Master thesis

Master Mathématiques de la Modélisation (LJLL)

UPMC – Sorbonne Université, Paris

Intern at Laboratory Jacques Louis Lions (LJLL)

Université Paris Diderot, Paris

Wealth distribution over the business cycle
A mean-field game with common noise

Student :

Thomas BOURANY[a]

Dates :

April 2018 – August 2018

Supervisor :

Yves ACHDOU

Professor (LJLL)

[a] *thomas.bourany@sciencespo.fr*, Website

Wealth distribution over the business cycle
A mean-field game with common noise

Thomas Bourany*

Master Mathématiques de la Modélisation (LJLL)

UPMC – Sorbonne Université, Paris

August 2018

Abstract

The standard economic framework with "heterogeneous agents" – the Aiyagari-Bewley model – has recently been reformulated as an example of Mean Field Game (MFG), by Achdou, Han, Lasry, Lions, and Moll (2017). One of the key question in such model is to understand the transmission of aggregate shocks – on macroeconomic dynamics or the shape of the wealth distribution. With aggregate risk, this framework can thus be understood as a MFG with "common noise". However, the resolution of such model is notoriously difficult, due to the "curse of dimensionality" arising when common noise interact with both the behavior and the distribution of agents. Economists usually simplify the model considering a finite set of moments of the measure (bounded-rationality à la Krusell-Smith) or using Projection and Perturbation methods (à la Reiter). In contrast, we use new methods to keep the full dimensionality and simulate the model using a discretization procedure for the common noise. Considering a tree structure or (optimal) quantization to represent the trajectories of the common noise with a finite number of shocks, we solve the MFG system using specific finite-differences methods for the two PDEs.

We apply this method to the standard framework, and two extensions (i) with Endogenous Labor Supply (ii) One Asset HANK model and we provide intuitions for (iii) the two Assets H.A. model (à la Kaplan-Moll-Violante). We show that such method might be relevant to analyze the transmission of large shocks on the economy.

*I would like to thank my supervisor Yves Achdou for his guidance and helpful comments all throughout this work. My gratitude goes also to Pierre Cardaliaguet and Xavier Ragot for their respective lectures on Mean Field Games and Heterogeneous agents models that introduced me to this field at the frontier between the two disciplines. I would also like to thank my classmates and the PhD students of LJLL and LPSM at Paris-Diderot and (UPMC)-Sorbonne University and at Sciences Po Paris. In particular, I thank Thibaut Montes for his introduction to optimal quantization and Julien Pascal and Ziad Kobeissi for interesting discussions about Mean Field Games.
All errors are mine.

1 Introduction

One of the recent development of macroeconomics has been to integrate agent heterogeneity in credible fashion to study the income and wealth distributions of households. Since the contribution of Bewley (1986) and Aiyagari (1994), this "heterogenous agents" literature could provide relevant answers to numerous economic questions: the causes and dynamics of inequality, the implication of market frictions to generate skewed wealth distributions or the distributional consequences of monetary or fiscal policies[1]. One aspect of this literature investigates the implication of aggregate shocks for income and wealth distribution. The present master thesis follows this tradition by providing new methods to solve such type of model using tools from the Mean Field Games literature.

If the first treatment of this topic goes back to Krusell and Smith (1998), some aspects are still subject to a lack of understanding. The main result of Krusell-Smith was to show that, in incomplete market with both uninsurable idiosyncratic risk and aggregate risk, the model features "approximate aggregation" in the behavior of agents: the macroeconomic aggregates can be almost perfectly described using only the mean of the wealth distribution, and the control of agent is almost independent of its state. Therefore, one is naturally inclined to wonder whether agent heterogeneity substantially changes the results at stake in "representative agent" models. This paper investigate these different questions.

In this project, we explore the implication of aggregate uncertainty – or "common noise" – in a standard Aiyagari-Bewley framework, as in Krusell-Smith model. As most heterogeneous agents model in economics, this framework is an interesting example of Mean Field Games (MFG)[2]. Developed simultaneously by Lasry and Lions (2007) and Huang, Malhamé, and Caines (2006), interested readers can find a very pedagogical and comprehensive treatment in the lecture notes Cardaliaguet (2018). A Mean-Field Game can be described a game with a large number of "small" symmetric players: the interaction between them is only reflected by the interaction between each agent and the "distribution" – i.e. the measure m_t – of the other agents:

$$\sup_{\alpha} \mathbb{E}\left[\int_0^{\infty} e^{-\rho t} f(t, X_t, m_t, \alpha_t) dt\right]$$
$$dX_t = b(t, X_t, m_t, \alpha_t)dt + \sigma(t, X_t, m_t)dB_t + \sigma^0(t, X_t, m_t)dW_t^0$$

where W_t and B_t^0 refers respectively to the idiosyncratic noise and the common noise.

[1] One could mention for instance Gabaix, Lasry, Lions, and Moll (2016) for dynamics of inequality, Kaplan and Violante (2014) for consumption and redistributive effects of fiscal stimulus, Heathcote, Storesletten, and Violante (2014) for the influence of incomplete insurance on risk-sharing, consumption and labor supply, Benhabib, Bisin, and Zhu (2011) on how both labor and capital idiosyncratic risk can generated Pareto wealth distribution and Bhandari, Evans, Golosov, and Sargent (2017) on optimal monetary and fiscal policy in incomplete markets

[2] A summary on how to obtain the standard MFG system is recalled in section A of appendix

This stochastic control problem boils down to a system of coupled partial differential equations:

- First, a Hamilton-Jacobi-Bellman (HJB) equation – where the agents make their choices $\alpha^\star$ taking the measure m_t of agents as given
- Second, a Fokker-Plank (FP, or Kolmogorov Forward) equation – where the measure m_t evolves, given optimal choices and the controlled process $X^{\alpha^\star}$.

The standard Aiyagari-Bewley model has recently been reformulated as an example of Mean Field Game (MFG), by Achdou, Han, Lasry, Lions, and Moll (2017). Economically, the problem is the following: households solve a typical control problem – consumption-saving problem – and are subject to idiosyncratic risk on their income. In an incomplete market and unable to hedge this risk, the agents use the only asset at their disposal – saving in capital – to self-insure. This capital stock is also used as an input factor by firms and 'priced' competitively at an interest rate. However, there is a limited amount households can borrow, and this 'credit constraint' forces the agents to (over)- self-insure: that what we call "precautionary saving".

This model can be simulated by solving numerically this system of two equations, using a finite-difference scheme we will describe thoroughly in the following. The two main particularities of the model are (i) a credit constraint that acts as a 'state-constraint', imposing a specific treatment of the PDE on the boundaries and (ii) the mean-field interaction through the interest rate acting as a coupling between the two PDEs that requires the search for an equilibrium fixed point.

When introducing aggregate uncertainty, this model is known as a Mean Field Game with "common noise", i.e. common source of random shocks that applies to the MFG system as the whole[3]. The treatment of such common noise is as old as the Krusell and Smith (1998) article but impose challenging difficulties in mathematics.

When common noise interacts with both the behavior and the distribution of agents, solving the model suffers from what is sometimes called "curse of dimensionality" by economists of the field. As explained in the latest book by Carmona and Delarue (2018): "*In order to account for the dependence of the equilibria upon the realization of the common noise, it is necessary to enlarge the space in which the fixed point has to be sought.*" The control problem become infinite-dimensional and solving this model has appeared to be notoriously difficult.

The present master thesis offers a new numerical method to deal with such issue and to simulate the MFG system with common noise. The main idea is the discretize the aggregate shocks. Following the approach of Y. Achdou[4], we approximate the common noise – that can

[3]Here we will use indifferently the term "common noise" and "aggregate risk/shocks/uncertainty"
[4]This method has originally been used for models of crowd motion with congestion in movie theater

be a two-state jump process or a Brownian motion – as a finite number of shocks using a tree structure. Between each shock – i.e. on each branch of the tree – the MFG will be a standard deterministic system of two evolution PDEs, the HJB and the FP coupled by the interest rate. The main challenge will be to link the behavior of the MFG in function of the anticipation of future aggregate shocks and in function of the evolution of past state variables (i.e. saving). A way to include this is to (i) change the terminal condition for the HJB: the final value function will now be the expectation of the future value functions over the different paths of the common noise and (ii) define the initial condition of the FP with the past wealth distribution before the realization of the common noise. With these two methods, we can compute the evolution of the MFG for different trajectories of aggregate productivity – i.e. the common noise in our setting.

This new method has several comparative advantages, even if the treatment of problem of MFG with common noise has a long history discussed in two separate literatures.

