well known Bellman equations were applied to solving the optimization problems of the agents within different microeconomic sectors in the macroeconomy. In the mean time in history, the finance theory in continuous time was thriving due to its popularity in asset pricing and the elegant combination of mathematics and portfolio theories. It was until the burst of the financial crisis of 2008 that financial sectors were added to the macroeconomic models to help people better understand how the real sector of the economy and the business cycles are influenced by the financial market's turbulence. One critical tool to construct such a connection is through the continuous time framework. Different from the optimization methods in discrete time, the stochastic calculus, the stochastic control, the numerical methods to solve the stochastic differential equations (SDEs), the ordinary differential equations (ODEs), and the partial differential equations (PDEs) are often needed to solve the macroeconomic optimization problems in continuous time. Therefore, the second half of this survey provide an overview of the latest progresses of the macroeconomic modeling in continuous time, especially the heterogeneous agent models (HAM), as well as the advanced computational methods to solve the models' key equations, the transition dynamics, and the impulse response functions. Thanks to the fast-growing computation power, the solutions that we can get from the nonlinear models are global solutions.

The scope of the questions that macroeconomics can explore has been greatly expanded due to the advances in the mathematical modeling and computational tools, from the macroeconomics models without micro-foundations, to the models that depict the households', firms', and governments' decision-making; from the static analysis to the dynamic analysis; from the deterministic equilibrium to the equilibrium with uncertainty; from the first-order approximation to the solutions locally, to the nonlinear solutions globally. Throughout this survey, we can find the efforts made by generations of scholars that clarify and promote our knowledge on these matters continuously. Even though the current coverings in this interdisciplinary subject can be thought as scattered, we look forward to more future studies coming to fill the gap in the unknown space.

The rest of this paper is organized as follows. Section 2 gives a review on mathematical programming for macroeconomics models with micro-foundations. Section 3 outlines some challenges for the discrete-time nonlinear dynamic stochastic general equilibrium models. Section 4 describes the dynamic programming methods for continuous-time macro-finance models. Finally, Section 5 concludes the paper with some remarks for future research.

§2 Mathematical Programming for Macroeconomics Models with Micro-foundations

The micro-foundation was first introduced in the macroeconomics models after the school of the new classical macroeconomics. The two generations of the new classical macroeconomists are the school of the rational expectation, such as Thomas J. Sargent, Robert E. Lucas Jr., and the school of the real business cycle, such as Edward C. Prescott and Neil Wallace, respectively. There are three key assumptions in the new classical macroeconomics: agents' maximization, rational expectation, and market clearing. The decision problems of the microeconomic sectors of the economy are combined in a general equilibrium model, thus serving as the micro-foundation of the macroeconomic model. Specifically, the determination of the equilibrium output and prices are from the solutions of the optimization problems of the households, the producers, the retailers, and the government. To solve these optimization problems in discrete time, the dynamic programming, such as the Bellman equation, is necessary.

The real business cycle school pointed that it is the real factors that cause the economic fluctuations, rather than the monetary factors. To answer the questions that how the real factors cause the economic fluctuations and what the underlying transmission mechanism is, the real business cycle theory emphasizes on the importance of the supply shock and the intertemporal substitution between the leisure and the work. The latter explains why a small change in the wage can cause large and long-term variations of the output and employment. During the debate of the macroeconomic school of thoughts, the new Keynesian school borrowed the merits of the new classical macroeconomics and developed its own framework to support Keynes's economic thoughts. The representatives of the new Keynesian school macroeconomists include, but not limited to, N. Gregory Mankiw, Lawrence H. Summers, Olivier Blanchard, Julio Rotemberg, Edmund S. Phelps, George A. Akerlof, Janet L. Yellen, Joseph Stiglitz, Ben S. Bernanke, David H. Romer and so on. In contrast to the new classical macroeconomics, the new Keynesian school does not assume that the market can clear once a shock hits the economy, since the supply of labor and products, the wage and the price adjust slowly. The new Keynesian school adopts the agents' optimization and rational expectation from the new classical macroeconomics, while following the Keynesian school to support that the fluctuations of the aggregate demand shift both the output and the price in the short-run. Therefore, the government's policy plays a key role to bring the aggregate demand back to the normal level during the economic recession.

As demonstrated in Gertler (2024), a typical macroeconomic model with micro-foundations is consisted of the sectors of households, firms, monetary, and fiscal authorities. The dynamic programming serves as an important method of solution for the optimization problems. Compared to the dynamic programming applied in microeconomics, the application in macroeconomics resides in a general equilibrium framework, which ultimately synthesizes the individual sectors' decisions into several types of equilibrium in macroeconomics, such as the competitive equilibrium, the stationary equilibrium, and the Markov equilibrium. We will compare these types of equilibrium later in Section 4.3.

2.1 Household's Utility Maximization

Suppose that a representative household chooses $\{C_t, L_t, M_t/P_t, B_{t+1}/P_t, K_{t+1}\}_{t\geq 0}$ to maximize its utility

$$\mathbb{E}_0\left\{\sum_{t=0}^{\infty}\beta^t \left[\frac{1}{1-\gamma}C_t^{1-\gamma} + \frac{a_m}{1-\gamma_m}\left(\frac{M_t}{P_t}\right)^{1-\gamma_m} - \frac{1}{1+\varphi}L_t^{1+\varphi}\right]\right\},\$$

subject to the budget constraint given by

$$C_{t} = \frac{W_{t}}{P_{t}}L_{t} + Z_{t}K_{t} + \Pi_{t} + \mathrm{TR}_{t} - \frac{M_{t} - M_{t-1}}{P_{t}} - \frac{\left(\frac{1}{R_{t}^{n}}\right)B_{t+1} - B_{t}}{P_{t}} - Q_{t}\left(K_{t+1} - K_{t}\right)$$

and assume that there is no Ponzi schemes, where β is the discount factor, which is assumed to satisfy $0 < \beta < 1$, and $\mathbb{E}_0(\cdot)$ denotes the mathematical expectation operator. Note that the solution of the optimization problem yields the control variable K_{t+1} as a policy function of the state variable K_t . In the aforementioned household's optimization problem, C_t is the household's consumption, W_t is the nominal wage, P_t is the price level, L_t is the labor (hours worked), Z_t is the rental cost of capital, K_t is the capital holding, Π_t is the profit from monopoly competitive firms, TR_t is the government transfer, M_t is the household's money holding, B_t is the bond holding, R_t^n is the gross nominal interest rate, $1/R_t^n$ is the price of one period discount bond earning the gross nominal return R_t^n , and Q_t is the price of capital. The parameters γ and γ_m are the coefficients of the relative risk aversion, and φ^{-1} stands for the Frisch elasticity of labor supply. The parameter a_m stands for the weight of real money balance in the utility. For some special cases, γ , γ_m , φ , and a_m might take some specific values. For example, when $\gamma = \gamma_m = \varphi = 1$, the constant relative risk aversion separable utility becomes the log separable utility. For the case of cashless economy, $a_m \to 0$.

2.2 Firms' Profit Maximization and Cost Minimization

There are three types of firms in the general equilibrium model: the final good firms, the intermediate good firms, and the capital producer.

2.2.1 Final Good Firms

Final good firms are competitive producers of a homogeneous good, Y_t , using intermediate goods, $Y_t(f)$. The production function that transforms intermediate goods into final output is given by

$$Y_t = \left[\int_0^1 Y_t(f)^{\frac{\varepsilon-1}{\varepsilon}} df\right]^{\frac{\varepsilon}{\varepsilon-1}}$$

where $\varepsilon > 1$ is the (constant) elasticity of substitution between intermediate goods. Note that this production function also exhibits constant returns to scale and diminishing marginal product for each input $(\varepsilon - 1)/\varepsilon < 1$. Each firm chooses $Y_t(f)$ to minimize costs $\int_0^1 P_t(f)Y_t(f)df$ for a given level of output $Y_t = \left[\int_0^1 Y_t(f)^{\frac{\varepsilon-1}{\varepsilon}}df\right]^{\frac{\varepsilon}{\varepsilon-1}}$ and given $P_t(f)$. The result is the following demand function for each intermediate good f:

$$Y_t(f) = \left[\frac{P_t(f)}{P_t}\right]^{-\varepsilon} Y_t.$$

Combining with the production function yields the following nominal price index for the final good:

$$P_t = \left[\int_0^1 P_t(f)^{1-\varepsilon} df\right]^{\frac{1}{1-\varepsilon}}$$

2.2.2 Intermediate Good Firms

There is a continuum of intermediate good firms, indexed by $f \in [0, 1]$. Each produces a differentiated good and is a monopolistic competitor. Each firm uses both labor $L_t(f)$ and capital $K_t(f)$ to produce output according to

$$Y_t(f) = A_t K_t(f)^{\alpha} L_t(f)^{1-\alpha}$$

where A_t is the level of productivity and $0 < \alpha < 1$ is the capital share. Firm f chooses inputs $K_t(f)$ and $L_t(f)$ to minimize the total cost given by

$$\frac{W_t}{P_t}L_t(f) + Z_tK_t(f)$$

subject to the output demand

$$A_i K_t(f)^{\alpha} L_t(f)^{1-\alpha} = \bar{Y}$$

where \bar{Y} is a given output level. The first order conditions for this optimization problem are as follows:

$$\frac{W_t/P_t}{(1-\alpha)Y_t(f)/L_t(f)} = \mathrm{MC}_t(f) \quad \text{and} \quad \frac{Z_t}{\alpha Y_t(f)/K_t(f)} = \mathrm{MC}_t(f),$$

where $\mathrm{MC}_t(f)$ is the Lagrange multiplier, interpretable as the marginal cost of producing output. By combining the two first order conditions, it is easy to conclude that $L_t(f)/K_t(f) = (1-\alpha)Z_tP_t/(\alpha W_t)$. With this condition and the production function $Y_t(f) = A_tK_t(f)^{\alpha}L_t(f)^{1-\alpha}$, we could rewrite $\mathrm{MC}_t(f)$ as $\mathrm{MC}_t(f) = \frac{1}{A_t} \left(\frac{W_t/P_t}{1-\alpha}\right)^{1-\alpha} \left(\frac{Z_t}{\alpha}\right)^{\alpha} \equiv \mathrm{MC}_t$. Note that the marginal cost and the gross markup are reciprocal; that is, $1 + \mu_t = \frac{1}{\mathrm{MC}_t}$, where μ_t is defined as the markup.

The intermediate good firms set prices on a staggered basis. Following Calvo (1983), each period a firm adjusts its price with probability $1 - \theta$ and keeps it fixed with probability θ . All firms have the same likelihood of adjustment. The adjustment probability is independent over time and across firms. The average time for a price remaining fixed is given by

$$(1-\theta)\sum_{i=1}^{\infty}\theta^{i-1}i = \sum_{i=0}^{\infty}\theta^i = \frac{1}{1-\theta}.$$

Firms that are able to adjust their prices choose $P_t(f), Y_t(f), K_t(f)$ and $L_t(f)$. These firms maximize expected discounted profits given the production technology and the demand curve.

They choose the optimal reset price $P_t^o(f)$ to maximize

$$\mathbb{E}_{t}\left\{\sum_{i=0}^{\infty}\theta^{i}\left[\Lambda_{t,i}\left(\frac{P_{t}^{o}(f)}{P_{t+i}}-\mathrm{MC}_{t+i}\right)Y_{t,t+i}(f)\right]\right\}$$

subject to

$$Y_{t,t+i}(f) = \left(\frac{P_t^o(f)}{P_{t+i}}\right)^{-\varepsilon} Y_{t+i},$$

where $\mathbb{E}_t(\cdot)$ is the mathematical expectation operator, $\Lambda_{t,i} = \beta^i C_{t+i}^{-\gamma} / C_t^{-\gamma}$ is the stochastic discount factor, ε is the elasticity of substitution between intermediate goods, and MC_{t+i} is the nominal marginal cost. Then, the first order condition is

$$\mathbb{E}_t \left\{ \sum_{i=0}^{\infty} \theta^i \Lambda_{t,i} Y_{t,t+i}(f) \left[\frac{P_t^o}{P_{t+i}} - (1+\mu) \mathrm{MC}_{t+i} \right] \right\} = 0,$$

where $1 + \mu = \varepsilon/(\varepsilon - 1)$ is the steady state gross markup. Given that (i) all firms that adjust in period t choose the same price P_t^0 and (ii) the average price of firms that do not adjust is simply last period's price level P_{t-1} , then, we can rewrite the price index as

$$P_t = \left[\theta \left(P_{t-1}\right)^{1-\varepsilon} + \left(1-\theta\right) \left(P_t^o\right)^{1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}}.$$

The entrance of P_{t-1} in P_t introduces nominal inertia.