First, the literature of computational economics have searched to reduce to dimensionality of the problem: Originally, Krusell and Smith (1998) approximated the wealth distribution using finite number of moments and solving the control problem using this finite/bounded-rational view of the economy. Later, Reiter (2010) used projection methods to approximate the measure of agents by simpler objects and implemented linearization and perturbation methods to simulate the economy for (necessarily small) aggregate shocks. With advances by Algan, Allais, and Den Haan (2008) and Winberry (2016) who use parametric families to approximate the measure, the combinaison of these methods have been relatively successful since the recent article by Ahn, Kaplan, Moll, Winberry, and Wolf (2018). One need to note that truncation methods have also been developed in economics, in particular by X. Ragot. These methods offer parsimonious and intuitive views of these complex models, and can be simulated using Dynare. See for example Challe, Matheron, Ragot, and Rubio-Ramirez (2016) or the recent handbook chapter Ragot (2018).

On a different world, in mathematics, the goal has rather been to reformulate the problem in infinite-dimension. A first approach using the *master equation* was introduced by P. L. Lions in his lectures at the *Collège de France*. This amounts to define a PDE on the state-space *times* the space of probability measures on the state and relying on differential calculus on the Wasserstein space. As such, it allows to deal with problem without or with common noise by treating the master equation of first or second order. This method, developed extensively in Cardaliaguet, Delarue, Lasry, and Lions (2017) proves to be very powerful.

Another approach by Carmona, Delarue, and Lacker (2016) would be to discretize the common noise in order to treat the problem with a finite space. The idea, building around the notion of weak MFG solution can be summarized as follow: First of all, without common noise, the standard fixed-point procedure used in MFG consists in (i) fixing a measure m_t, (ii) solving the optimal control problem $\alpha_t^\star$, (iii) finding the corresponding probability law of the controlled process $X_t^{\alpha^\star}$, (iv) iterating to find a fixed point $m_t = \mathcal{L}(X_t)$. With the introduction of the common noise, the measure considered is now a measure *conditional* on the common noise $\mathcal{L}(X_t|B_t^0)$. However, conditioning introduces a major technical difficulty in the procedure above, and the discretization is a way to bypass such complication. When refining the discretization, the authors can recover the solution at the limit only in the *weak* sense: the solution of the weak MFG m_t may not be adapted to the common noise and the fixed point condition need to be weaken. However, under the assumption of pathwise uniqueness, existence of strong solution can be recovered. In section 3.1, we explain the details of these approaches, and a brief introduction to the master equation, with a tentative to adapt these results to the Krusell-Smith model.

This project and the numerical scheme associated follows this last approach, discretizing the common noise and refining the discretization to approximate the solution in the weak sense. This method has the first and main advantage to preserve the non-linearity of the system. This bypasses the traditional use of linearization techniques when one uses perturbation methods. Moreover, the second strength of our procedure is to be unambiguous on the anticipations of agents. With our discretization of the common noise, we compute effectively the different trajectories of aggregate risk and measure how agents expect and react to these future shocks. Even with a rough discretization, we can have a precise view of how agents would hedge aggregate risk and can proved relevant in future research in mathematical finance.

However, one of the main challenge of this approach is when the discretization procedure is made finer to approximate the Brownian motion/Jump process of the common noise. The number of trajectories will grow exponentially and thus the computation time. A goal for further research will be to develop novel techniques to be able to reach a greater accuracy of the MFG with common noise. We offer novel ideas, combining recent techniques of optimal quantization and the resolution of the Forward-Backward system or using the treatment of the Master equation with projection methods. Even though these methods are not yet implemented, the procedures are explained in the section 3.3 and deserves further research.

To prove how this method can be relevant for MFG and heterogeneous agents models (H.A.) in economics, we first describe our algorithm in section 3.2 and apply this method to the main frameworks of the recent economics literature in section 4. First, I apply this model to the standard setting of Krusell-Smith. In this model, developed in detail in section 4.1, the agents simply choose consumption & saving in presence of credit constraint and both idiosyncratic and aggregate risk. When solving the model, we study the importance of aggregate shocks on the shape of the wealth distribution and how precautionary saving can change the reaction of households and the dynamics of aggregate variables.

We provide a natural extension when labor supply in endogenous as discussed in section 4.2. This model can be the analogous H.A. model of the Real Business Cycle framework – one of the traditional building block of macroeconomics – and we study how labor supply with precautionary saving can change business cycle dynamics.

We cover a last model in section 4.3, where we consider a version of the one-asset HANK model – i.e. Heterogeneous agents New Keynesian model– as discussed in B. Moll's lecture notes and in the PHACT project. The interesting feature of this model is that – in absence of capital – precautionary motives interacts with rigid prices, creating an amplification channel of productivity shocks.

We end this paper by providing some insights for the two-assets model as developed in Kaplan and Violante (2014) and Kaplan, Moll, and Violante (2018). Moreover, to explain how the prediction of these models differ quantitatively from the "representative agent" counterpart, we provide in section 4 a comparison with these standard macro model, namely the Brock-Mirman model, the RBC model and the standard New Keynesian model.

Before developing the mathematical setting in section 3.1, we now introduce and describe insection 2 how the four Heterogeneous Agents models are good example of Mean Field Games.

2 Heterogeneous agents models: examples of MFGs

2.1 The Aiyagari-Bewley model

In one of its simplest formulation, the model shows the economy, composed of a continuum of households. They face idiosyncratic and uninsurable income shocks, and are subject to a credit (borrowing) constraint. They thus solve the following stochastic control problem:

$$\begin{aligned}
&\sup_{c_t} \mathbb{E}_0 \int_0^\infty e^{-\rho t} u(c_t) dt \\
\text{subject to :} \quad & da_t = (z_t w_t + r_t a_t - c_t) dt && \text{(Budget constraint)} \\
\text{and} \quad & a_t \geq \underline{a} && \text{(Credit constraint)}
\end{aligned}$$

where a is wealth – the state variable we'll focus on – c_t consumption, ρ rate of time preference, $u(\cdot)$ a utility function, supposed increasing and concave ($u' > 0, u'' < 0$). For the simulation, the utility will be chosen to adopt a specific functional form : Constant Relative Risk Aversion (CRRA) utility, i.e. $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$.

The income is composed of a wage w_t and productivity factor z_t. The wage w_t, and interest rate r_t, will adjust in general equilibrium, considering the firm's side (cf. next subsection). The productivity z_t is subject to shocks that are (i) idiosyncratic (i.e. its law is drawn i.i.d.[5]) and (ii) uninsurable, (agents can not hedge/cover against this risk). z intuitively represents the state of the agent, for instance, employment $z_{high} = z_2$ and unemployment $z_{low} = z_1 = \frac{1}{2} z_2$. In this presentation, it is modeled simply as a jump processes with two states $\{z_1, z_2\}$ (with intensities λ_1, λ_2, the higher the intensity, the higher the proba to jump). However, one can generalize it any stochastic process: e.g. diffusion $dz_t = b(z_t)dt + \sigma(z_t)dB_t$, Poisson or Levy processes. The only numerical constraint is that it has a bounded domain[6].

One of the specificities of this problem is the credit constraint, which is a *state* constraints: $a_t \geq \underline{a}$. State constraint is a complicated problem for control theory. In our situation, intuitively, the optimal strategy of the agent might be (and will be) to move on the constraint ($\partial\Omega$) and stay there (poverty trap). Mathematically, it is not possible to find a PDE and a boundary condition on $\partial\Omega$ even in the sense of distribution.

According to Soner (1986), the mapping from the state to the set of admissible controls $a \mapsto \mathcal{A}_a$(and $c \in \mathcal{A}_a$) 'will have a complicated structure', and its regularity may not be insured in general. In our situation it will be the case, but it will implicitly impose a constraint on the derivative of the value function at the boundary.

[5]But the etymology implies that agents might react in their own way to a similar shock

[6]In case of diffusion, it will be reflected after a limit

This will result on both (i) a Dirac mass on the boundary and (ii) an explosion near the boundary. Economically, there is a lower bound on the value of this borrowing limit: $\underline{a} \geq -z_1 w_t / r_t$. The latter term represents the natural borrowing limit – the credit that an agent could repay if it would fall permanently in the low income state (and repaying its debt at rate r_t.

2.1.1 The supply side – a neoclassical firm

The supply side is driven by a representative firm, produce goods following a production function of aggregate productivity, capital and labor $F(A, K, L)$, supposed to be concave in the two last variables. Given δ the depreciation of capital and A productivity level, the firm is producing in a perfectly competitive environment, and thus maximize its profit:

$$\Pi(K_t, L_t) = \sup_{K,L} \; F(A_t, K_t, L_t) - (r_t + \delta)K_t - w_t L_t$$

The price of demanded input factors – i.e. r_t interest rate, w_t wage – will be determined by the First-Order-Condition of this optimization problem:

$$r_t = \partial_K F(A_t, K_t, L_t) - \delta \qquad\qquad w_t = \partial_L F(A_t, K_t, L_t)$$

In this article, we suppose that (i) production function is simply given by the Cobb-Douglas production function: $F(A_t, K, L) = A_t\, K_t^\alpha\, L_t^{1-\alpha}$, and (ii) the effective labor supply is fixed and equal to the average productivity of households: $z_{av} = \frac{z_1\lambda_2 + z_2\lambda_1}{\lambda_1 + \lambda_2}$.

The optimality relation described above reduce to :

$$\begin{aligned} K(r) :&= \left(\frac{\alpha A}{r+\delta}\right)^{\frac{1}{1-\alpha}} z_{av} && \text{Capital demand} \\ w &= (1-\alpha)\, A\, K^\alpha z_{av}^{-\alpha} && \text{Wage} \\ r &= \alpha\, A\, K^{\alpha-1} z_{av}^{1-\alpha} - \delta && \text{Interest rate} \end{aligned}$$

Wage and interest rate are indeed constant when productivity A and all parameters are constant. Let us look at the Mean Field Game formulation[7] of the problem.

2.1.2 A mean field game formulation

The stochastic control problem boils down to a system of two partial differential equation: (i) a Hamilton-Jacobi-Bellman (HJB) equation and (ii) a Fokker-Plank (or Kolmogorov Forward

[7] Cf. appendix for a general setting

– FP). When parameters are constant, the stationary equilibrium system is described by the equations:

$$\rho v_j(a) = \max_c u(c) + \partial_a v_j(a)(z_j w + ra - c) + \lambda_j (v_{-j}(a) - v_j(a)) \qquad \text{[HJB]}$$

$$0 = \frac{d}{da}[s_j(a)\, g_j(a)] + \lambda_j g_j(a) - \lambda_{-j} g_{-j}(a) \qquad \text{[FP]}$$

$$S(r) := \int_{\underline{a}}^{\infty} a\, g_1(a) da + \int_{\underline{a}}^{\infty} a\, g_2(a) da = K(r) \qquad \text{[Market clearing]}$$

where j indicate the z-state of the agent and $-j$ the opposite state.

However, when the system is subject to shocks – under either (i) perfect foresight/deterministic transitions, (ii) MIT shocks i.e. unanticipated (zero probability) shocks or (iii) anticipated common noises – the system dynamics is described by the two following PDEs:

$$\rho\, v_j(t,a) = \partial_t v_j(t,a) + \max_c u(c) + \partial_a v_j(t,a)\, s_j(t,a) + \lambda_j (v_{-j}(t,a) - v_j(t,a)) \qquad \text{[HJB]}$$

$$0 = \partial_t g_j(t,a) + \frac{d}{da}[s_j(t,a)\, g_j(t,a)] + \lambda_j g_j(t,a) - \lambda_{-j} g_{-j}(t,a) \qquad \text{[FP]}$$

$$S(t,r) := \int_{\underline{a}}^{\infty} a\, g_1(t,a) da + \int_{\underline{a}}^{\infty} a\, g_2(t,a) da = K(t,r) \qquad \text{[Market clearing]}$$

$$s_j(t,a) = z_j w_t + r_t\, a - c_j(t,a) \qquad c_j(t,a) = (u')^{-1}(\partial_a v_j(t,a)) \qquad \text{[FOC]}$$

$$v_j(T,\cdot) = v^{\infty} \qquad g_j(0,\cdot) = g^0 \qquad \partial_a v_j(\underline{a}) \geq u'(z^j w + r\underline{a}) \qquad \text{[Boundary conditions]}$$

In the two cases, the third equations describe the market clearing, i.e. the Walrasian adjustment of prices (i.e. interest rate) to equalize supply of saving $S(r)$ and demand for capital $K(r)$. The last equation represents the optimality condition of the control variable c. The HJB actually features an optimization problem, and, since the objective $u(c) - pc$ is concave in c, the optimum is reached for: $u'(c^\star) = p$ and thus $c^\star = (u')^{-1}(p)$.

Here the state-constraint does not show up in the HJB (as could be the case in discrete time) but it appears in the boundary condition. It comes from the optimality of the maximization in the HJB – holding everywhere, and in particular at $\underline{a}$ – and the FOC is given by $u'(c^j(\underline{a})) = \partial_a v^j(\underline{a})$, and the state-constraint affecting the control: $s^j(a) = z_t^j w + r\, a - c^j(a) \geq 0$. Since u is concave, its derivative is decreasing, yielding the boundary condition:
$u'(z^j w + r\underline{a})) \leq u'(c^j(\underline{a})) = \partial_a v^j(\underline{a})$.

2.2 A preview of the extensions

2.2.1 H.A. – Endogenous labor supply (RBC)

We now explore a natural extension of the standard Aiyagari-Bewley model with endogenous labor supply. In the 80's, researchers like Kydland and Prescott (and many others) built on the Brock-Mirman neoclassical growth model to initiate the Real-Business Cycle (RBC) model. This model was particularly successful to reproduce the main features of macroeconomic fluctuations. In the same spirit, the consumption/labor choice with heterogenous agents may imply interesting aggregate dynamics and that is what we explore in this first extension. When adapted to the MFG problem, the control problem is the following:

$$\max_{\{c_t\}_{t_0}^{\infty}, \{\ell_t\}_{t_0}^{\infty}} \mathbb{E}_0 \int_0^{\infty} e^{-\rho t} u(c_t, \ell_t) dt$$

$$\text{subject to :} \qquad da_t = (z_t w_t \ell_t + r_t a_t - c_t)dt \qquad \text{(Budget constraint)}$$

$$\text{and} \qquad a_t \geq \underline{a} \qquad \text{(Credit constraint)}$$

with ℓ is the labor supplied by household to firms. $u(c, \ell)$, the utility function, is supposed increasing and concave ($u' > 0, u'' < 0$) in c and decreasing and concave in ℓ. The specific functional form is chosen to be Constant Relative Risk Aversion (CRRA, γ) utility in consumption and separable in labor, with a constant Frisch elasticity (ϕ) of labor supply:

$$u(c, \ell) = \frac{c^{1-\gamma}}{1-\gamma} - \frac{\ell^{1+\frac{1}{\phi}}}{1+\frac{1}{\phi}}$$

The rest of the model, in particular the firm side remains mostly similar. The main difference lies in the presence of labor in the production function. The pricing of interest rate and wages operates in terms of the Capital/Labor ratio K_t/L_t

$$Y_t = A_t K_t^{\alpha} L_t^{1-\alpha} \qquad \text{Production fct}$$

$$\frac{K_t}{L_t}(r) := \left(\frac{\alpha A_t}{r_t + \delta}\right)^{\frac{1}{1-\alpha}} = \left(\frac{(1-\alpha)A_t}{w_t}\right)^{-\frac{1}{\alpha}} \qquad \text{Capital /Labor ratio}$$

$$w_t = (1-\alpha)\, A_t\, (K_t/L_t)^{\alpha} \qquad r_t = \alpha\, A_t\, (K_t/L_t)^{\alpha-1} - \delta \qquad \text{Wage/Interest}$$

The interaction of endogenous labor supply and credit constraint will have unattended effects for aggregate fluctuation as we will observe in section 4.

2.2.2 H.A. – New Keynesian Model (One Asset HANK)

If the RBC was the most successful macroeconomic model of the 80's-early 90's, the New Keynesian (NK) model soon took over in the late 90's and 2000's. This framework stemmed from menu costs and imperfect competition models developed earlier. As a synthesis of Neoclassical and Keynesian thought, it kept a similar Household side as the RBC model. The main changes concentrated on firm behavior: firms choose their price level in a staggered manner – known as the Calvo-Yun price setting, where not all firms can choose to update their price. A reason for its success was that monetary policy was not neutral anymore and change in the path of interest rate could have an impact on household and firms dynamics. A general equilibrium summary can be found in Woodford (2003) or Galí (2015). When introducing heterogenous agents, the MFG model is the following:

$$\max_{\{c_t\}_{t_0}^{\infty},\{\ell_t\}_{t_0}^{\infty}} \mathbb{E}_0 \int_0^{\infty} e^{-\rho t} u(c_t, \ell_t) dt$$

$$\text{subject to :} \quad db_t = \big((1-\tau) z_t w_t \ell_t + r_t^b b_t + T_t + \Pi_t - c_t\big) dt \qquad \text{(Budget constraint)}$$

$$\text{and} \quad b_t \geq \underline{b} \qquad \text{(Credit constraint)}$$

Most of the variables are identical as the previous models. Here we choose to include government: T_t is a lump-sum government transfer to all households redistributed from tax τ on labor. Firms are owned by households such that firms profits Π_t are transferred to households, proportional to their income level z. The asset of household side is closer to the Huggett model: there is no capital in the simplest form of the NK, but the bond supply B^g is fixed, and (here) can be changed by the issuance of government debt.