2.2.3 Capital Producer

The capital producer's optimization problem is to choose the investment level I_t such that the profit from producing the capital is maximized:

$$\max_{I_t} [Q_t J_t - I_t]$$

subject to

$$J_t = I_t - \frac{1}{2}c\left(\frac{I_t}{K_t} - \delta\right)^2 K_t,$$

where J_t stands for the technology for producing new capital goods, and the capital producer invests I_t units of final output and rents K_t units of capital to produce J_t units of new capital. c is the adjustment cost parameter, δ is the capital depreciation, and $\frac{1}{2}c(I_t/K_t - \delta)^2 K_t$ reflects increasing marginal costs of producing new capital goods after the depreciation. Then, the first order condition yields

$$\frac{I_t}{K_t} = \delta + \frac{1}{c} \left(1 - \frac{1}{Q_t} \right).$$

2.3 Monetary and Fiscal Authorities

The central bank sets the nominal interest rate according to the following simple feedback rule

$$1 + r_t^n = (1+r) \left(\frac{P_t}{P_{t-1}}\right)^{\phi_\pi} \left(\frac{Y_t}{Y_t^*}\right)^{\phi_y} e^{v_t},$$

where Y_t^* is the natural (i.e. flexible price equilibrium) level of output with $\phi_{\pi} > 1$ and $\phi_y > 0$, r_t^n is the nominal interest rate, and r is the zero inflation steady state nominal interest rate.

The fiscal policy is given by

$$G_t = \bar{G},$$

where G_t is the government expenditure, and \overline{G} is an exogenously given level of the government spending.

The government budget constraint is

$$G_t = T_t + (M_t - M_{t-1}) / P_t,$$

where T_t is the government's tax revenue, and M_t is the monetary supply.

2.4 Competitive Equilibrium

The resource constraints in the economy contain the income and expenditure constraint

$$Y_t = C_t + I_t + G_t,$$

and the evolution of capital constraint

$$K_{t+1} = I_t - \frac{1}{2}c\left(\frac{I_t}{K_t} - \delta\right)^2 K_t + (1 - \delta)K_t.$$

A competitive equilibrium is defined as an allocation $(Y_t, L_t, C_t, I_t, K_{t+1})$ and a price system $(Z_t, W_t, P_t, P_t^o, r_t^n, Q_t, \mu_t)$ such that all agents are maximizing subject to their respective constraints, all markets clear, and all resource constraints are satisfied, given P_{t-1}, A_t , and K_t .

In practice, it is convenient to express the equilibrium as a system of 10 equations for $(Y_t, C_t, I_t, L_t, P_t, P_t^o, r_t^n, Q_t, \mu_t, K_{t+1})$, given the predetermined states P_{t-1}, A_t, K_t . It is useful to group the equations into aggregate demand, aggregate supply and policy blocks as follows. First, the aggregate demand block has the following 4 equations:

$$Y_t = C_t + I_t + \bar{G}, \qquad (\text{resource constraint})$$

 $C_t = \mathbb{E}_t \left\{ \left[(1+r_t^n) \frac{P_t}{P_{t+1}} \beta \right]^{-\sigma} C_{t+1} \right\},\$ (consumption Euler equation) $\frac{I_t}{K_t} = \delta + \frac{1}{c} \left(1 - \frac{1}{Q_t} \right),$

(link between asset prices and investment)

and

$$\mathbb{E}_{t}\left\{\Lambda_{t,+1}\left(1+r_{t}^{n}\right)\frac{P_{t}}{P_{t+1}}\right\} = \mathbb{E}_{t}\left\{\Lambda_{t,+1}\left(\frac{Z_{t}+(1-\delta)Q_{t+1}}{Q_{t}}\right)\right\},$$
(marginal cost of funds = marginal return to capital)

with $Z_t = \alpha Y_t / [(1 + \mu_t) K_t]$, $\Lambda_{t,+1} = \beta C_{t+1}^{-\gamma} / C_t^{-\gamma}$, and $\sigma = 1/\gamma$. These equations define an investment-savings curve that relates spending inversely to the real rate $(1 + r_t^n) P_t / P_{t+1}$ and expectations of the future. Second, the aggregate supply block has the following 5 equations as follows:

> $Y_t = A_t K^{\alpha}_t L^{1-\alpha}_t V_t.$ (production function) $(1-\alpha)\frac{Y_t}{L_t} = (1+\mu_t)\frac{L_t^{\varphi}}{C_t^{-\gamma}},$ (labor market equilibrium)

$$P_{t} = \left[\theta \left(P_{t-1}\right)^{1-\varepsilon} + \left(1-\theta\right) \left(P_{t}^{o}\right)^{1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}}, \qquad (\text{price adjustment})$$

$$\mathbb{E}_t \left\{ \sum_{i=0}^{\infty} \theta^i \Lambda_{t,i} \left[\frac{P_t^o}{P_{t+i}} \right]^{-\varepsilon} Y_{t+i} \left[\frac{P_t^o}{P_{t+i}} - \frac{1+\mu}{1+\mu_{t+i}} \right] \right\} = 0,$$
(Phillips curve)

and

$$K_{t+1} = I_t - \frac{1}{2}c\left(\frac{I_t}{K_t} - \delta\right)^2 K_t + (1 - \delta)K_t, \qquad (\text{evolution of capital})$$

with $V_t = \left[\int_0^1 \left(\frac{P_t(f)}{P_t} \right)^{-\varepsilon} df \right]^{-1}$ and $[1 + \mu_{t+i}]^{-1} = MC_{t+i}$. V_t reflects the misallocation of intermediate inputs due to relative price dispersion. Note that $V_t = 1$ in the zero inflation steady state. Finally, the policy block is given by

$$1 + r_t^n = (1+r) \left(\frac{P_t}{P_{t-1}}\right)^{\phi_\pi} \left(\frac{Y_t}{Y_t^*}\right)^{\phi_y} e^{\upsilon_t}, \qquad \text{(interest rate rule)}$$

where Y_t^* denotes the level of output in the flexible price equilibrium (the natural output). Given the exogenous processes for the log of productivity A_t satisfying the AR(1) model:

$$\ln A_t = \rho_a \ln A_{t-1} + \varepsilon_{at},$$

and the monetary shock v_t satisfying the AR(1) model:

$$v_t = \rho_m v_{t-1} + \varepsilon_{mt}.$$

Therefore, the equilibrium system with 10 unknown variables $(Y_t, C_t, I_t, L_t, P_t, P_t^o, r_t^n, Q_t, \mu_t, K_{t+1})$ in the above 10 equations is complete.

§3 Computational Challenges of the Discrete-Time DSGE models

3.1 Log-linearization

To solve this typical nonlinear dynamic stochastic general equilibrium models, the most common practice since Kydland and Prescott (1982) and King et al. (1988) is to approximate the solutions using linear methods, especially the log-linearization method. As an illustration, we follow the model that has been set up above and show the procedure of the log-linearization as below.

Let X with no time subscript, no star superscript, and no tilde denote the level of a variable at the zero-inflation steady state, where $P_t/P_{t-1} = 1^1$. Let $\widetilde{X}_t \equiv \ln X_t - \ln X$ stand for the log-linear deviation of a variable from its zero-inflation steady state. Specially, since the interest rate has already been in a percentage, its log-linearization follows $\widetilde{R}_t^n \equiv \ln R_t^n$, where $R_t^n = 1 + r_t^n$ denotes the gross nominal interest rate. Further, we have the net interest rate $r_t^n \approx \widetilde{R}_t^n$ and $\widetilde{r}_t^n = dr_t^n = r_t^n - r$. Let $\hat{\mu}_t = \mu_t - \mu$ denote the deviation of the markup from its steady state level. Let X_t^* denote the level of a variable in the flexible price equilibrium (the natural level). Let $\rho \equiv -\log \beta$ and $r \approx \beta^{-1} - 1$. Log-linearize around the steady state with zero inflation and we can write the log-linearized equilibrium into three blocks as follows.

The first block is for the aggregate demand given by

$$\widetilde{Y}_t = \frac{C}{Y}\widetilde{C}_t + \frac{I}{Y}\widetilde{I}_t, \qquad (\text{resource constraint})$$

$$\tau(r^n - \tau - r) + \widetilde{C} \qquad (\text{consumption Fully equation})$$

 $\widetilde{C}_{t} = \mathbb{E}_{t} \left\{ -\sigma \left(r_{t}^{n} - \pi_{t+1} - \rho \right) + \widetilde{C_{t+1}} \right\}, \qquad \text{(consumption Euler equation)}$ $\widetilde{I}_{t} - \widetilde{K_{t}} = \frac{1}{c^{\delta}} \widetilde{Q_{t}}, \qquad \text{(link between asset prices and investment)}$

$$\widehat{T}_{t}^{n} - \pi_{t+1} - \rho = (1 - \tau) \left(-\widehat{\mu}_{t} + \widetilde{Y}_{t} - \widetilde{K}_{t} \right) + \tau \widetilde{Q_{t+1}} - \widetilde{Q_{t}},$$

(marginal cost of funds = marginal return to capital) where $\widetilde{Z}_t = -\widehat{\mu}_t + \widetilde{Y}_t - \widetilde{K}_t$, $\pi_t = \widetilde{P}_t - \widetilde{P}_{t-1}$, and $\tau = 1 - \delta \left[\frac{\alpha Y}{(1+\mu)K} + (1-\delta)\right]^{-1} = (1-\delta)/[Z+(1-\delta)].$

The second block is for the aggregate supply formulated as

 $\widetilde{Y}_t = \widetilde{A}_t + \alpha \widetilde{K}_t + (1 - \alpha) \widetilde{L}_t, \qquad (\text{production function})$

$$\widetilde{Y}_t - \widetilde{L}_t = \widehat{\mu}_t + \varphi \widetilde{L}_t + \gamma \widetilde{C}_t,$$
 (labor market equilibrium)

¹Note that in the zero-inflation steady state $r^n = r = \rho$.

$$\pi_t = -\lambda \widehat{\mu}_t + \beta \mathbb{E}_t \{ \pi_{t+1} \}, \qquad (\text{price adjustment and Phillips curve})$$

where $\lambda = (1 - \theta)(1 - \theta\beta)/\theta$, and

$$\widetilde{K_{t+1}} = \delta \widetilde{I}_t + (1-\delta)\widetilde{K}_t, \qquad (\text{evolution of capital})$$

Finally, the last one is for the monetary policy provided by

$$r_t^n = \rho + \phi_\pi \pi_t + \phi_y \left(\widetilde{Y}_t - \widetilde{Y}_t^* \right) + \upsilon_t.$$
 (interest rate rule)

For readers with further interests in the details of the log-linearization method, we provide the detailed technical derivations in Appendix for the above results.

The log-linearization method is a simple and convenient tool to study the equilibrium dynamics. If the shocks driving aggregate fluctuations are small and an interior stationary solution exists, the first-order approximations provide adequate answers to questions such as local existence, determinacy of equilibrium and the size of the second moments of endogenous variables; see, for example, Schmitt-Grohé and Uribe (2004) for details. However, there are several problems that exist in the process of log-linearization. First, the log-linearization omits the higher-order terms. Therefore, the topics that are closely related with the higher-order terms, such as the risk terms, are not able to be fully studied. Due to this shortage, the model's connection with the asset prices and the financial risk is very limited. In addition, the first-order approximation techniques are not well suited to handle questions such as welfare comparisons across alternative stochastic or policy environments.

Second, the log-linearization is not able to provide the global solution for nonlinear models. It is only able to solve the model locally, at the cost of missing the most optimal solution due to missing the true global optimum². This becomes a relevant issue not only for its qualitative and quantitative economic implications but also from an econometric and statistical perspective. When concerned about the estimation of the structural parameters of the model, an econometrician/statistician is more interested in studying the global shape of the approximated likelihood function. This will not be possible if the solution of the model is built from a local approximation. Furthermore, as shown in Fernández-Villaverde and Rubio-Ramírez (2005), it is possible to obtain a better fit of the model to the data as well as more accurate point estimates of the moments of the model by exploiting the nonlinear

²One exception is that the equivalent linearization representation of a nonlinear problem generates the same global optimum but more efficiently.

structure of the economic and mathematical model, which can only be achieved through the use of global methods; see, for instance, Parra-Alvarez (2018) for details.

Finally, we mark that even if the performance of linear methods is disappointing along a number of dimensions, linearization in levels is preferred to log-linearization for both the benchmark calibration and the highly nonlinear cases in some real applications, as argued in Aruoba et al. (2006).

3.2 Nonlinear Methods of Solutions

The ability to find global solutions and estimate highly nonlinear DSGE models is of critical importance for central banks and policy makers with their interests in quantifying the impacts of economic policies in a DSGE model. Since Taylor and Uhlig (1990) and Coleman (1990, 1991), a number of nonlinear solution methods in discrete time have been proposed, including the perturbation methods proposed by Judd and Guu (1997) and the projection methods initiated by Judd (1992), as alternatives to the linear approaches and to the value function iteration. The perturbation method, formally introduced by Fleming (1971), has been applied extensively to economic models by Judd and co-authors. Note that a first order perturbation is equivalent to linearization when performed in levels, and is equivalent to log-linearization when performed in logs. Aruoba et al. (2006) found that higher order perturbations display a much superior performance over linear methods for a trivial marginal cost. These findings are based on the computations in first, second, and fifth order, both in levels and in logs. The projection method, on the other hand, is found to be more stable and accurate than the perturbation method. For the projection method, Aruoba et al. (2006) found that finite elements perform very well for all parameterizations. It is extremely stable and accurate over the range of the state space even for high values of the risk aversion and the variance of the shock. This property is crucial in estimation procedures, where accuracy is required to obtain unbiased estimates. Chebyshev polynomials share all the good results of the finite elements method and are easier to implement. However, in a model where policy functions have complicated local behavior, finite elements might outperform Chebyshev polynomials.

Other methods to solve the DSGE models globally include the state space based approach, as in Krusell and Smith (1997, 1998) with a parametric law of motion and as in Den Haan and Rendahl (2010) with a nonlinear law of motion, and the simulation based approach as in Judd et al. (2011) and Maliar et al. (2011), which solve a model only in the realized ergodic state space in the equilibrium. For more details, the reader is referred to the aforementioned papers.

Even though the nonlinear methods as mentioned above can overcome some shortages of the linear methods by improving the approximation, there exist some limitations as noted by several papers. For example, Taylor and Uhlig (1990) found that the nonlinear solution methods for solving the stochastic growth models are not satisfactory in answering the volatility related questions, such as the relative volatility of investment and consumption. Schmitt-Grohè and Uribe (2004) derived a second-order approximation to the policy function of a general class of dynamic, discrete time, rational expectations models, and showed that the coefficients on the terms linear and quadratic in the state vector in a second-order expansion of the decision rule are independent of the volatility of the exogenous shocks. Therefore, up to the second order, the presence of uncertainty affects only the constant term of the decision rules.

In parallel, a desire to understand the economic phenomena that can not easily be captured by linear models makes nonlinear models more relevant for empirical macroeconomics, especially since the end of the great moderation, such as the financial crisis and the COVID-19 pandemic. As noted by Aruoba et al. (2013), there are several types of nonlinearities that can appear in a nonlinear DSGE model. The first is for the case that decision rules display curvature and possibly asymmetries such as non-convex adjustment costs, and the second is for kinks in decision rules, such as the zero lower bound on interest rates, credit constraints, borrowing constraints, and default. In addition to this categorization, we add a third type of nonlinearities, that are the nonlinearities related with uncertainty, such as the stochastic volatility, time-varying risk premia, rare disasters, Poisson jumps, and Markov switches; see Fernández-Villaverde et al. (2011) and Rudebusch and Swanson (2012) for examples.

More recently, Fernández-Villaverde and Levintal (2018) used a mixture of projection and perturbation methods for computing the equilibrium of DSGE models with rare disasters. They found that the Taylor projection delivers the best accuracy/speed tradeoff compared to the third-order perturbations and the Smolyak collocation. Cao et al. (2023) introduced the global DSGE (GDSGE) framework and a novel global solution method, called simultaneous transition and policy function iterations, for solving DSGE models. In the GDSGE, the state variables and their global domain need to be specified. The algorithm solves jointly for policy and transition functions over the iterations and is a pure projection method using wealth share as an endogenous state variable with an implicit law of motion, different from the standard policy function iteration algorithms as in Coleman (1990, 1991) and Judd (1992). Auclert et al. (2021) and Lee (2024) solved the DSGE models globally in the sequence space. Additionally, Lee (2024) developed the repeated transition method to accurately compute the sequence of the conditional expectation of economic agents utilizing the ergodicity of DSGE models. Neither a parametric law of motion nor parametrized expectation is necessary for the implementation. The method is flexibly applicable to standard macro models with and without micro-level heterogeneity, especially for solving models with substantial nonlinearities in aggregate fluctuations, as the method does not rely on a (potentially misspecified) parametric form of the aggregate law of motions. This method provides a novel angle that a nonlinear model with complex endogenous aggregate states (e.g., heterogeneous-agent models) can be solved using the sufficient statistic approach, and the validity of the approach can be tested based on some theory. For further explorations, there are approaches which adopt the machine learning and deep learning techniques, as in Han et al. (2021), Azinovic et al. (2022), and Fernández-Villaverde et al. (2023), as well as the adaptive sparse grids, as in Winschel and Krätzig (2010) and Brumm and Scheidegger (2017).

The estimation of the DSGE models can be categorized into likelihood based approaches and moments based approaches; see DeJong and Dave (2007) and Canova (2007). The likelihood based approaches use nonlinear filters for the construction of the likelihood function, such as the particle filter and the Kalman filter. Farmer (2021) developed the discretization filter for approximating the likelihood of nonlinear, non-Gaussian state space models. The major difficulty that arises when studying nonlinear state space models is that the likelihood cannot be evaluated recursively in closed form as it can in linear models with the Kalman filter. The discretization filter solves this problem by constructing a discrete-valued Markov chain that approximates the dynamics of the state variables.

The moments based approaches for estimating the nonlinear DSGE models include the generalized method of moments (GMM); reviewed as in Ruge-Murcia (2013), the instrumental variables approach; see Canova (2007), the simulated method of moments (SMM); see Duffie and Singleton (1993), Ruge-Murcia (2007), and Ruge-Murcia (2012), and the indirect inference; see Smith (1993), Dridi et al. (2007), and Creel and Kristensen (2011).

Finally, as noted in Andreasen, et al. (2017), the higher-order approximations often

generate explosive sample paths because of the resulting unstable steady states in the approximated system. The presence of explosive behavior complicates any model evaluation because no unconditional moments exist in this approximation. Any estimation method using unconditional moments, such as GMM or SMM, is inapplicable because it relies on finite moments from stationary and ergodic probability distributions. Non-explosive sample paths are also required for likelihood methods, for instance, when using the particle filter outlined in Fernández-Villaverde and Rubio-Ramírez (2007). To overcome this issue, Andreasen et al. (2017) applied pruning to perturbation approximations of any order and showed how pruning greatly facilitates the inference of DSGE models.

§4 Dynamic Programming Methods for Continuous-Time Macro-Finance Models

The continuous-time models can be dated back to the finance literature since the seminal works by Robert C. Merton and others in 1970s. During 1990s, the continuous-time models were successfully applied to the growth and the neoclassical investment theories. Since the financial crisis in 2008, there has been a boom of continuous-time methods in macroeconomics, especially in the fields of business cycle and financial market. These literatures connect the areas that are seemingly disconnected in the past: finance, macroeconomics, and mathematics as well as statistics. As a bridge, the continuous-time methods can provide a promising framework to integrate asset pricing theories studied in the finance literature to the real side of the economy studied in macroeconomics, together with mathematical and statistical tools.

Theoretically, macroeconomics models in continuous time are preferred over the discretetime models because of their analytical tractability. The continuous-time methods transform optimal control problems into stochastic differential equations, such as the Hamilton-Jacobi-Bellman (HJB) equation, the Kolmogorov forward (KF) equation, and the Black-Scholes model. Solving these SDEs is much simpler than solving the Bellman or the Chapman-Kolmogorov equations in discrete time. Compared with the discrete-time framework, the elegant and powerful mathematics such as differential equations and stochastic processes can be well applied in the continuous-time framework. In fact, it is possible to derive closed-form solutions for a wider class of models in continuous time without the need for strong parametric restrictions. On the handling of borrowing constraints, the continuous-time framework is advantageous in the presence of occasionally binding constraints, as these are dealt with using boundary conditions rather than inequalities of the optimality conditions. Finally, continuous-time methods are suited to studying optimal stopping problems and situations where actions are taken infrequently because they entail a fixed cost impulse control problem as described in Stokey (2009), such as a country's decision to default on its sovereign debt as discussed in Parra-Alvarez (2018).

In particular, the continuous-time framework is well applied in the heterogeneous agent models as in the work of Achdou et al. (2022). When recast in continuous time, heterogeneous agent models boil down to systems of two coupled SDEs. The first SDE is the HJB equation for the optimal choices of a single individual who takes the evolution of the distribution and hence prices as given. An individual's consumption saving decision depends on the evolution of the interest rate which is in turn determined by the evolution of the distribution. And, the second SDE is the KF equation characterizing the evolution of the distribution, given optimal choices of individuals. The evolution of the distribution depends on individuals' saving decisions. More generally, this approach is to cast heterogeneous agent models in terms of the mathematical theory of the mean field games (MFG) initiated by Huang et al. (2003, 2007) and Lasry and Lions (2007). The system of coupled HJB and KF equations is known as the backward-forward MFG system. The two equations run in opposite directions in time: the HJB equation runs backward and looks forward. It answers the question "given an individual's valuation of income and wealth tomorrow, how much will she save today and what is the corresponding value function today?" In contrast, the KF equation runs forward and looks backward. It answers the question "given the wealth distribution, savings decisions and the random evolution of income today, what is the wealth distribution tomorrow?"

Computationally, continuous time has resurfaced as a popular environment for economic models because of its efficiency in numerical analysis. Solving a workhorse incomplete markets model in continuous time is much faster compared to its discrete-time counterpart as argued in Rendahl (2022). Financial frictions in macroeconomics require the nonlinear techniques. This is especially important for the macro-finance models which include the financial intermediary in the continuous-time framework, such as Chen (2010), Brunnermeier and Sannikov (2014), Phelan (2016), Drechsler, Savov, and Schnabl (2018), He and Krishnamurthy (2019), Hansen et al. (2024) and D'Avernas and Vandeweyer (2024). Modeling this class of problems rarely leads to analytical solutions and needs to resort to the numerical techniques that provide accurate and fast solution methods.

Continuous time imparts a number of computational advantages relative to the discrete time. First, in the continuous-time models, the optimality conditions (the first-order conditions) that describe the equilibrium allocations of a stochastic economy are deterministic; see, Chang (2010). Since the HJB equation does not contain future values of the value function and the optimal policies only depend on the current value function, there is no need to approximate expected values numerically. Hence, the computational cost and the numerical errors can be reduced. This feature can also tame the "curse of dimensionality", with which the standard dynamic programming in discrete-time models often struggle with, given that the dynamic programming equation does not include any composition of functions or expectation operators in the continuous-time framework; see, for example, Doraszelski and Judd (2012). The efficiency of the perturbation methods and the projection methods can be improved in continuous time as well. Since there is no need to approximate composition of unknown functions, neither to numerically approximate the integrals associated with expected values, the approximations use much less computing time in both perturbation and projection methods. Specifically, Parra-Alvarez (2018) assessed the performance of the firstand second-order perturbation and the projection methods to compute an approximated solution of continuous-time DSGE models based on the maximized HJB equation and the first-order conditions. It is found that the fit of perturbation deteriorates when the degree of nonlinearities increases and the approximated value is different from that obtained by global methods. Despite perturbation being only locally accurate, the increase in the order of approximation improves substantially the goodness of fit. Aruoba (2006) pointed that the projection methods are more accurate and robust than perturbations for a wide range of values of the state-space centered around the deterministic steady state, similarly to the discrete-time case. Their accuracy extends to different degrees of nonlinearities.

Second, the continuous-time problems with discretized state space are very sparse. The sparse structure of the implicit method (with finite differences) is by many considered the most important practical implication of continuous time, such as the sparsity of transition matrices as discussed in Rendahl (2022).

Third, the viscosity solutions and finite difference methods can handle the non-differentiable and non-convex problems in continuous time without the need to change the algorithm, while these problems are difficult to handle in the standard discrete time methods. To compare and contrast the performance of the continuous-time methods and the discrete-time methods further, Rendahl (2022) compared the value function iteration for discrete time, and the explicit and implicit finite difference methods for continuous time. The implicit finite difference method is the continuous time equivalent to Howard's improvement algorithm in discrete time, yet the implicit method is faster. Since the HJB equation does not contain future values of the value function, the update of the value function can be formulated as the solution to a system of linear equations. Through the sparse matrix operations, this system can be solved efficiently. The explicit finite difference method in continuous time is where the value function iteration in discrete time converges to. Since the value function iteration is less efficient, the explicit method generally has slow convergence.

Mathematically or statistically, the continuous-time models do not impose a priori a perfect synchronization of decisions among economic agents, since the decision interval of the model is not tied to the observation interval in the data. They also allow for a clear distinction between flows and stocks in the economy; see, the paper by Parra-Alvarez (2018) for more discussions.