We chose to provide a detailed description of the New Keynesian model in the following[8]. The production side features a continuum of firms and each of them is producing a variety/type of intermediate input. Each of them will choose how to price these inputs but will be subject to 'imperfect competition': a final good producer will aggregate these good with the CES[9] technology:

$$Y = \left(\int_0^1 y_j^{\frac{\varepsilon-1}{\varepsilon}} dj \right)^{\frac{\varepsilon}{\varepsilon-1}}$$

[8] This description of the New Keynesian model in continuous time is inspired from B. Moll's PhD lecture 2

[9] CES: Constant elasticity of substitution, here the parameter ε is the degree at which a consumer will be willing to substitute a good for another

The demand (from the final good producer) for a particular good is given by[10]

$$y_j(p_j) = \left(\frac{p_j}{P}\right)^{-\varepsilon} \qquad \text{and} \qquad P = \left(\int_0^1 p_j^{1-\varepsilon} dj\right)^{\frac{1}{1-\varepsilon}}$$

The intermediate goods producers have a linear production[11] and and thus solves the following maximization problem :

$$y_{j,t} = A_t N_{j,t} \qquad \Rightarrow \qquad \max_{p_j} p_j y_j(p_j) - w_t N_{j,t} = \max_p p\left(\frac{p}{P_t}\right)^{-\varepsilon} Y_t - w_t \frac{1}{A_t}\left(\frac{p}{P_t}\right)^{-\varepsilon} Y_t$$

$$p_{j,t} = P_t = \frac{\varepsilon}{\varepsilon - 1}\frac{w_t}{A_t} \qquad \forall j$$

Let us introduce sticky prices, where firms now faces a quadratic adjustment cost $\Theta_t(\frac{\dot{p}}{p})$ (where θ is the degree of price stickiness). They now solve a (deterministic) control problem for the choice of $\dot{p}$:

$$\Pi_t(p) = p\left(\frac{p}{P_t}\right)^{-\varepsilon} Y_t - w_t \frac{1}{A_t}\left(\frac{p}{P_t}\right)^{-\varepsilon} Y_t \qquad \Theta_t\left(\frac{\dot{p}}{p}\right) = \frac{\theta}{2}\left(\frac{\dot{p}}{p}\right)^2 P_t Y_t$$

$$v^f(p_0) = \max_{(p_t)_{t\geq 0}} \int_0^\infty e^{-\int_0^t i_s ds}\left[\Pi_t(p_t) - \Theta_t\left(\frac{\dot{p}_t}{p_t}\right)\right] dt$$

To derive the optimal price setting, one can state the Hamiltonian $\mathcal{H}(p, \dot{p}, \eta)$ and use the Pontryagin Maximum Principle:

$$\mathcal{H}(p_t, \dot{p}_t, \eta_t) = \Pi_t(p_t) - \Theta_t\left(\frac{\dot{p}_t}{p_t}\right) + \eta_t \dot{p}_t$$

$$\dot{p}_t^\star \in \operatorname*{argmin}_{\dot{p}_t} \mathcal{H}(p_t, \dot{p}_t, \eta_t) \qquad \Rightarrow \qquad \eta_t = \theta \frac{1}{p_t}\left(\frac{\dot{p}_t^\star}{p_t}\right) P_t Y_t$$

$$\dot{\eta}_t = -\frac{\partial \mathcal{H}(p_t, \dot{p}_t, \eta_t)}{\partial p_t} = i_t \eta_t \; - \left[(1-\varepsilon)\left(\frac{p}{P_t}\right)^{-\varepsilon} Y_t + \varepsilon \frac{w_t}{A_t}\frac{1}{p_t}\left(\frac{p}{P_t}\right)^{-\varepsilon} Y_t + \frac{\theta}{p_t}\left(\frac{\dot{p}_t}{p_t}\right)^2 P_t Y_t\right]$$

Since the control problem is identical for all firms (thanks to CRS), we use the fact that $p_t = P_t$ and thus, defining inflation $\pi_t = \frac{\dot{P}_t}{P_t}$, we obtain:

$$\theta \; \pi_t Y_t = \eta_t \qquad \Rightarrow \qquad \dot{\eta}_t = \theta \; (\dot{\pi}_t Y_t + \pi_t \dot{Y}_t)$$

$$\dot{\eta}_t = i_t \eta_t - \left[(1-\varepsilon) Y_t + \varepsilon \frac{w_t}{P_t A_t} Y_t + \theta \pi_t^2 Y_t\right]$$

And rearranging – defining marginal cost $m_t = \frac{w_t}{P_t A_t}$ and $\bar{m} = \frac{\varepsilon - 1}{\varepsilon}$ – we obtain the New

[10] The final good producer is behaving competitively and thus solves a maximization problem of the form : $\max_{(y_j)_{j\in[0,1]}} Y - \int_0^1 y_j p_j dj$

[11] With the two assets model, one could have a production similar to the Neoclassical model $Y = AK^\alpha L^{1-\alpha}$. Here, there is no capital, the important thing for NK model is to have constant return to scale to make sure price setting is homogenous for all producers.

Keynesian Phillips Curve, determining the inflation dynamics:

$$\Big(i_t - \pi_t - \frac{\dot{Y}_t}{Y_t}\Big)\pi_t = \frac{\varepsilon}{\theta}(m_t - \bar{m}) + \dot{\pi}_t$$

To close the model, it remains to determine the policy variables[12], i.e. the evolution of interest rate by monetary authorities and public debt dynamics/taxes by fiscal policy.

The central bank determines the path of nominal interest rate i_t, affecting the real return r_t of asset a_t (which corresponds to government bonds here). One can choose the conventional policy rule:

$$r_t := i_t - \pi_t \qquad \text{(Fisher relation)}$$

$$i_t = \rho + \phi_\pi \pi_t \qquad \text{(Taylor rule), with } \phi_\pi > 1$$

The government is subject to a budget constraint yield bond (public debt) dynamics:

$$\dot{B}_t^g + G_t + T_t = \tau_t w_t L_t + r_t B_t^g$$

Again, the debt choice $\dot{B}^g$ or government spending G_t could be given by a policy rule. For simplicity, we choose the policy $\dot{B}_t^g, G_t = 0$ and the lump-sum transfers to households T_t to adjust with the dynamics of interest rate and labor.

With these features concerning firms/central bank/government, added to similar market clearing on bond $B_t^g + B_t^h = 0$ with $B_t^h = \sum_j \int a g_j da$ and $N_t(w) = L_t := \sum_j \int \ell_j g_j da$

2.2.3 HANK – Two-assets model

Another extension, developed by Kaplan and Violante (2014) and Kaplan, Moll, and Violante (2018) consists in a control problem with two assets: one liquid asset b_t (bond/deposit) and one illiquid asset a_t(housing/equity). It has been particularly successful in explaining wealth distribution. Our version of the rest of the problem is similar to the RBC model.

The two-assets H.A. New Keynesian version is an extension developed by Kaplan, Moll, and Violante (2018) but not described here.

[12] Even though one of the main objectives in macroeconomics would be to determine the optimal policy – by posing a control problem on top of the MFG – a frequent short cut is to use ad-hoc policy rules that can approximate optimal policies in 'representative agent' setting

$$\max_{\{c_t\}_{t_0}^{\infty},\{\ell_t\}_{t_0}^{\infty},\{d_t\}_{t_0}^{\infty}} \mathbb{E}_0 \int_0^{\infty} e^{-(\rho+\zeta)t} u(c_t, \ell_t) dt$$

$$\begin{aligned}
\text{subject to :} \quad & db_t = \big((1-\tau)z_t w_t \ell_t + (r_t^b(b) + \zeta)b_t + T - d_t - \chi(d_t, a_t) - c_t\big)dt && \text{(Budget – Liquid)} \\
& da_t = ((r_t^a + \zeta)a_t) + d_t)dt && \text{(Budget – Illiquid)} \\
& b_t \geq \underline{b} && \text{(Credit constraint – Liquid)} \\
\text{and} \quad & a_t \geq \underline{a} = 0 && \text{(Credit constraint – Illiquid asset)} \\
& r_t^b(b) = r_t^b + \mathbb{1}_{\{b<0\}}\kappa && \text{(Discrimination – lending/borrowing)}
\end{aligned}$$

with ρ rate of time preference and ζ the exogenous rate of death (and birth) of new agents, and $u(c, \ell)$ the same utility as above. Note that $r_t^b(b)$ introduces a "soft" borrowing constraint at zero since borrowing rate is higher than saving rate. Moreover, the main difference between liquid and illiquid assets lies in the higher return of illiquid asset[13]. However, this gain is counterbalanced by the transaction cost $\chi(d, a)$ one has to pay to convert liquid wealth into illiquid wealth. This cost is always positive and strictly convex, and its expression have strong economic justifications[14]

$$\chi(d, a) = \chi_0|d| + \frac{\chi_1}{1+\chi_2}\left|\frac{d}{a}\right|^{1+\chi_2}|a|$$

The rest of the model is similar to the RBC-endogenous labor supply.

[13] Illiquid asset – representing equity – allows to gain the firms-profits in the New Keynesian version of this two-assets model

[14] As explained in Kaplan, Moll, and Violante (2018): *This transaction cost has two components: The linear component generates an inaction region in households' optimal deposit policies because for some households the marginal gain from depositing or withdrawing the first dollar is smaller than the marginal cost of transacting $\chi_0 > 0$. The convex component ($\chi_1 > 0, \chi_2 > 0$) ensures that deposit rates are finite, and hence household's holdings of assets never jump. Finally, scaling the convex term by illiquid assets a delivers the desirable property that marginal costs $\chi(d, a)$ are homogeneous of degree zero in the deposit rate d/a so that the marginal cost of transacting depends on the fraction of illiquid assets transacted, rather than the raw size of the transaction.*

3 Introducing common noise / aggregate shocks

We now consider the same framework, when introducing the common noise – or aggregate shock to use to economics idiom. This is of particular interest for economists to understand how the economy 'reacts' to exogenous shocks affecting the system as a whole. In the Krusell-Smith model and its extensions, we consider the level of productivity A_t (TFP: total factor productivity) to be subject to random fluctuations. This affects firms in their production process, and thus prices of the factors (interest rate and wages) and in turn the behavior of households.

What stochastic process will follow the aggregate productivity? The form of the process will matter a lot for the anticipations of agents (who are perfectly forward looking and thus draw expectations about the future of the economy). The simplest way would be to consider a simple Brownian motion dB_t starting at a initial level $\bar{A}_0$ and stopped at a deterministic stopping time T (assuming to be the stationary equilibrium). This will be the situation considered in the following mathematical analysis and the first simulation.

However, we also simulate the model using a Jump process with two states, meant to represent the state of economic "booms" (growth) or "busts" (recession). This process was indeed the option in the original article by Krusell-Smith, where both idiosyncratic and aggregate risks were modeled by Markov Chains. Moreover, to study the transmission mechanisms after a transitory shock, economists often consider stationary processes as $AR(1)$ processes ($X_{t+1} = \mu + \varphi X_t + \sigma\varepsilon_t$). The closest process in continuous time would be the Ornstein-Uhlenbeck diffusion process, that mean-reverts to the value μ: $dX_t = -\theta(X_t - \mu)dt + \sigma dB_t$, or the "Jump-Drift Process". $dX_t = -\theta(X_t - \mu)dt + \varepsilon dN_t$ where dN_t is a jump process ($dN = 1$ with intensity λ) and $\varepsilon \sim \mathcal{N}(0, \sigma^2)$. This latest is the one we consider for the plot of the Impulse Response Functions (IRF) in the results in section 4.

3.1 General mathematical framework

We restate the MFG problem with common noise. To provide the general framework, we can consider the case where the idiosyncratic state z_t and common state A_t are diffusion processes. We denote m_t the measure of the states (a_t, z_t, A_t) and $\langle m_t, a\rangle$ the first moment of the measure w.r.t the states a_t.

$$\sup_{c_t} \mathbb{E}_0 \int_0^\infty e^{-\rho t} u(c_t) dt \tag{1}$$

$$\text{s.t.} \quad da_t = \Big(z_t \underbrace{\left(A_t\, \kappa_w \, \langle m_t, a\rangle^\alpha\right)}_{\equiv w_t} + \underbrace{\left(A_t\, \kappa_r \langle m_t, a\rangle^{\,\alpha-1} - \delta\right)}_{\equiv r_t} a_t - c_t\Big) dt \qquad \& \qquad a_t \geq \underline{a} \tag{wealth}$$

$$dz_t = b^{idio}(z_t) dt + \sqrt{2\sigma^{idio}} dB_t \tag{labor productivity}$$

$$dA_t = b^{com}(A_t) dt + \sqrt{2\sigma^{com}} dW_t^0 \tag{aggreg. productivity}$$

where we normalized labor supply $\langle m_t, z\rangle$ to a constant z_{av} and defined the constant $\kappa_w = (1-\alpha) z_{av}^{-\alpha}$ and $\kappa_r = \alpha z_{av}^{1-\alpha}$ to alleviate a bit the notations. The consumption c_t is the control of this representative agent.

We can rewrite our framework, using the standard notations from stochastic control/MFG: $X_t \equiv (a_t, z_t, A_t)$ is now a state on $\mathbb{R}^3$, the optimal control $\alpha^\star(t, x, m, p) \equiv c_t^\star$ and we use the vector formulation:

$$b(X_t, m_t, \alpha_t^\star) \equiv \begin{pmatrix} z_t\, A_t\, \kappa_w \, \langle m_t, a\rangle^\alpha + \left(A_t\, \kappa_r \langle m_t, a\rangle^{\alpha-1} - \delta\right) a_t - c_t^\star \\ b^{idio}(z_t) \\ b^{com}(A_t) \end{pmatrix}, \qquad \begin{aligned} \widetilde{\sigma}^{idio} &= (0, \sqrt{2\sigma^{idio}}, 0)^T \\ \widetilde{\sigma}^{com} &= (0, 0, \sqrt{2\sigma^{com}})^T \end{aligned}$$

Therefore, the dynamics of the state is the following:

$$\begin{aligned} dX_t &= b(X_t, m_t, \alpha_t^\star) dt + \widetilde{\sigma}^{idio} dB_t + \widetilde{\sigma}^{com} dW_t^0 \\ dX_t &= -D_p \mathcal{H}\big(t, X_t, m_t, D_x v_t(X_t)\big) dt + \widetilde{\sigma}^{idio} dB_t + \widetilde{\sigma}^{com} dW_t^0 \end{aligned} \tag{2}$$

where v_t is the value function of the control problem and $\mathcal{H}$ the hamiltonian given below and B_t and W_t^0 respectively the idiosyncratic and common noise – both are 1-dimensional Wiener processes (Brownian motion), due to the value of $\widetilde{\sigma}^{idio}$ and $\widetilde{\sigma}^{com}$. This formalism is closer to the usual notation of the MFG literature.

By posing the Hamiltonian $\mathcal{H}(t, x, m, p) = \max_\alpha \big(u(\alpha) + b(x, m, \alpha) \cdot p\big)$ we see that we are in a "strongly coupled" Mean Field Game problem: the coupling between the agents states and the measure of the other players depend on their strategies. Indeed, the prices (interest rate and wages) will depend on the measure and will affect the optimal choice. For results on strongly coupled MFG, see Bertucci et al. (2018).

Finite vs. infinite-horizon. In economics, models are often formalized in infinite horizon, and the cost is subject to the discount factor ρ, i.e. the lower ρ, the more patient/forward-looking the agents. However, for various reasons, economists are inclined to consider a stationary equilibrium – a "steady-state" – and analyze the transmission effects of a one-time shock – an "unexpected and transitory deviation from steady state" – in a given number of periods (in discrete time). In the following, we adopt an intermediary setting: for a finite time period $[0, T]$, the system is subject to aggregate shocks, and thereafter, when $t > T$, the MFG system returns to a situation without aggregate uncertainty – the common noise is "switched off". The main difference compared to usual economic models is to consider multiple trajectories of common noise: there are thus many different "steady-states". The advantage is also to consider the MFG with common noise in finite-time horizon, a setting more familiar for applied mathematicians.

Therefore, the MFG without common noise, when $\widetilde{\sigma}^{com} = \underline{0}$, is characterized by the usual[15] system of PDE:

$$\begin{cases} -\partial_t v(t,x) + \rho\, v(t,x) = \mathcal{H}\big(t,x,m,D_x v\big) + \sigma^{idio}\partial^2_{x_2} v(t,x) \qquad \text{on } [0,T]\times\mathbb{X} \\ \partial_t m(t,x) - \operatorname{div}\Big[D_p\mathcal{H}\big(t,x,m,D_x v\big)\, m(t,x)\Big] - \sigma^{idio}\partial^2_{x_2} m(t,x) = 0 \qquad \text{on } [0,T]\times\mathbb{X} \\ v(T,\cdot) = v^\infty \qquad m(0,\cdot) = m^0 \qquad \text{on } \mathbb{X} \end{cases} \tag{3}$$

where $\mathbb{X} := [\underline{a}, \infty)\times\mathbb{R}\times[0,\infty)$, and v^∞ is the value function for the stationary control problem. The Hamiltonian is given above by $\mathcal{H}(t,x,m,p) = \max_\alpha \big(u(\alpha) + b(x,m,\alpha)\cdot p\big)$. Note that we adopt this notation since, in our special setting, $\sigma^{idio}\partial^2_{x_2} v(t,x) = \frac{1}{2}\operatorname{tr}\big(\widetilde{\sigma}\widetilde{\sigma}'^{idio} D^2_x v(t,x)\big)$ and $\sigma^{idio}\partial^2_{x_2} m(t,x) = \frac{1}{2} D^2_x\big(\widetilde{\sigma}\widetilde{\sigma}'^{idio} m(t,x)\big)$. We emphasize the fact that the MFG equilibria is deterministic is this context. Let us figure out what would be the setting when a transitory – finite-horizon – aggregate shock occurs.