4.1 Hamilton-Jacobi-Bellman Equations

The HJB model is one of the most popular mathematical and statistical models in macroeconomics to characterize the macroeconomic activities; see, for example, Fernández-Villaverde and Nuño (2021) for details. The basic setup is that an agent maximizes the following

$$\max_{\{\alpha_t\}_{t=0}^{\infty}} \mathbb{E}_0\left[\int_0^\infty e^{-\rho t} u\left(\alpha_t, x_t\right) dt\right]$$

subject to the law of motion for the state

$$dx_t = \mu_t \left(x_t, \alpha_t \right) dt,$$

where $x_t \in \mathbb{X} \subset \mathbb{R}^N$ is the state variable (such as K_t in the new Keynesian model), $\alpha_t \in \mathbb{A} \subset \mathbb{R}^M$ is the control variable (such as C_t , L_t , and K_{t+1} in the new Keynesian model), $\alpha_t = \alpha_t(x_t)$ is the policy function, $\rho > 0$ is the discount factor (recall that in the notation of new Keynesian model: $\beta = e^{-\rho}$), $\mu(\cdot)$ is the drift, and $u(\cdot)$ is the instantaneous utility. It satisfies the HJB equation as follows

$$\rho V(x) = \frac{\partial V}{\partial t} + \max_{\alpha} \left\{ u(\alpha, x) + \mu(x, \alpha)^{\top} \nabla_x V(x) \right\}$$

with a transversality condition $\lim_{T\to\infty} e^{-\rho T} V_T(x) = 0$, where V(x) stands for the value function at time t with the time subscript dropped, which might abuse notation. Here, A^{\top} denotes the transpose of a vector or matrix A and $\operatorname{tr}(\cdot)$ stands for the trace of a matrix. Note that in the HJB equation, we use x to denote the vector of the state variables (dimension $N \times 1$), $\mu(x, \alpha)$ as the vector of the drift terms (dimension $N \times 1$), and $\sigma(x, \alpha)$ as the vector of the risk terms (dimension $N \times 1$), and $\sigma^2(x, \alpha)$ stands for the $N \times N$ variance-covariance matrix. Finally, we use $\nabla_x V(x)$ to denote the gradient of V(x) (dimension $N \times 1$) and $\Delta_x V(x)$ to denote the Hessian matrix of V(x) (dimension $N \times N$).

In a diffusion format, popular in the finance literature, the state is now

$$dx_{t} = \mu_{t} \left(x_{t}, \alpha_{t} \right) dt + \sigma_{t} \left(x_{t}, \alpha_{t} \right) d\mathbb{W}_{t}$$

where \mathbb{W}_t is the standard Brownian motion. Then, the HJB equation is given by

$$\rho V(x) = \frac{\partial V}{\partial t} + \max_{\alpha} \left\{ u(\alpha, x) + \mu(x, \alpha)^{\top} \nabla_x V(x) + \frac{1}{2} \operatorname{tr} \left(\Delta_x V(x) \sigma^2(x, \alpha) \right) \right\}.$$

To characterize a possible extension, one can use the following jump model

$$dx_t = \mu_t \left(x_t, \alpha_t, s_t \right) dt,$$

where s_t is a two-state continuous-time Markov chain $s_t \in \{s_1, s_2\}$ and the Poisson process jumps from state 1 to state 2 with intensity λ_1 and vice-versa with intensity λ_2 , which is the so-called famous Markov switching (diffusion) model in the finance literature. Thus, the HJB equation for this case is

$$\rho V_i(x) = \frac{\partial V_i}{\partial t} + \max_{\alpha} \left\{ u(\alpha, x) + \mu \left(x, \alpha, s_i \right)^\top \nabla_x V_i(x) \right\} + \lambda_i \left(V_j - V_i \right)$$

for i, j = 1, 2 for $i \neq j$, where $V_i(x) \equiv V(x, s_i)$ denotes the value function at time t and state s_i , with the time subscript dropped. When the HJB equation includes both the volatility and jumps, we can have the following jump-diffusion process:

$$dx_t = \mu_t \left(x_t, \alpha_t, s_t \right) dt + \sigma_t \left(x_t, \alpha_t \right) dW_t,$$

where s_t and $\sigma_t(x_t, \alpha_t) dW_t$ constitute a Lévy process. Hence, the HJB equation that we want to solve numerically is given by

$$\rho V_i(x) = \frac{\partial V_i}{\partial t} + \max_{\alpha} \left\{ u(\alpha, x) + \mu \left(x, \alpha, s_i \right)^\top \nabla_x V_i(x) + \frac{1}{2} \operatorname{tr} \left(\Delta_x V_i(x) \sigma^2(x, \alpha) \right) \right\} + \lambda_i \left(V_j - V_i \right)$$

with a transversality condition $\lim_{T\to\infty} e^{-\rho T} V_T(x) = 0$, and some boundary conditions defined by the dynamics of x_t .

4.2 Fokker-Plank Equations

Given a stochastic process x_t with an associated infinitesimal generator \mathcal{A} , its probability density function g(x) is defined as: $\mathbb{P}_{t_0}[x_t \in \Omega] = \int_{\Omega} g(x) dx$ for any $\Omega \subset \mathbb{X}$ following the dynamics: $\partial g/\partial t = \mathcal{A}^* g$, where \mathcal{A}^* is the adjoint operator of \mathcal{A} . Note that we omit the time subscript of the function g(x). Let x_t be a stochastic process given by the SDE as follows

$$dx_{t} = \mu_{t}\left(x_{t}\right)dt + \sigma_{t}\left(x_{t}\right)d\mathbb{W}_{t},$$

which is a diffusion model or the so-called Black-Scholes model, widely used in the finance literature. The evolution of the associated density is given by:

$$\frac{\partial g}{\partial t} = \mathcal{A}^* g = -\frac{\partial}{\partial x} \left[\mu(x)g(x) \right] + \frac{1}{2} \frac{\partial^2}{\partial x^2} \left[\sigma^2(x)g(x) \right]$$

with initial value $g_0(x_t) = \delta(x_t - x_0)$, which is the well known Kolmogorov forward equation.

Now, consider the case with a Poisson jump. To do so, let x_t be a stochastic process given by:

$$dx_t = \mu_t \left(x_t, s_t \right) dt,$$

where s_t is a two-state continuous-time Markov chain $s_t \in \{s_1, s_2\}$ with intensities λ_1 and λ_2 , respectively. Then, the evolution of the density is given by:

$$\partial g_i / \partial t = \mathcal{A}^* g = -\frac{\partial}{\partial x} \left[\mu \left(x, s_i \right) g_i(x) \right] - \lambda_i g_i(x) + \lambda_j g_j(x)$$

for i, j = 1, 2 for $j \neq i$ with initial value $g_{1,0}(x_t) = \delta(x_t - x_0)$ and $g_{2,0}(x_t) = 0$. Next, for x_t following a Lévy process

$$dx_{t} = \mu_{t} \left(x_{t}, s_{t} \right) dt + \sigma_{t} \left(x_{t} \right) d\mathbb{W}_{t}$$

the evolution of the density is given by:

$$\partial g_i/\partial t = \mathcal{A}^* g = -\frac{\partial}{\partial x} \left[\mu\left(x, s_i\right) g_i(x) \right] - \lambda_i g_i(x) + \lambda_j g_j(x) + \frac{1}{2} \frac{\partial^2}{\partial x^2} \left[\sigma^2(x) g_i(x) \right].$$

4.3 Markov Equilibrium

Similar to the discrete-time model, a competitive equilibrium in the continuous time model is characterized by the market prices together with the allocations, such that given prices, agents optimize and markets clear. From the perspective of the macroeconomic dynamics, the continuous-time framework can be used to further study the stationary equilibrium and

Type of equilibrium	Property	Steady state	Dynamics of state variables
Competitive equilibrium	Deterministic (no shock)	Converge to a deterministic steady state (stationary value); Time-invariant competitive equilibrium	Deterministic evolution
Stationary equilibrium	Stochastic (subject to shocks, but volatilities are at constant levels); Time-invariant Markov equilibrium	Converge to a stochastic steady state (the value of the state variable with the highest probability from the stationary distribution)	Evolution with constant volatilities eg. $\sigma_t = \sigma$
Markov equilibrium	Stochastic (subject to shocks and volatilities are time-varying)	Converge to a stationary distribution	Evolution with time-varying volatilities eg. $\sigma_t \neq \sigma$

Table 1: Types of equilibrium in macroeconomics models

the Markov equilibrium due to its advanced stochastic features. In particular, the three types of equilibrium in macroeconomics models are summarized in Table 1 below.

The stationary equilibrium is reached when the volatilities in the economic system are at constant levels, even though the economy is still subject to the fundamental shocks. A Markov equilibrium is a set of functions for the control variables, the monetary and fiscal policies, the drifts and diffusions such that the agents' optimal controls solve their respective HJB equations given the law of motion of the state variable. In contrast to the stationary equilibrium, the Markov equilibrium is where there is uncertainty in the value of the state variable and its law of motion is endogenously solved by other structural variables in the model. The state variable reaches a stationary distribution in the Markov equilibrium, in contrast to the case where the state variable reaches a steady state value in the competitive equilibrium. A stochastic steady state under the stationary distribution is the value of the state variable with the highest frequency in the stationary distribution³.

To solve the Markov equilibrium in continuous time, firstly we need to solve the optimization problems and derive the allocations and prices as smooth functions of the state variable. Then, define the stochastic process for the state variable and derive the law of motion of the state variable, which is determined by the structural variables in the model

 $^{^{3}}$ Note that the stationary equilibrium in our paper is not equivalent to the stationary value, which stands for the deterministic steady state under the competitive equilibrium. Indeed, Achdou et al. (2022) defined the stationary equilibrium as the time-invariant competitive equilibrium. However, we refer this concept as the stationary value in this paper.

using Ito's lemma. Finally, solve for the equilibrium by converting the stochastic differential equations into a system of ODEs in the asset prices to find a numerical solution if there is no explicit solution; see, for example, the book by Zhu et al. (2013) for more details. The ODEs can be solved using appropriate boundary conditions.

4.4 Solution Methods

There are several methods to solve the SDEs (the HJB equations and the KF equations) in the continuous-time models. First, the finite difference methods can approximate the derivatives by differences; see, for instance, the book by Zhu et al. (2013) for details. Second, the perturbation method can use a Taylor expansion of order q to solve the SDEs around the deterministic steady state. Third, the projection method can project the value function over a subspace of functions.

As an illustrative example, an implicit upwind finite difference scheme is used to solve a simple heterogeneous agent model, the so-called Huggett (1993) model. This scheme converges to the viscosity solution of the problem, as long as it satisfies three properties: monotonicity, stability, and consistency. To be specific, an agent maximizes:

$$\max_{\{c_t\}_{t=0}^{\infty}} \mathbb{E}_0\left[\int_0^\infty e^{-\rho t} u\left(c_t\right) dt\right]$$

subject to $da_t = (s_t + R_t a_t - c_t) dt$, which yields a global solution, where $s_t \in \{s_1, s_2\}$ is a Markovian chain with intensities $s_1 \to s_2 : \lambda_1$ and $s_2 \to s_1 : \lambda_2$, a_t is the agent's wealth at the period t, c_t is the consumption, $\rho > 0$, $u(\cdot)$ is the momentary utility function, R_t is the gross real interest rate, and s_t is the idiosyncratic endowment. Then the HJB equation is as follows

$$\rho V_i(a) = \max_{c} \left\{ u(c) + \mu_i(a) V'_i(a) \right\} + \lambda_i \left(V_j(a) - V_i(a) \right)$$

for i = 1 and 2, where $\mu_i(a)$ is the drift with $\mu_i(a) = s_i + Ra - c(a)$ and $V'_i(a)$ is the derivative of $V_i(a)$. When the market clears, the aggregate income normalizes to one; that is, $\mathbb{E}[s_t] = 1$. Total assets in zero net supply is $\sum_{i=1}^2 \int ag_t(a, s_i) da = 0$, where $g_t(a, s_i)$ is the income wealth density. Using the implicit finite difference scheme, we can write the finite difference approximation of the HJB equation as:

$$\frac{V_{i,j}^{n+1} - V_{i,j}^{n}}{\Delta} + \rho V_{i,j}^{n+1} = u \left(c_{i,j}^{n} \right) + \left(V_{i,j}^{n+1} \right)' \left(s_j + Ra_i - c_{i,j}^{n} \right) + \lambda_j \left(V_{i,-j}^{n+1} - V_{i,j}^{n+1} \right),$$

where the derivative $V'_{i,j} = V'_j(a_i)$ is approximated with either a forward or a backward difference approximation

$$V'_{i,j,F} \equiv \frac{V_{i+1,j} - V_{i,j}}{\Delta a}$$
, and $V'_{i,j,B} \equiv \frac{V_{i,j} - V_{i-1,j}}{\Delta a}$