3.1.1 MFG system with common noise – Stochastic PDEs

When $\widetilde{\sigma}^{com} > \underline{0}$ the effect of the common noise is to randomize the MFG system. Therefore, the measure m_t becomes a random flow of measures, and can be now considered as a flow of conditional marginal measures of $(X_t)_{t\geq 0}$ given the realization of the common noise W^0_t, i.e. formally $m_t = \mathcal{L}(X_t|W^0)$. We formalize the MFG system with common noise as follow (using the notation of stochastic processes with time subscript). The reference for the derivation of such system is Cardaliaguet, Delarue, Lasry, and Lions (2017) and Carmona and Delarue (2014).

[15] The standard MFG system is recalled in appendix section A

The forward Fokker-Planck Stochastic PDE describing the dynamics of the conditional laws of the state – given the common noise – is the following :

$$\begin{aligned} dm_t(x) &= \mathrm{div}\left[D_p\mathcal{H}\big(t,x,m,D_xv\big)\,m_t(x)\right]dt - \mathrm{div}\left[\widetilde{\sigma}^{com}\,dW_t^0\,m_t(x)\right] + \tfrac{1}{2}D_x^2\big(\widetilde{\sigma}\widetilde{\sigma}'^{idio}m_t(x)\big) + \tfrac{1}{2}D_x^2\big(\widetilde{\sigma}\widetilde{\sigma}'^{com}m_t(x)\big)dt \\ &= \mathrm{div}\left[D_p\mathcal{H}\big(t,x,m,D_xv\big)\,m_t(x)\right]dt \;-\; \sqrt{2\sigma^{com}}\partial_{x_3}m_t(x)\,dW_t^0 \;+\; \big(\sigma^{idio}\partial_{x_2}^2m_t(x) \;+\; \sigma^{com}\partial_{x_3}^2m_t(x)\big)dt \\ m_0(\cdot) &= m^0 \end{aligned} \tag{4}$$

where the passage from the first to the second line applies in our special setting (and own notations). Compared to the previous framework, the common noise creates two novel terms: (i) a first order term $\sqrt{2\sigma^{com}}\partial_{x_3}m_t\,dW_t^0$ representing the direct effect of the common noise on the state of each agent (and hence on the measure) and (ii) a diffusion term $\sigma^{com}\partial_{x_3}^2m_t$ coming from the correlation between the state x_3 of all the agents – and issued from the bracket term in the Itô formula.

Given the forward dynamics of the state eq. (2), the solution of the Mean Field Game usually consists in four steps:

1. Fix an arbitrary measure m_t, that should now be a stochastic process (over $\mathcal{P}(\mathbb{X})$) adapted to the filtration $\mathcal{F}^0$ generated by the common noise W^0.
2. Solve the standard stochastic control problem eq. (1) with random coefficient, subject to the dynamics eq. (2)
3. When an optimal control exists in feedback form $\alpha^\star(t,x,m)$, plug it into the Fokker Planck above eq. (4) to obtain the evolution of m_t'.
4. Find a fixed point where this solution m' is precisely the m we started from. It reduces to a search for a flow of random measure such that $m_t = \mathcal{L}(X_t\;|\mathcal{F}^0)$

In presence of common noise, the difficulty is twofold. First, the fixed point problem is performed in an infinite dimensional space $[\mathcal{C}([0,T],\mathcal{P}(\mathbb{X}))]^\Omega$ where Ω denotes the underlying probability space carrying the common noise. This space being too large, the use of compactness arguments may fail, preventing the proof of existence by mean of Schauder's Theorem – which is the standard method to solve the case without common noise. Second, the control problem should be solved in the space of stochastic processes, i.e. adapted to $\mathcal{F}^0$. Let us see now how to deal with such issue.

When the flow of (random) measure m_t is fixed we look for an admissible control such that, for each (t,x), the controlled state $(X^{t,x,\alpha})_t$ is solution of the SDE starting from $X_t = x$:

$$dX_s = b(X_s,m_s,\alpha_s)ds + \widetilde{\sigma}^{idio}dB_s + \widetilde{\sigma}^{com}dW_s^0$$

The conditional cost and the value function being:

$$J^m_{(t,x)}((\alpha_s)_t^T) = \mathbb{E}\left[\int_t^\infty e^{-\rho s} u(\alpha_s) ds \,\middle|\, \mathcal{F}^0_t\right] \qquad\qquad v^m(t,x) = \sup_{\{\alpha_s\}_t^T} J^m_{(t,x)}((\alpha_s)_t^T)$$

Under the suitable regularity assumption, for $x \in \mathbb{X}$, one can show that $v^m(t,x)$ is a $(\mathcal{F}^0_s)_{t\le s\le T}$–semi-martingale and one can use the Itô formula and invoke a generalization of the Dynamic Programming Principle in this random environment.

We define the random operator:

$$\begin{aligned}\mathcal{L}(t,x,m,p,M,q^0) &= \sup_\alpha \left[\, u(\alpha) + b(x,\mu,\alpha)\cdot p\,\right] + \tfrac{1}{2}\operatorname{tr}\left(\widetilde{\sigma}\widetilde{\sigma}'^{idio} M\right) + \tfrac{1}{2}\operatorname{tr}\left(\widetilde{\sigma}\widetilde{\sigma}'^{com} M\right) + \widetilde{\sigma}^{com}\cdot q^0 \\ &= \mathcal{H}\big(t,x,m,D_x v\big) + \sigma^{idio} M_{2,2} + \sigma^{com} M_{3,3} + \sqrt{2\sigma^{com}}\, q^0_3\end{aligned}$$

where $\mathcal{H}$ is again the same hamiltonian (since the diffusion term is uncontrolled). We introduces an additional term $Z^{0,m}$, where the couple $(v^m, Z^{0,m})$ is solution of the Backward SDE – parametrized by m – of the following form:

$$v^m_t(x) = v^\infty(x) + \int_t^T \mathcal{L}\big(t,x,m,D_x v^m(s,x), D_x^2 v^m_s(x), D_x Z^{0,m}_s(x)\big) ds + \int_t^T Z^{0,m}_s(x) dW^0_s$$

As usual in the theory of Backward Stochastic Differential Equation, "$-\mathcal{L}$" is the generator of the BSDE, $v^\infty(x)$, the terminal condition – i.e. the value function of the MFG without common noise here – and $Z^{0,m}$ that allows the solution to be adapted the the filtration $(\mathcal{F}^0_s)_s$. Written differently, the dynamics of the value function $v^m(t,x)$ – with the notation for stochastic processes – yields the following Stochastic Hamilton-Jacobi-Bellman :

$$\begin{aligned} dv^m_t &= \rho\, v^m_t\, dt - \mathcal{L}\big(t,x,m,D_x v^m(t,x), D_x^2 v^m_t(x), D_x Z^{0,m}_t(x)\big) dt + Z^{0,m}_t(x) dW^0_t \\ &= \Big[\rho v^m_t - \mathcal{H}\big(t,x,m,D_x v^m\big) - \sigma^{idio}\partial^2_{x_2} v^m_t - \sigma^{com}\partial^2_{x_3} v^m_t - \sqrt{2\sigma^{com}}\partial_{x_3} Z^{0,m}_t\Big] dt + Z^{0,m}_t dW^0_t \quad \text{on } \mathbb{X} \ \&\ t\in[0,T] \\ v^m_T &= v^\infty \qquad \text{on } \mathbb{X}\end{aligned}$$

where the random field $Z^{0,m}$ is a 1-dimensional process, just like W^0, but is a function of the state x. Note that the (random) infinitesimal generator $\mathcal{L}(\cdot)dt + Z^{0,m} dW^0_t$ is also the adjoint operator of the generator driving the Stochastic Fokker Planck eq. (4).

By verification arguments, one could find the optimal strategy in feedback form when looking for the supremum of the Hamiltonian[16].

[16] We also precise that the optimal control $\alpha^\star$ realizing the supremum of the hamiltonian is only a function of $(t,x,m,D_x v)$ since the volatility terms of both the idiosyncratic and common noises are uncontrolled. If it would be the case, the optimal strategy would then depend on $D^2_x v^m$ and $Z^{0,m}$.

Let us provide an alternative interpretation for the random vector field $Z_t^0(\cdot)$ and describe the two new terms – compared to the HJB without common noise. The term $Z_t^0 dW_t^0$ is reminiscent of the theory of Backward SDE, since the HJB runs backward : it guarantees that the solution $(v_t)_0^T$ is adapted to the filtration generated by the common noise W^0, since in presence of aggregate uncertainty the agents can not foresee the future of the economy. The second term $-\sqrt{2\sigma^{com}}\partial_{x_3} Z_t^0$ could be understood as a first-order effect of this additional noise taken at the value x. It could well be understood as the "effect of common uncertainty" – a diffusion here – which here transmit to the value function only through the state x_3. Mathematically, this extra term can be explained by the Itô-Wentzell for random field. This formula[17] allows to perform the chain rule for random fields applied to random processes. When doing so, the derivative of the diffusion term of the random field – i.e. the value effect of uncertainty – interact its effect on the state, i.e. $\partial_{x_3} Z_t$ and $\sqrt{2\sigma^{com}}$ respectively here. This term is thus cancelled out by the term $-\sqrt{2\sigma^{com}}\partial_{x_3} Z_t^0$ in the dynamics of the value function. For further reference, see chapter 4 and appendix of Cardaliaguet, Delarue, Lasry, and Lions (2017).

Dropping the m superscript when the fixed point is found, we rewrite the system of Forward-Backward SDE or Stochastic PDE :

$$\begin{cases} dv_t &= \Big[\rho\, v_t - \mathcal{H}\big(t,x,m,D_x v_t\big) - \sigma^{idio}\partial_{x_2}^2 v_t - \sigma^{com}\partial_{x_3}^2 v_t - \sqrt{2\sigma^{com}}\partial_{x_3} Z_t^0\Big]\, dt + \, Z_t^0 dW_t^0 \qquad \text{on } \mathbb{X} \quad \& \; t\in[0,T] \\ dm_t &= \Big[\operatorname{div}\big(D_p\mathcal{H}\big(t,x,m,D_x v\big)\, m_t\big) + \sigma^{idio}\partial_{x_2}^2 m_t + \sigma^{com}\partial_{x_3}^2 m_t\Big] dt - \sqrt{2\sigma^{com}}\,\partial_{x_3} m_t\, dW_t^0 \qquad \text{on } \mathbb{X} \quad \& \; t\in[0,T] \\ v_T &= v^\infty \qquad\qquad m_0 = m^0 \qquad \text{on } \mathbb{X} \end{cases} \tag{5}$$

[17] Briefly, take a random vector field Ψ_t, function of a state x, expanded under an Itô form: $d\Psi_t(x) = f_t(x)dt + g_t(x)dW_t^0$. When considering an Itô process : $dX_t = b_t dt + \sigma^i dB_t + \sigma^c dW_t^0$, one would like to know the Itô decomposition of $\Psi_t(X_t)$. Under suitable assumption, and when the Brownians are 1-d, the Itô-Wentzell formula provides this expansion:

$$d\Psi_t(X_t) = \Big[D\Psi_t(X_t)\cdot b_t + \tfrac{1}{2}\operatorname{tr}\big(D^2\Psi_t(X_t)(\sigma^i\sigma'^i + \sigma^c\sigma'^c)\big) + f_t(X_t) + Dg(X_t)\cdot\sigma^c\Big]\, dt + D\Psi_t(X_t)(\sigma^i dB_t + \sigma^c dW_t^0) + g_t(X_t)dW_t^0$$

3.1.2 Master equation

We now discuss how the recent literature of Mean Field Game would handle the problem with common noise. This approach consists in describing the Nash equilibrium of the continuum of players as a single equation, known as the *master equation*. This equation is the limiting form of the Nash-system of a $N-$players stochastic differential game – that would take the form of N quasilinear parabolic PDE. When the limit is drawn, the resulting master equation is set over the state-space $\mathbb{X}$ times the space of probability measure $\mathcal{P}(\mathbb{X})$. The Mean Field Game problem is thus reformulated in a single equation in infinite-dimension, that integrates *both* the control of the players and the evolution of their distribution. This formulation was introduced by P. L. Lions in his lectures at the *Collège de France*, and is studied in details in Cardaliaguet, Delarue, Lasry, and Lions (2017). This novel description proves to be very flexible and powerful to integrate the complexity of the system. It allows to tackle the convergence problem of N-players Nash equilibria toward the MFG system – an difficult question that has remained opened until recently – or games with both major and minor players – a setting that could have many different applications in economics. Moreover, this formulation is the right setting to understand MFG with common noise, and we try to provide informal intuitions in this section.

Master equation for MFG without common noise

To provide an introduction, the master equation of the deterministic MFG system found in eq. (3) would be :

$$\begin{cases} -\partial_t U(t,x,m) + \rho U(t,x,m) - \mathcal{H}(t,x,m,D_x U) - \sigma^{idio}\partial^2_{x_2} U(t,x,m) - \displaystyle\int_{\mathbb{X}} \sigma^{idio}\partial_{y_2}[D_m U(t,x,m;y)]\, m(dy) \\ \qquad + \displaystyle\int_{\mathbb{X}} D_m U(t,x,m;y)\cdot D_p\mathcal{H}\big(t,y,m,D_x U(t,y,m)\big)m(dy) = 0 \qquad \text{in } [0,T]\times\mathbb{X}\times\mathcal{P}(\mathbb{X}) \\ U(t,x,m) = v^{\infty}(x,m) \qquad \text{in } \mathbb{X}\times\mathcal{P}(\mathbb{X}) \end{cases} \tag{6}$$

There are several interpretations for the solution $U : [0,T]\times\mathbb{X}\times\mathcal{P}(\mathbb{X}) \mapsto \mathbb{R}$:

- This represents the value function of the representative player i in a $N-$players differential game, when $N \to \infty$. In this Nash-system, the value function would depend on the $N-1$ others states $\{x_1, \ldots, x_{i-1}, x_{i+1} \ldots, x_N\} \equiv m_x^{N,i}$ when the states vector is embedded in the space $\mathcal{P}(\mathbb{X})$ and be represented by the empirical distribution $m_x^{N,i} = \frac{1}{N}\sum_{j\neq i}\delta_{x_j}$.
- This value, as solution of the master equation, will be constructed *via* the methods of characteristics: Starting from $(t_0, m_{(0)})$ one define $U(t_0,x,m_{(0)}) := v(t_0,x)\forall x \in \mathbb{X}$ where (v,m) is solution of the MFG system eq. (3). We can define $U(t,x,m_t)$ all along the characteristic (v,m). When assuming that U is smooth enough – while it is actually difficult to prove – one can show that U will be solution of the master equation, as we will explain in the case with common noise – following the method of Cardaliaguet et al. (2017).

- This function on the Wasserstein space can also be viewed as the decoupling field of the MFG system. This approach, particularly relevant in the case of Forward Backward SDE (FBSDE) is well treated in the 2nd volume of Carmona and Delarue (2018). If we will provide intuitions in the case of common noise, one can already emphasize that this function will allow to express the backward component of the system as a deterministic function of the first. Indeed, some form of Markov property holds on the enlarged space $\mathbb{X} \times \mathcal{P}(\mathbb{X})$. This intuition holds both in the case with or without common noise, where – in the latest – the system of SPDEs (v_t^m, m_t) as seen in eq. (5) can in fact be viewed a FBSDE. The decoupling field $v_t^m(x) = U(t, x, m_t)$ and U can be shown – using Itô in random environment – to solve the master equation.

Let us also do some remarks on the first order master equation:

- If the usual time and space derivative are taken for $U(\cdot, \cdot, m)$, we now require differential calculus in the space of probability measure to understand the formalism behind the mapping $\mathbb{X} \ni y \mapsto D_m U(t, x, m; y)$. We will try to provide intuitions below and the interested readers can find very good introductions in Cardaliaguet (2018) and in Vol. 1, Chap 4. of Carmona and Delarue (2018).
- Therefore, note that this equation consists in two parts: the four first terms involve the local terms (t, x) and are the same as the one which appear in the Hamilton-Jacobi-Bellman equation: they correspond to the local control problem. Let us recall that the Hamiltonian in our problem is given by $\mathcal{H}(t, x, m, p) = \sup_\alpha u(\alpha) + b(x, m, \alpha) \cdot p$.
- The two last terms, however, are non-local terms and are defined as integrals over the whole space and integrate the coupling between the agents. The derivatives with respect to the measure correspond to the flows of the agents as could be seen in the Fokker Planck equation.

Aggregation: Master equation simplified

In the case without common noise, to provide further intuitions, we could now *suppose* that one can aggregate the economy. The measure of agent would be represented by some moments of the distribution $\langle h, m \rangle$, for example the first moment $K_t = \overline{m}_t = \langle a, m_t \rangle$.