Under the upwind scheme, using the forward difference approximation whenever the drift of the state variable is positive and the backward difference approximation whenever it is negative, we have

$$\mu_{i,j,F} = s_j + Ra_i - c_{i,j,F}, \quad \mu_{i,j,B} = s_j + Ra_i - c_{i,j,B},$$

and

$$V'_{i,j} = V'_{i,j,F} \mathbf{1}_{\{\mu_{i,j,F} > 0\}} + V'_{i,j,B} \mathbf{1}_{\{\mu_{i,j,B} < 0\}} + \bar{V}'_{i,j} \mathbf{1}_{\{\mu_{i,j,F} \leqslant 0 \leqslant \mu_{i,j,B}\}},$$

where $\bar{V}'_{i,j} = u' (s_j + Ra_i)$. Thus, the HJB equation can be written as:

$$\frac{V_{i,j}^{n+1} - V_{i,j}^{n}}{\Delta} + \rho V_{i,j}^{n+1} = u \left(c_{i,j}^{n} \right) + \left(V_{i,j,F}^{n+1} \right)' \left[s_{j} + Ra_{i} - c_{i,j,F}^{n} \right]^{+} + \left(V_{i,j,B}^{n+1} \right)' \left[s_{j} + Ra_{i} - c_{i,j,B}^{n} \right]^{-} + \lambda_{j} \left[V_{i,-j}^{n+1} - V_{i,j}^{n+1} \right]$$

This equation constitutes a system of $2 \times I^4$ linear equations, and it can be written in matrix notation using the following steps

$$\frac{V_{i,j}^{n+1} - V_{i,j}^{n}}{\Delta} + \rho V_{i,j}^{n+1} = u\left(c_{i,j}^{n}\right) + \frac{V_{i+1,j}^{n+1} - V_{i,j}^{n+1}}{\Delta a}\left(\mu_{i,j,F}^{n}\right)^{+} + \frac{V_{i,j}^{n+1} - V_{i-1,j}^{n+1}}{\Delta a}\left(\mu_{i,j,B}^{n}\right)^{-} + \lambda_{j}\left[V_{i,-j}^{n+1} - V_{i,j}^{n+1}\right].$$

Collecting terms with the same subscripts on the right-hand side, it leads to the following

$$\frac{V_{i,j}^{n+1} - V_{i,j}^{n}}{\Delta} + \rho V_{i,j}^{n+1} = u\left(c_{i,j}^{n}\right) + V_{i-1,j}^{n+1} x_{i,j} + V_{i,j}^{n+1} y_{i,j} + V_{i+1,j}^{n+1} z_{i,j} + V_{i,-j}^{n+1} \lambda_{j},$$

where $x_{i,j} = -(\mu_{i,j,B}^n)^- /\Delta a, y_{i,j} = -\left[(\mu_{i,j,F}^n)^+ - (\mu_{i,j,B}^n)^-\right] /\Delta a - \lambda_j$, and $z_{i,j} = (\mu_{i,j,F}^n)^+ /\Delta a$. Thus, this equation can be written in the matrix notation as $\frac{1}{\Delta}(v^{n+1} - v^n) + \rho v^{n+1} = u^n + \mathbf{A}^n v^{n+1}$, where

$$\mathbf{A}^{n} = \begin{pmatrix} y_{1,1} & z_{1,1} & 0 & \dots & 0 & \lambda_{1} & 0 & 0 & \dots & 0 \\ x_{2,1} & y_{2,1} & z_{2,1} & 0 & \dots & 0 & \lambda_{1} & 0 & 0 & \dots \\ 0 & x_{3,1} & y_{3,1} & z_{3,1} & 0 & \dots & 0 & \lambda_{1} & 0 & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ddots & \ddots & x_{I,1} & y_{I,1} & 0 & 0 & 0 & 0 & \lambda_{1} \\ \lambda_{2} & 0 & 0 & 0 & y_{1,2} & z_{1,2} & 0 & 0 & 0 \\ 0 & \lambda_{2} & 0 & 0 & 0 & x_{2,2} & y_{2,2} & z_{2,2} & 0 & 0 \\ 0 & 0 & \lambda_{2} & 0 & 0 & 0 & x_{3,2} & y_{3,2} & z_{3,2} & 0 \\ 0 & 0 & \ddots \\ 0 & \cdots & \cdots & 0 & \lambda_{2} & 0 & \cdots & 0 & x_{I,2} & y_{I,2} \end{pmatrix}_{2I \times 2I}$$

,

⁴We assume that $i = 1, \ldots, I$ and $j = 1, \ldots, J$.

which is from the final HJB iteration, and $u^n = (u(c_{1,1}^n), \ldots, u(c_{1,1}^n), u(c_{1,2}^n), \ldots, u(c_{I,2}^n))^{\top}$. This system can be written as $\mathbf{B}^n v^{n+1} = b^n$, $\mathbf{B}^n = (\frac{1}{\Delta} + \rho) \mathbf{I} - \mathbf{A}^n$, and $b^n = u^n + \frac{1}{\Delta} v^n$, which can be solved very efficiently using sparse matrix (since \mathbf{A}^n is a sparse matrix) routines.

To solve the KF equation, we have to solve the following ODE using the finite difference method

$$0 = -\frac{d}{da} \left[\mu_j(a)g_j(a) \right] - \lambda_j g_j(a) + \lambda_{-j}g_{-j}(a).$$

where $g(\cdot)$ is the density function. With the Poisson jumps, $g_j(\cdot)$ denotes the density function under state j, and $g_{-j}(\cdot)$ is the density function under state -j. Then, the ODE can be approximated by

$$-\frac{g_{i,j}\left(\mu_{i,j,F}^{n}\right)^{+}-g_{i-1,j}\left(\mu_{i-1,j,F}^{n}\right)^{+}}{\Delta a}-\frac{g_{i+1,j}\left(\mu_{i+1,j,B}^{n}\right)^{-}-g_{i,j}\left(\mu_{i,j,B}^{n}\right)^{-}}{\Delta a}-g_{i,j}\lambda_{j}+g_{i,-j}\lambda_{-j}=0.$$

Collecting terms, we obtain

$$g_{i-1,j}z_{i-1,j} + g_{i,j}y_{i,j} + g_{i+1,j}x_{i+1,j} + g_{i,-j}\lambda_{-j} = 0,$$

where $x_{i+1,j} = -(\mu_{i,j+1,B}^n)^- /\Delta a$, $y_{i,j} = -\left[(\mu_{i,j,F}^n)^+ - (\mu_{i,j,B}^n)^-\right] /\Delta a - \lambda_j$, and $z_{i-1,j} = (\mu_{i,j-1,F}^n)^+ /\Delta a$. This approximation can be written in the matrix form $\mathbf{A}^\top g = 0$ where \mathbf{A}^\top is the transpose of the intensity matrix \mathbf{A} ($\mathbf{A} = \lim_{n \to \infty} \mathbf{A}^n$) from the HJB equation.

The matrix \mathbf{A} captures the evolution of the discrete Poisson process (continuous-time Markov chain). To find the stationary distribution, one solves the eigenvalue problem $\mathbf{A}^{\mathrm{T}}g = 0$, a system of 2 × J linear equations. The reason for using the transpose of the intensity matrix \mathbf{A} can be made more precise by the differential operators, so that one can write the HJB equation in terms of a differential operator \mathcal{A} , the infinitesimal generator of the process. Similarly, the Kolmogorov forward equation can be written in terms of an operator \mathcal{A}^* , the adjoint of the operator \mathcal{A} in the HJB equation, which is the infinite-dimensional analogue of a matrix transpose. \mathbf{A} is simply the discretized infinitesimal generator, whereas \mathbf{A}^{\top} is the discretized version of its adjoint, the Kolmogorov forward operator.

4.5 Transition Dynamics and Impulse Responses

In this section, we provide several recent methods in the existing literature on computing the transition dynamics and impulse response functions in the continuous-time models. The computational method for solving transition dynamics from an arbitrary initial condition can also be used to compute nonlinear impulse responses to unanticipated aggregate shocks, the so-called MIT shocks, i.e., an unanticipated (zero probability) shock followed by a deterministic transition as in Krusell and Smith (1998). Recently, Fernández-Villaverde, et al. (2023) further extended the model in Krusell and Smith (1998) by proposing a nonparametric perceived law of motion and updated with machine learning, and Boppart et al. (2018) and Auclert et al. (2021) used the linearized counterpart to compute linear impulse responses to small MIT shocks in order to obtain further speed gains. Additionally, both Auclert et al. (2021) and Oskolkov (2023) used the sequence-space Jacobians to compute the impulse responses. For the type of the representative agent model, He and Krishnamurthy (2019) provided a method to compute the impulse responses in a continuous-time framework. Now, let us introduce these methods in detail.

When there is aggregate uncertainty, Krusell and Smith (1998) proposed a bounded rationality method. Households in the model approximate the distribution by a number of its moments, e.g., the mean $\int_0^{\infty} \int_{\underline{s}}^{\overline{s}} ag(a, s) dads = K_t$, where s_t is the labor productivity following the Ornstein-Uhlenbeck process as $ds_t = \theta (\hat{s} - s_t) dt + \sigma dB_t$. The production function is given by $Y_t = F(A_t, K_t, L_t) = A_t K_t^{\alpha} L_t^{1-\alpha}$ with the aggregate total factor productivity (TFP) shock A_t , following a diffusion process as $dA_t = \mu_A(A_t) dt + \sigma_A(A_t) dW_t$. The capital K_t evolves according to $dK_t = \mu_K (K_t, A_t) K_t dt$. Then, the HJB simplifies to

$$\rho V(a, s, A, K) = \max_{c \ge 0} u(c) + [ws + Ra - c] \frac{\partial V}{\partial a} + \theta(\hat{s} - s) \frac{\partial V}{\partial s} + \frac{\sigma^2}{2} \frac{\partial^2 V}{\partial s^2} + \mu_A(A) \frac{\partial V}{\partial A} + \frac{\sigma^2_A(A)}{2} \frac{\partial^2 V}{\partial A^2} + K \mu_K(K, A) \frac{\partial V}{\partial K}.$$

Suppose the perceived law of motion (PLM) has a simple parametric (linear) form:

$$\mu_K(K, A; \boldsymbol{\theta}) = \theta_0 + \theta_1 K + \theta_2 K A + \theta_3 A.$$
(PLM)

Then, begin with an initial guess of $\boldsymbol{\theta}^0 = (\theta_0^0, \theta_1^0, \theta_2^0, \theta_3^0)$, and set n := 0,

- 1. Given $\mu_K(K, A; \theta^0)$, solve the HJB equation and obtain matrix **A**.
- 2. Conduct the Monte Carlo simulation to obtain the simulated data $\{A_m\}_{m=0}^M$ based on the model $\Delta A_m = \mu_A (A_{m-1}) \Delta t + \sigma_A (A_{m-1}) \sqrt{\Delta t} \varepsilon_m$, where $\varepsilon_m \sim \mathcal{N}(0, 1)$.
- 3. Compute the dynamics of the distribution using the KF equation and use it to obtain aggregate capital: $\int_0^\infty \int_{\underline{s}}^{\underline{s}} ag(a, s) dads = K_t.$

4. Run an ordinary least squares $\Delta K_m/K_m = \mu_K(K_m, A_m; \theta) \Delta t$ over the simulated sample $\{A_m, K_m\}_{m=0}^M$ to update coefficients $\boldsymbol{\theta}^{n+1}$. If $\boldsymbol{\theta}^{n+1} = \boldsymbol{\theta}^n$ or they are very close, stop, otherwise go back to step 1.

For models with aggregate nonlinear dynamics, which is a general form of PLM, Fernández-Villaverde et al. (2023) extended the Krusell and Smith (1998) methodology and proposed a nonparametric perceived law of motion to globally compute and estimate the heterogeneous agent model, updated using machine learning such as a neural network. As claimed by Fernández-Villaverde et al. (2023), their algorithm can approximate the PLM arbitrarily well; see, for instance, the paper by Fernández-Villaverde et al. (2023) for details.

Households consider a PLM of aggregate debt B_t :

$$dB_t = h\left(B_t, N_t\right) dt,$$

where $h(B_t, N_t) = \mathbb{E}[dB_t | B_t, N_t]/dt$ and $N_t = K_t - B_t$, the net wealth (i.e., inside equity) of the expert, which is the difference between his assets (capital) and liabilities (debt). Given the PLM, the household's HJB equation becomes to the following:

$$\rho V_i(a, B, N) = \max_c \frac{c^{1-\gamma} - 1}{1-\gamma} + \mu_i(a, B, N) \frac{\partial V_i}{\partial a} + h(B, N) \frac{\partial V_i}{\partial B} + \mu^N(B, N) \frac{\partial V_i}{\partial N} + \lambda_i \left[V_j(a, B, N) - V_i(a, B, N) \right] + \frac{\left[\sigma^N(B, N) \right]^2}{2} \frac{\partial^2 V_i}{\partial N^2},$$

where the net wealth N_t evolves as $dN_t = \mu^N (B_t, N_t) dt + \sigma^N (B_t, N_t) dW_t$, and the household's saving a_t follows $da_t = (w_t s_t + R_t a_t - c_t) dt = \mu_i (a_t, B_t, N_t) dt$ and $B_t \equiv \int a dG_t(a, s)$. Instead of using the projection method to approximate the PLM: $h(\boldsymbol{\mu}; \theta) = \theta_0 + \sum_{q=1}^Q \theta_q \psi_q(\boldsymbol{\mu})$, Fernández-Villaverde et al. (2023) approximated the PLM with a neural network as $h(\boldsymbol{\mu}; \theta) = \theta_0 + \sum_{q=1}^Q \theta_q^1 \psi_q(\boldsymbol{\mu})$, $\theta_0^1 + \sum_{q=1}^Q \theta_q^1 \phi \left(\theta_{0,q}^2 + \sum_{i=1}^D \theta_{i,q}^2 \mu^i\right)$, where $\phi(\cdot)$ is an activation function, such as $\phi(x) = \log(1 + e^x)$, and θ is selected to minimize the quadratic error function $\mathcal{E}(\theta; \boldsymbol{\mu}, \mathbf{h})$; that is,

$$\theta^* = \arg\min_{\theta} \mathcal{E}(\theta; \boldsymbol{\mu}, \mathbf{h}) = \arg\min_{\theta} \sum_{j=1}^{J} \mathcal{E}\left(\theta; \mu_j, \hat{h}_j\right) = \arg\min_{\theta} \frac{1}{2} \sum_{j=1}^{J} \left\| h\left(\mu_j; \theta\right) - \hat{h}_j \right\|^2.$$

Fernández-Villaverde et al. (2023) solved the HJB equation with a first-order, implicit upwind scheme using the finite difference method. Note that some other machine learning or deep learning methods can be applied to solve the above system too.

In a linearized system, Boppart et al. (2018) employed the MIT shock and obtained the first-order perturbation solution just by computing transitional dynamics. The initial state

is the deterministic steady state $f_{ss}(\cdot)$. The aggregate TFP evolves with time according to:

$$\Delta A_0 = \mu_A(A_0) \,\Delta t + \sigma_A(A_0) \,\sqrt{\Delta t}, \quad \text{and} \quad \Delta A_t = \mu_A(A_t) \,\Delta t, \ t > 0,$$

where $A_0 = A_{ss}$. If the model is approximately linear, the response to an MIT shock is the impulse response function of the model. The method can be extended to the case with n shocks $d\mathbf{A} \equiv (dA_0, dA_1, \dots, dA_n)^{\top}$.

Auclert et al. (2021) proposed a general and highly efficient method for solving and estimating the general equilibrium heterogeneous-agent models with aggregate shocks in discrete time. The model is set up in the sequence space by assuming perfect foresight with respect to aggregates. The approach relies on the rapid computation of sequence space Jacobians, the derivatives of perfect-foresight equilibrium mappings between aggregate sequences around the steady state. This algorithm can be combined with a systematic approach of composing and inverting Jacobians to solve for general equilibrium impulse responses. A rapid procedure is obtained for likelihood-based estimation and computation of nonlinear perfect-foresight transitions.

Equilibrium in the sequence space can always be expressed as a solution to a nonlinear system $\mathbf{F}(\mathbf{X}, \mathbf{Z}) = 0$, where \mathbf{X} represents the time path of endogenous variables (usually aggregate prices and quantities) and \mathbf{Z} represents the time path of exogenous shocks. Obtaining the impulse responses of unknowns to shocks, $d\mathbf{X} = -\mathbf{F}_{\mathbf{X}}^{-1}\mathbf{F}_{\mathbf{Z}}d\mathbf{Z}$, requires computing the Jacobians $\mathbf{F}_{\mathbf{X}}$ and $\mathbf{F}_{\mathbf{Z}}$, which are formed by combining Jacobians from different parts of the model.

Similarly, Oskolkov (2023) employed the methods from Kaplan et al. (2018) and Auclert et al. (2021) to analyze nonlinear solutions for aggregate one-time unanticipated shocks. In particular, this paper works in the sequence space and computes the sequence-space Jacobians by solving the linearized version of the coupled system of equations. For estimation, a sample path of x(t) is integrated from a simulated sequence of shocks dW given the parameters, which come from the calibration and determine the steady state. To estimate the parameters of the processes, the simulated method of moments (SMM) of McFadden (1989) is used, which is to compute the moments of these series, to compare them to the moments of simulated sequences, and to look for a combination of parameters that minimize a quadratic distance.

Finally, in a representative agent model, He and Krishnamurthy (2019) studied the effect of -1% shock in σdW_t , which means the fundamental shock leads capital to fall exogenously by 1%. The computational algorithm to calculate the impulse response functions is as follows, where the focus is on the mean path by shutting future shocks to zero.

- 1. Compute the benchmark path of these variables without any shocks, but still subject to the endogenous drift of the state variable in the model. In other words, calculate the benchmark path for the realizations of $dW_{t+m} = 0$ for $m \ge 0$.
- 2. Compute the shocked path of these variables given the initial shock $\sigma dW_t = -1\%$, but setting future realizations of shocks to be zero, i.e., $dW_{t+m} = 0$ for m > 0.
- 3. Calculate and plot the (log) difference between the path with the shock and the mean path without any shock. This computation is meant to mimic a deviation-from-steadystate computation that is typically plotted in impulse-response functions in the linearnon-stochastic models.

Note that in traditional linear models, the impulse-response functions are independent of future shocks. However, the impact of a shock depends on future shocks in nonlinear models. For more on the difference between impulse responses in linear models with a non-stochastic steady state and those nonlinear models with a stochastic steady state; see, for example, the papers by Koop et al. (1996) and Borovička et al. (2011). An alternative method to calculate the impulse response functions in the stochastic nonlinear models is to calculate the expected impact of the initial shock $\sigma dW_t = -1\%$ on the variable at t+m by integrating over all possible future paths.

4.6 A More General Framework

Achdou et al. (2022) extended the study to the backward-forward mean field games system in *n* dimensions, which is a natural generalization of the equations for the Bewley-Huggett-Aiyagari (BHA) models, proposed by Bewley (1987), Hugget (1993), and Aiyagari (1994), respectively. The mathematical MFG literature typically writes this system using the language of the modern theory of SDEs, especially the vector calculus notation, described as follows. For more details about modeling BHA type models, the reader is referred to the papers by Kirkby (2018) and Hansak (2023).

The mathematics literature typically only considers the case where the state variables follow diffusion processes rather than processes featuring jumps. Under this assumption, a general backward-forward MFG system in n dimensions is:

$$\rho V = \max_{\alpha} \left\{ r(x, \alpha, g) + \sum_{i=1}^{n} \alpha_i \partial_i V \right\} + \frac{1}{2} \sum_{i=1}^{n} \sigma_i^2(x) \partial_{ii} V + \partial_t V$$

in $\mathbb{R}^n \times (0,T)$, where we use the short-hand notation $\partial_a v = \partial v / \partial a$, and so on, and

$$\partial_t g = -\sum_{i=1}^n \partial_i \left(\alpha_i^*(x,g)g \right) + \frac{1}{2} \sum_{i=1}^n \partial_{ii} \left(\sigma_i^2(x)g \right)$$

in $\mathbb{R}^n \times (0, T)$, with

$$g_0 = g(0)$$
 and $V_T = V(x, T)$

in \mathbb{R}^n , where V(x,t) is the value function, g(x,t) the density, $x \in \mathbb{R}^n$ an *n*-dimensional state vector, $r(x, \alpha, g)$ a period return function, $\sigma_i^2(x)$ a diffusion coefficient, $\alpha \in \mathbb{R}^n$ a control vector and α^* its optimally chosen policy function. The first equation is the HJB equation, the second equation is the KF equation and the equations in the third line are the initial condition on the density and the terminal condition on the value function. The system iterates backward-forward in the sense that given the steady state value function, the system updates backward using the HJB equation to obtain the policies. Given the initial distribution, the system updates forward using the KF equation to propagate the distribution.

Now, we define three useful operators: the gradient ∇ , the Laplacian Δ and the divergence. First, for a function $f : \mathbb{R}^n \to \mathbb{R}$, the gradient vector is the vector of first derivatives $\nabla f := \left[\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right]^\top$. Second, for a function $f : \mathbb{R}^n \to \mathbb{R}$, the Laplacian is the sum of all the unmixed second derivatives $\Delta f := \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}$. Third, for a vector-valued function $\mathbf{v} : \mathbb{R}^n \to \mathbb{R}^n$, i.e. $\mathbf{v}(x_1, \ldots, x_n) = \left[v_1(x_1, \ldots, x_n), \ldots, v_n(x_1, \ldots, x_n)\right]^\top$, the divergence of \mathbf{v} is $\operatorname{div}(\mathbf{v}) := \sum_{i=1}^n \frac{\partial v_i}{\partial x_i}$. Note that $\Delta f = \operatorname{div}(\nabla f)$.

Then, the backward-forward MFG system above describes general heterogeneous agent models without aggregate uncertainty. However, in many economically interesting situations, it is important to allow for aggregate risk in addition to idiosyncratic risk as in Den Haan (1997) and Krusell and Smith (1998). Fortunately, the theory of MFGs has also studied that case, with mathematicians referring to aggregate uncertainty as "common noise". In the most general case, such MFGs can be written in terms of the so-called "Master equation" as in Cardaliaguet et al. (2019). This Master equation is an equation on the space of measures, i.e. it is an equation that is set in infinite-dimensional space. In the case without aggregate uncertainty, the Master equation reduces to the backward-forward MFG system.

§5 Conclusion

This selective review outlined the mathematical/statistical tools and the computational methods in mathematics for solving both discrete-time and continuous-time models in macroeconomics. As the mathematical tools and computational methods are more advanced in continuous time, we see a bright future for the macroeconomics modeling in continuous time. For example, in addition to a neural network as employed by Fernández-Villaverde et al. (2023), some advanced machine learning such as deep learning methods or AI methods can be applied to this field too, especially for nonlinear models, which can attract some young scholars and Ph.D. students in economics, mathematics and statistics, to find their own interesting research topics for continuous-time models in macroeconomics. In addition, the well-developed asset pricing theories in continuous time also shed light on the inclusion of financial risk analysis in the macroeconomics models. Looking forward, more work can be for sure done to study the connection between the real economy's business cycles and the financial market's fluctuations within a continuous-time general equilibrium framework. Even though the discrete-time model and the continuous-time model may share similar computational accuracy, we still have much confidence on the scope of the analysis that can be done only in continuous time which can not be substituted by the discrete-time model, such as the formulation of problems related with uncertainty. Finally, similar to the model specification problem for conventional stochastic diffusions, well studied in the literature such as the pioneer work by Aït-Sahalia (1996), it would be interesting to consider some possible model specification tests for heterogenous agent models with aggregate shocks under full or partial information as addressed in Cai et al. (2024), which is definitely warranted as future research. The reader is referred to the paper by Cai, Mei and Wang (2024) for more discussions.

Reference

- Achdou, Y., J. Han, J.M. Lasry, P.L. Lions and B. Moll (2022). Income and wealth distribution in macroeconomics: A continuous-time approach. *Review of Economic Studies*, 89(1):45-86.
- [2] Aiyagari, S.R.(1994). Uninsured idiosyncratic risk and aggregate saving. Quarterly Journal of Economics, 109(3):659–684.
- [3] Aït-Sahalia, Y. (1996). Testing continuous-time models of the spot interest rate. Review

of Financial Studies, 9(2):385–426.

- [4] Andreasen, M.M., J. Fernándes-Villaverde, and J.F. Rubio-Ramírez (2017). The pruned state-space system for non-linear DSGE models: Theory and empirical applications. *Review of Economic Studies*, 85(1):1-49.
- [5] Aruoba, S.B., J. Fernández-Villaverde and J.F. Rubio-Ramírez (2006). Comparing solution methods for dynamic equilibrium economies. *Journal of Economic Dynamics* & Control, **30**(12):2477-2508.
- [6] Auclert, A., B. Bardóczy, M. Rognlie, and L. Straub (2021). Using the sequence-space Jacobian to solve and estimate heterogeneous agent models. *Econometrica*, 89(5):2375-2408.
- [7] Azinovic, M., L. Gaegauf, and S. Scheidegger (2022). Deep equilibrium nets. International Economic Review, 63(4):1471-1525.
- [8] Bewley, T. (1987). Stationary monetary equilibrium with a continuum of independently fluctuating consumers. Contributions to Mathematical Economics in Honor of Gérard Debreu 79.
- [9] Boppart, T., P. Krusell and K. Mitman (2018). Exploiting MIT shocks in heterogeneous agent economies: The impulse response as a numerical derivative. *Journal of Economic Dynamics and Control*, 89(1):68-92.
- [10] Borovička, J., L.P. Hansen, M. Hendricks and J.A. Scheinkman (2011). Risk-price dynamics. *Journal of Financial Econometrics*, 9(1):3-65.
- [11] Brunnermeier, M.S. and Y. Sannikov (2014). A macroeconomic model with a financial sector. American Economic Review, 104(2): 379-421.
- [12] Brumm, J. and S. Scheidegger (2017). Using adaptive sparse grids to solve highdimensional dynamic models. *Econometrica*, 85(5):1575-1612.
- [13] Cai, Z., H. Mei and R. Wang (2024). Model specification tests of heterogenous agent models with aggregate shocks under partial information. *Working Paper*, Department of Economics, University of Kansas. https://econpapers.repec.org/paper/ kanwpaper/202405.htm.
- [14] Cao, D., W. Luo, and G. Nie (2023). Global DSGE models. Review of Economic Dynamics, 51 (December): 199-225.
- [15] Calvo, G.A. (1983). Staggered prices in a utility-maximizing framework. Journal of Monetary Economics, 12(3):383-398.
- [16] Canova F. (2007). Methods for Applied Macroeconomic Research. Princeton University Press, New York.

- [17] Cardaliaguet, P., F. Delarue, J.-M. Lasry, and P.L. Lions (2019). The Master Equation and the Convergence Problem in Mean Field Games. Princeton University Press.
- [18] Chang, F.-R. (2010). Stochastic Optimization in Continuous Time. Cambridge University Press, Cambridge, UK.
- [19] Chen, H. (2010). Macroeconomic conditions and the puzzles of credit spreads and capital structure. *Journal of Finance*, 65(6):2171-2212.
- [20] Coleman, W.J. (1990). Solving the stochastic growth model by policy-function iteration. Journal of Business & Economic Statistics 8(1): 27-29.
- [21] Coleman, W.J. (1991). Equilibrium in a production economy with an income tax. Econometrica 59(4): 1091-1104.
- [22] Creel, M. and Kristensen, D. (2011). Indirect likelihood inference. *Working Paper*, Barcelona GSE.
- [23] D'Avernas, A. and Q. Vandeweyer (2024). Treasury bill shortages and the pricing of short-term assets. *Journal of Finance*, 79(6):4083-4141.
- [24] Den Haan, W.J. (1997). Solving dynamics models with aggregate shocks and heterogeneous agents. *Macroeconomic Dynamics*, 1(2):355-386.
- [25] Den Haan, W. J. and P. Rendahl (2010). Solving the incomplete markets model with aggregate uncertainty using explicit aggregation. *Journal of Economic Dynamics and Control*, 34(1): 69-78.
- [26] DeJong D.N. with C. Dave (2007). Structural Macroeconometrics. Princeton University Press, New York.
- [27] Doraszelski, U. and K.L. Judd (2012). Avoiding the curse of dimensionality in dynamic stochastic games. *Quantitative Economics*, 3(1): 53-93.
- [28] Drechsler, T., A. Savov and P. Schnabl (2018). A model of monetary policy and risk premia. Journal of Finance, 73(1):317-373.
- [29] Dridi, R., Guay, A. and Renault, E. (2007). Indirect inference and calibration of dynamic stochastic general equilibrium models. *Journal of Econometrics*, **136**(2):397-430.
- [30] Duffie, D. and K.J. Singleton (1993). Simulated moments estimation of Markov models of asset prices. *Econometrica*, 61(4):929-952.
- [31] Farmer (2021). The discretization filter: A simple way to estimate nonlinear state space models. *Quantitative Economics*, **12**(1):41-76.
- [32] Fernández-Villaverde, J. and G. Nuño (2021), Lectures Notes on Continuous-Time Methods in Macroeconomics, Department of Economics, University of Pennsylvania.

- [33] Fernández-Villaverde, J., Guerrón-Quintana, P., Rubio-Ramírez, J. F. and Uribe, M. (2011). Risk matters: the real effects of volatility shocks. *American Economic Review*, 101(6):2530-2561.
- [34] Fernández-Villaverde, J., S. Hurtado and G. Nuño (2023). Financial frictions and the wealth distribution. *Econometrica*, 91(3):869-901.
- [35] Fernández-Villaverde, J. and O. Levintal (2018). Solution methods for models with rare disasters. *Quantitative Economics*, **9**(2): 903-944.
- [36] Fernández-Villaverde, J. and J.F. Rubio-Ramírez (2005). Estimating dynamic equilibrium economies: Linear versus nonlinear likelihood. *Journal of Applied Econometrics*, 20(7):891-910.
- [37] Fernández-Villaverde, J. and J. F. Rubio-Ramírez (2007). Estimating macroeconomic models: A likelihood approach. *Review of Economic Studies*, 74(4):1-46.
- [38] Fleming, W.H. (1971). Stochastic control for small noise intensities. SIAM Journal on Control and Optimization, 9(3):473-517.
- [39] Gertler, M. (2024). Lecture Notes on Macroeconomic Theory. Department of Economics, New York University.
- [40] Han, J., Y. Yang and W. E (2021). DeepHAM: A global solution method for heterogeneous agent models with aggregate shocks. SSRN Electronic Journal.
- [41] Hansak, A.P. (2023). Bewley-Hugget-Aiyagari models: Basic properties and numerical solution methods. *Working Paper*, Vienna Graduate School of Economics, University of Vienna.
- [42] Hansen, L.P., P. Khorrami and F. Tourre (2024). Comparative valuation dynamics in models in production economies: Long-run uncertainty, heterogeneity, and market Frictions. *Working Paper*, Department of Economics, University of Chicago.
- [43] He, Z. and A. Krishnamurthy (2019). A macroeconomic framework for quantifying systemic risk. American Economic Journal: Macroeconomics, 11(4):1-37.
- [44] Huang, M., P. E. Caines and R. P. Malhamé (2003). Individual and mass behavior in large population stochastic wireless power control problems: Centralized and Nash equilibrium solutions. In 42nd IEEE International Conference on Decision and Control (IEEE Cat. No. 03CH37475), Volume 1, pp. 98–103. IEEE.
- [45] Huang, M., P.E. Caines and R.P. Malhamé (2007). Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ε -Nash equilibria. *IEEE Transactions on Automatic Control*, **52**(9):1560–1571.
- [46] Huggett, M. (1993). The risk-free rate in heterogeneous agent incomplete insurance economies. Journal of Economic Dynamics and Control, 17(5-6):953-969.

- [47] Judd, K.L. (1992). Projection methods for solving aggregate growth models. Journal of Economic Theory, 58(2):410-452.
- [48] Judd, K.L. and S.-M. Guu (1997). Asymptotic methods for aggregate growth models. Journal of Economic Dynamics and Control, 21(6):1025-1042.
- [49] Judd, K.L., L. Maliar, and S. Maliar (2011). Numerically stable and accurate stochastic simulation approaches for solving dynamic economic models: Approaches for solving dynamic models. *Quantitative Economics*, 2(2), 173-210.
- [50] Kaplan, G., B. Moll and G.L. Violante (2018). Monetary policy according to HANK. American Economic Review, 108(3):697-743.
- [51] King, R.G., C.I. Plosser and S.T. Rebelo (1988). Production, growth and business cycles: I. The basic neoclassical model. *Journal of Monetary Economics*, 21(2-3):195-232.
- [52] Kirkby, R. (2018). Bewley-Hugget-Aiyagari models: Computation, simulation, and uniqueness of general equilibrium. *Macroeconomic Dynamic*, 23(6):1-40.
- [53] Koop, G., M.H. Pesaran and S.M. Potter (1996). Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics*, 74(1):119-147.
- [54] Krusell, P. and A.A. Smith (1997). Income and wealth heterogeneity, portfolio choice, and equilibrium asset returns. *Macroeconomic Dynamics* 1(2):387-422.
- [55] Krusell, P. and A.A. Smith (1998). Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy*, **106**(5):867-896.
- [56] Kydland, F.E. and E.C. Prescott (1982). Time to build and aggregate fluctuations. *Econometrica*, **50**(6):1345-1370.
- [57] Lasry, J.M. and P.-L. Lions (2007). Mean field games. Japanese Journal of Mathematics, 2(2):229-260.
- [58] Lee (2024). Dynamically consistent global nonlinear solution method in the sequence space and applications, *Working Paper*, University of Cambridge.
- [59] Maliar, S., L. Maliar, and K. Judd (2011). Solving the multi-country real business cycle model using ergodic set methods. *Journal of Economic Dynamics and Control*, 35(2):207-228.
- [60] McFadden, D. (1989). A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica*, **57**(5):995–1026.
- [61] Oskolkov, A. (2023). Heterogeneous impact of the global financial cycle. *Working Paper*, Department of Economics, University of Chicago.

- [62] Parra-Alvarez, J.C. (2018). A comparison of numerical methods for the solution of continuous-time DSGE models. *Macroeconomic Dynamics*, 22(6):1555-1583.
- [63] Phelan, G. (2016). Financial intermediation, leverage, and macroeconomic instability. American Economic Journal: Macroeconomics, 8(4):199-224.
- [64] Rendahl, P. (2022). Continuous versus discrete time: Some computational insights. Journal of Economic Dynamics & Control, 144(November 104522):1-16.
- [65] Rudebusch, G.D. and Swanson, E.T. (2012). The bond premium in a DSGE Model with long-run real and nominal risks. *American Economic Journal: Macroeconomics*, 4(1):1-43.
- [66] Ruge-Murcia, F. (2007). Methods to estimate dynamic stochastic general equilibrium models. Journal of Economic Dynamics and Control, 31(8):2599-2636.
- [67] Ruge-Murcia, F. (2012). Estimating nonlinear DSGE models by the simulated method of moments: with an application to business cycles. *Journal of Economic Dynamics* and Control, 35(6):914-938.
- [68] Ruge-Murcia, F. (2013). Generalized method of moments estimation of DSGE Models. in Hashimzade, N. and Thornton, M., (eds), *Handbook of Research Methods and Appli*cations in Empirical Macroeconomics, Northampton, MA: Edward Elgar Publishing.
- [69] Schmitt-Grohé, S. and M. Uribe (2004). Solving dynamic general equilibrium models using a second-order approximation to the policy function. *Journal of Economic Dynamics & Control*, 28(4):755-775.
- [70] Smith, J.A.A. (1993). Estimating nonlinear time-series models using simulated vector autoregressions. Journal of Applied Econometrics, 8(S1):S63-S84.
- [71] Stokey, N.L. (2009). The Economics of Inaction: Stochastic Control Models with Fixed Costs. Princeton University Press, New York.
- [72] Taylor, J.B. and H. Uhlig (1990). Solving nonlinear stochastic growth models: A comparison of alternative solution methods. *Journal of Business & Economic Statistics*, 8(1):1-17.
- [73] Winschel, V. and M. Krätzig (2010). Solving, estimating, and selecting nonlinear dynamic models without the curse of dimensionality. *Econometrica*, **78**(2):803-821.
- [74] Zhu, Y.-L., X. Wu, I-L. Chern and Z.-Z. Sun (2013). Derivative Securities and Difference Methods. Second Edition, Springer, Berlin.

Appendix. The details of log-linearization

In the following detailed derivation of the equations in Section 3.1, we use the general formula of the log-linearization as follows: $\widetilde{X}_t \equiv \ln X_t - \ln X \approx (X_t - X)/X$, so that $X_t = Xe^{\widetilde{X}_t}$, $X_t \approx X\left(1 + \widetilde{X}_t\right)$, and $e^{\widetilde{X}_t} \approx 1 + \widetilde{X}_t$. Also, $\ln(1 + X_t) \approx X_t$ and $d\ln X_t = \ln X_t - \ln X = X_t - X/X = dX_t/X \approx \widetilde{X}_t$. Next, we derive the equations listed in Section 3.1.

- 1. The resource constraint can written as follows: $Y_t = C_t + I_t + G$ and $Y\left(1 + \widetilde{Y}_t\right) = C\left(1 + \widetilde{C}_t\right) + I\left(1 + \widetilde{I}_t\right) + G$. At the steady state, we have that Y = C + I + G. Thus, $Y\widetilde{Y}_t = C\widetilde{C}_t + I\widetilde{I}_t$ and $\widetilde{Y}_t = C\widetilde{C}_t/Y + I\widetilde{I}/Y$.
- 2. The consumption Euler equations become to the following equations

$$C_{t} = E_{t} \left\{ \left[(1+r_{t}^{n}) \frac{P_{t}}{P_{t+1}} \beta \right]^{-\sigma} C_{t+1} \right\}, \quad Ce^{\widetilde{C_{t}}} = E_{t} \left\{ \left[e^{\widehat{1+r_{t}^{n}}} \times \frac{e^{\widetilde{P_{t}}}}{e^{\widetilde{P_{t+1}}}} \times \beta \right]^{-\sigma} Ce^{\widetilde{C_{t+1}}} \right\},$$

$$1 + \widetilde{C}_{t} = E_{t} \left\{ \left[\left(1 + \widehat{1+r_{t}^{n}} \right) \left(1 + \widetilde{P_{t}} - \widetilde{P_{t+1}} \right) \beta \right]^{-\sigma} \left(1 + \widetilde{C_{t+1}} \right) \right\},$$

$$\ln \left(1 + \widetilde{C}_{t} \right) = E_{t} \left\{ -\sigma \left[\ln \left(1 + \widehat{1+r_{t}^{n}} \right) + \ln \left(1 - \pi_{t+1} \right) + \ln \beta \right] + \ln \left(1 + \widetilde{C_{t+1}} \right) \right\},$$
and
$$\widetilde{C}_{t} = E_{t} \left\{ -\sigma \left(\widetilde{1+r_{t}^{n}} - \pi_{t+1} - \rho \right) + \widetilde{C_{t+1}} \right\}.$$

$$\widetilde{C}_t = E_t \left\{ -\sigma \left(1 + \overline{r_t^n} - \pi_{t+1} - \rho \right) + \widetilde{C_{t+1}} \right\},$$

where $\pi_t = \widetilde{P}_t - \widetilde{P}_{t-1}$ and $\ln \beta = -\rho$. Since $1 + r_t^n = e^{\widetilde{1+r_t^n}}$, then, $\ln(1+r_t^n) = \widetilde{1+r_t^n}$, so that $r_t^n \approx \widetilde{1+r_t^n}$ and

$$\widetilde{C}_t = E_t \left\{ -\sigma \left(r_t^n - \pi_{t+1} - \rho \right) + \widetilde{C_{t+1}} \right\}$$

3. The link between asset prices and investment is given by

$$\frac{I_t}{K_t} = \delta + \frac{1}{c} \left(1 - \frac{1}{Q_t} \right), \quad \frac{Ie^{\widetilde{I}_t}}{Ke^{\widetilde{K}_t}} = \delta + \frac{1}{c} \left(1 - \frac{1}{Qe^{\widetilde{Q}_t}} \right),$$
$$IQe^{\widetilde{I}_t}e^{\widetilde{Q}_t} = \delta QKe^{\widetilde{Q}_t}e^{\widetilde{K}_t} + \frac{1}{c} \left(QKe^{\widetilde{Q}_t}e^{\widetilde{K}_t} - Ke^{\widetilde{K}_t} \right),$$

and

$$IQ\left(1+\widetilde{I}_{t}\right)\left(1+\widetilde{Q}_{t}\right) = \delta QK\left(1+\widetilde{Q}_{t}\right)\left(1+\widetilde{K}_{t}\right) + \frac{1}{c}\left[QK\left(1+\widetilde{Q}_{t}\right)\left(1+\widetilde{K}_{t}\right) - K\left(1+\widetilde{K}_{t}\right)\right]$$

At the steady state, we have

$$\frac{I}{K} = \delta + \frac{1}{c} \left(1 - \frac{1}{Q} \right).$$

Therefore,

$$IQ\widetilde{I}_t - \frac{1}{c}K\widetilde{Q}_t = IQ\widetilde{K}_t \text{ and } \widetilde{I}_t - \widetilde{K}_t = \frac{1}{c}\frac{K}{IQ}\widetilde{Q}_t.$$

In the equilibrium, we have

$$\frac{I}{K} - \delta = 0,$$

which also implies Q = 1 with the steady state condition. Thus,

$$\widetilde{I}_t - \widetilde{K}_t = \frac{1}{c\delta}\widetilde{Q}_t.$$

4. Marginal cost of funds and marginal return to capital are given by

$$E_t \left\{ \Lambda_{t,+1} \left(1 + r_t^n \right) \frac{P_t}{P_{t+1}} \right\} = E_t \left\{ \Lambda_{t,+1} \left(\frac{Z_t + (1-\delta)Q_{t+1}}{Q_t} \right) \right\}$$

and

$$\ln(1+r_t^n) + \ln P_t - \ln P_{t+1} = \ln[Z_t + (1-\delta)Q_{t+1}] - \ln Q_t$$

Take total differential to obtain the following

$$d\ln(1+r_t^n) + d\ln P_t - d\ln P_{t+1} = d\ln[Z_t + (1-\delta)Q_{t+1}] - d\ln Q_t$$

and

$$dr_t^n + \widetilde{P}_t - \widetilde{P_{t+1}} = \frac{[dZ_t + (1-\delta)dQ_{t+1}]}{Z + (1-\delta)} - \widetilde{Q}_t r_t^n - \pi_{t+1} - \rho = (1-\tau)\widetilde{Z}_t + \tau \widetilde{Q_{t+1}} - \widetilde{Q}_t.$$

Since $dr_t^n = r_t^n - r = \tilde{r_t^n}, Z_t = \alpha Y_t / [(1 + \mu_t)K_t]$, and at the steady state $r = \rho$, it easy to see that

$$r_t^n - \rho - \pi_{t+1} = \frac{Z}{Z + (1-\delta)}\widetilde{Z}_t + \frac{1-\delta}{Z + (1-\delta)}\widetilde{Q_{t+1}} - \widetilde{Q_t}$$

Now, define $\tau = (1-\delta) \left[\frac{\alpha Y}{(1+\mu)K} + (1-\delta) \right]^{-1} = (1-\delta)/[Z+(1-\delta)]$. Then, $1-\tau = Z/[Z+(1-\delta)]$. Thus, $r_t^n - \pi_{t+1} - \rho = (1-\tau)\widetilde{Z}_t + \tau \widetilde{Q}_{t+1} - \widetilde{Q}_t$. Since $\widetilde{Z}_t = d\ln Z_t = d\ln \left(\frac{1}{1+\mu_t}\alpha \frac{Y_t}{K_t}\right) = -d\ln(1+\mu_t) + d\ln Y_t - d\ln K_t = -\widehat{\mu}_t + \widetilde{Y}_t - \widetilde{K}_t$, and $d\ln(1+\mu_t) = d\mu_t = \widehat{\mu}_t$, in particular, we arrive at

$$r_t^n - \pi_{t+1} - \rho = (1 - \tau) \left(-\widehat{\mu_t} + \widetilde{Y}_t - \widetilde{K}_t \right) + \tau \widetilde{Q_{t+1}} - \widetilde{Q_t}$$

5. Production function is $Y_t = A_t K_t^{\alpha} L_t^{1-\alpha} V_t$, so that $\ln Y_t = \ln A_t + \alpha \ln K_t + (1-\alpha) \ln L_t + \ln V_t$. Taking total differential leads to

$$\frac{dY_t}{Y} = \frac{dA_t}{A} + \alpha \frac{dK_t}{K} + (1-\alpha)\frac{dL_t}{L}.$$

Since $V_t = \left[\int_0^1 \left(\frac{P_t(f)}{P_t}\right)^{-\varepsilon} df\right]^{-1}$, it vanishes in a first-order log linearization around a zero inflation steady state due to the reason that the deviation of $\ln (P_t(f)/P_t)$ must average to zero. Thus,

$$\widetilde{Y}_t = \widetilde{A}_t + \alpha \widetilde{K}_t + (1 - \alpha) \widetilde{L}_t.$$

- 6. Labor market equilibrium is $(1 \alpha)\frac{Y_t}{L_t} = (1 + \mu_t)\frac{L_t^{\varphi}}{C_t^{-\gamma}}$, then, $\ln(1 \alpha) + \ln Y_t \ln L_t = \ln(1 + \mu_t) + \varphi \ln L_t + \gamma \ln C_t$. By taking total differential, one obtains that $\widetilde{Y}_t \widetilde{L}_t = \widehat{\mu}_t + \varphi \widetilde{L}_t + \gamma \widetilde{C}_t$.
- 7. Price adjustment and Phillips curve can be simplified as follows. It follows from $P_t = \left[\theta \left(P_{t-1}\right)^{1-\varepsilon} + (1-\theta) \left(P_t^o\right)^{1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}}$ that $\widetilde{P}_t = \theta \widetilde{P}_{t-1} + (1-\theta) \widetilde{P}_t^o$, which can be transformed into $\pi_t = \widetilde{P}_t \widetilde{P}_{t-1} = (1-\theta) \left(\widetilde{P}_t^o \widetilde{P}_t\right)/\theta$. Also, it is easy to see from $E_t \sum_{i=0}^{\infty} \theta^i \Lambda_{t,i} \left(P_t^o/P_{t+i}\right)^{-\varepsilon} Y_{t+i} \left[P_t^o/P_{t+i} (1+\mu)/(1+\mu_{t+i})\right] = 0$ that

$$\widetilde{P_t^o} = (1 - \theta\beta)E_t \sum_{i=0}^{\infty} (\theta\beta)^i \left(\widetilde{\mathrm{MC}_{t+i}} + \widetilde{P_{t+i}}\right) = (1 - \theta\beta) \left(\widetilde{\mathrm{MC}_t} + \widetilde{P_t}\right) + \theta\beta E_t \left\{\widetilde{P_{t+1}^o}\right\}.$$

Therefore,

$$\widetilde{P_t^o} - \widetilde{P}_t = (1 - \theta\beta)\widetilde{\mathrm{MC}}_t + \theta\beta E_t \left\{ \widetilde{P_{t+1}^o} - \widetilde{P_{t+1}} + \widetilde{P_{t+1}} - \widetilde{P}_t \right\}$$

Thus,

$$\frac{\theta}{1-\theta}\pi_t = (1-\theta\beta)\widetilde{\mathrm{MC}}_t + \theta\beta E_t \left\{\frac{\theta}{1-\theta}\pi_{t+1} + \pi_{t+1}\right\}$$

and

$$\pi_t = \frac{(1-\theta)(1-\theta\beta)}{\theta} \widetilde{\mathrm{MC}}_t + \beta E_t \{\pi_{t+1}\} = -\lambda \widehat{\mu}_t + \beta E_t \{\pi_{t+1}\},$$

where $\lambda = (1 - \theta)(1 - \theta\beta)/\theta$.

8. Evolution of capital is approximated by

$$K_{t+1} = I_t - \frac{1}{2}c\left(\frac{I_t}{K_t} - \delta\right)^2 K_t + (1 - \delta)K_t,$$
$$K\widetilde{e^{K_{t+1}}} = Ie^{\widetilde{I}_t} - \frac{1}{2}c\left(\frac{I}{K}e^{\widetilde{I}_t\widetilde{K}_t} - \delta\right)^2 Ke^{\widetilde{K}_t} + (1 - \delta)Ke^{\widetilde{K}_t}$$

and

$$K\left(1+\widetilde{K_{t+1}}\right) = I\left(1+\widetilde{I_t}\right) - \frac{1}{2}c\left[\frac{I}{K}\left(1+\widetilde{I_t}-\widetilde{K_t}\right) - \delta\right]^2 K\left(1+\widetilde{K_t}\right) + (1-\delta)K\left(1+\widetilde{K_t}\right).$$

At the steady state, the fact that $K = I - \frac{1}{2}c\left(\frac{I}{K} - \delta\right)^2 K + (1 - \delta)K$ implies that $K\widetilde{K}\widetilde{K}_{t+1} = I\widetilde{I}_t - cI\left(\frac{I}{K} - \delta\right)\left(\widetilde{I}_t - \widetilde{K}_t\right) + [K - I - (1 - \delta)K]\widetilde{K}_t + (1 - \delta)K\widetilde{K}_t$ and

$$\widetilde{K_{t+1}} - \widetilde{K_t} = \left[\frac{I}{K} - c\frac{I}{K}\left(\frac{I}{K} - \delta\right)\right]\left(\widetilde{I_t} - \widetilde{K_t}\right).$$

In the equilibrium, $I/K = \delta$, so that $\widetilde{K_{t+1}} = \delta \widetilde{I}_t + (1-\delta)\widetilde{K}_t$.

9. Interest rate rule is given by $1 + r_t^n = (1+r) \left(\frac{P_t}{P_{t-1}}\right)^{\phi_{\pi}} \left(\frac{Y_t}{Y_t^*}\right)^{\phi_y} e^{v_t}$. Then, $\ln(1+r_t^n) = \ln(1+r) + \phi_{\pi} \left(\ln P_t - \ln P_{t-1}\right) + \phi_y \left(\ln Y_t - \ln Y_t^*\right) + v_t$, which yields that $r_t^n = r + \phi_{\pi} \left(\widetilde{P}_t - \widetilde{P}_{t-1}\right) + \phi_y \left(\widetilde{Y}_t - \widetilde{Y}_t^*\right) + v_t$. Since in the equilibrium, $r = \rho$, thus, $r_t^n = \rho + \phi_{\pi} \pi_t + \phi_y \left(\widetilde{Y}_t - \widetilde{Y}_t^*\right) + v_t$.