The aggregation phenomenon is well-known is economics, under some specific conditions. One typical example is when there is *no credit constraint*: the state space is given by $\widetilde{\mathbb{X}} = \mathbb{R} \times \mathbb{R} \times [0, \infty)$ (instead of $\mathbb{X} = [\underline{a}, \infty) \times \mathbb{R} \times [0, \infty)$ with a state-constraint).

In such case, the economy with idiosyncratic risk but without state-constraint is isomorphic to the representative agent model (with and without aggregate shocks). The "average" households will accumulate or diminish freely their asset/capital stock in function of sequence of shocks they faces. This evolution will be "priced" and the interest rate will reflect the excess or shortage of capital. Therefore, the control problem reduces to a finite-dimensional problem – as we will show thereafter.

The probabilistic treatment of such phenomenon – using FBSDE and Stochastic Pontryagin Principle – is covered in Carmona and Delarue (2018), in section 3.6. of Vol 1. There[18], the Backward SDE for the adjoint (co-state) variable – representing in a sense marginal utility – is in fact deterministic and can be decoupled from the Forward SDE of the dynamic of the state $X_t = (a_t, z_t, A_t)^T$. The decoupling field is thus function – solving the adjoint equation – of the first moment of the distribution : $Y_t = g(\langle m_t, a\rangle)$. We use the same logic here, *assuming that* the value function – solution of the master equation – is only function of some moment of the distribution of agents $K^{h,m} = \langle h, m_t\rangle$, i.e. :

$$K^{h,m} = \langle h, m_t\rangle = \int_{\widetilde{\mathbb{X}}} h(x)m_t(dx) \qquad\qquad U(t,x,m) = \overline{U}(t,x,K^{h,m}) = \overline{U}(t,x,\langle h,m\rangle)$$

The main idea behind aggregation in these MFG would be to obtain the same solution for an arbitrary function $U(\cdot, m)$ or for a function $\overline{U}(\cdot, \langle h,m\rangle)$ considering only the first moment of the distribution $h(x) = x_1 \equiv a$ and $h : \mathbb{X} \mapsto \mathbb{R}$, and $\overline{U} : [0,T]\times\widetilde{\mathbb{X}}\times\mathbb{R} \mapsto \mathbb{R}$. Indeed, in the Krusell-Smith model (with or without credit constraint) the dynamics of the model is only coupled through the first-moment of the measure, in the interest rate and wage. One could extend the procedure used in the following, to any functions $U(m) = \overline{U}(g_1(m), \dots, g_k(m))$ with g_k any monomial function $g_k(m) = \Pi_j \langle h_j^k, m_t\rangle$. In this setting, one would obtain a finite dimensional PDE, as we will see now.

Derivative in the space of measure. We now briefly use two simple examples to compute the derivative of functions in the Wasserstein space, i.e. to give a sense of $D_m U$. The idea is to represent a function of a measure $g(m)$ as a function of the random variable following this law $\widetilde{g}[X]$ and $X \sim m$ – called "lifting" or "extension". The derivative w.r.t. to the measure is the derivative of this lifted function : $D_m g(m,y) \equiv D_y \widetilde{g}([X], y)$. One can also use the notion using "intrinsic derivative" $\frac{\delta g}{\delta m}$ and $D_m g(m,y) \equiv D_y \frac{\delta g}{\delta m}(m,y)$. Note that we perform the following computations in a very informal way, and complete references are found in Cardaliaguet (2018)

- If $g_1(m) = \langle h, m\rangle = \int_{\mathbb{R}^d} h(x)m(dx)$. The lifting is $\widetilde{g_1}[X] = \int_{\mathbb{R}^d} h(x)m(dx) = \mathbb{E}(h(X))$. Its derivative would be $D\widetilde{g_1}[X](Z) = \mathbb{E}(D\widetilde{g_1}[X] \cdot (Z - X)) = \mathbb{E}(Dh(X) \cdot (Z - X))$. The resulting derivative w.r.t. m is therefore :

  $D_m g_1(m; y) = D_y h(y) \qquad$ with $\qquad D_m g_1(m; \cdot) : \mathbb{R}^d \mapsto \mathbb{R}^d, y \to D_m g_1(m; y)$
- A simple variation: $g_2(m) = \psi(\langle h, m\rangle)$. Using the chain rule with the method above:

[18] Note that the model is also solved in section 6.7 using optimal control of McKean Vlasov dynamics instead of Mean Field Game (i.e. Pareto equilibrium instead of Nash Equilibrium)

$$D_m g_2(m;y) = \psi'(\langle h, m\rangle) D_y h(y) \qquad \text{with} \quad D_m g_2(m;\cdot) : \mathbb{R}^d \mapsto \mathbb{R}^d, y \to D_m g_2(m;y)$$

Using this simple example, for the master equation, the derivative of the solution is thus:

$$D_m U(t,x,m;y) := (\partial_{m_1} U, \partial_{m_2} U, \partial_{m_3} U)^T \qquad \partial_{m_i} U(t,x,m;y) = \partial_{K^h} \overline{U}(t,x,K^{h,m}) \partial_{y_i} h(y)$$

Applying these techniques and after some simple calculus, the master equation with aggregation $K^h = \langle h, m\rangle$ becomes:

$$\begin{cases} -\partial_t \overline{U}(t,x,K^h) + \rho\, \overline{U}(t,x,K^h) - \overline{\mathcal{H}}(t,x,K^h, D_x\overline{U}) - \sigma^{idio} \partial^2_{x_2} \overline{U}(t,x,K^h) \\ \qquad - \sigma^{idio} \partial_K \overline{U}(t,x,K^h) \displaystyle\int_{\widetilde{\mathbb{X}}} \partial^2_{y_2} h(y)\, m(dy) \\ \qquad + \partial_K \overline{U}(t,x,K^h) \displaystyle\int_{\widetilde{\mathbb{X}}} \nabla h(y) \cdot D_p \overline{\mathcal{H}}(t,x,K^h, D_x\overline{U})\, m(dy) = 0 \qquad \text{in } [0,T]\times\mathbb{X}\times\mathbb{R} \\ \overline{U}(t,x,K^{h,m}) = v^\infty(x,m) \qquad \text{in } \mathbb{X}\times\mathbb{R} \quad \text{with} \ \ \langle h, m\rangle = K^{h,m} \end{cases} \tag{7}$$

Result: Aggregation

1. If the coupling of the MFG system is reduced to the dependence in some moments of the distribution, i.e. $K^{h,m} = \langle h, m_t\rangle$ and $U(t,x,m) = \overline{U}(t,x,K^{h,m})$, then the value function of the MFG can be represented by the function $\overline{U}$ solving the master equation eq. (7)
2. If one can aggregate the economy, through the first moment of the measure: $K = \bar{m}_t = \langle a, m_t\rangle$, then[19] the master equation reduces to a standard Hamilton Jacobi Bellman equation, on extended space $\widetilde{\mathbb{X}}\times\mathbb{R}$ and is given by :

$$\begin{cases} -\partial_t \overline{U}(t,x,K) + \rho\, \overline{U}(t,x,K) - \overline{\mathcal{H}}(t,x,K, D_x\overline{U}) - \sigma^{idio} \partial^2_{x_2} \overline{U}(t,x,K) \\ \qquad + \partial_K \overline{U}(t,x,K) \displaystyle\int_{\widetilde{\mathbb{X}}} \partial_{p_1} \overline{\mathcal{H}}(t,y,K, D_x\overline{U})\, m(dy) = 0 \qquad \text{in } [0,T]\times\mathbb{X}\times\mathbb{R} \\ \overline{U}(t,x,K) = v^\infty(x,m) \qquad \text{in } \widetilde{\mathbb{X}}\times\mathbb{R} \quad \text{with} \ \ \langle a, m\rangle = K \end{cases} \tag{8}$$

Note that, when the measure is unknown, the integral term $\int_{\widetilde{\mathbb{X}}} \partial_{p_1} \overline{\mathcal{H}} m(dy)$ may still be a highly non-linear function of the state, and can not be simplified easily. The aggregation result in economics would search to express the controlled drift $\partial_{p_1} \mathcal{H}$ as a linear function of the state $h(x)$. If this property holds, then the master equation would not feature any more measure terms, but only functions of K^h.

[19] Since $h(x) = x_1$, we thus have $\nabla_x h(x) = (1,0,0)^T$ and $\Delta h(x) = 0$

Master equation for MFG with common noise

We now provide intuitions for the master equation with common noise, i.e. when $\sigma^{com} > 0$. The corresponding MFG system is developed eq. (5), where the two PDE are stochastic. The master equation, a contrario will be a deterministic PDE on $[0,T]\times\mathbb{X}\times\mathcal{P}(\mathbb{X})$ and it will account for the evolution of the common noise and its influence on the distribution.

$$
\begin{cases}
-\partial_t U(t,x,m) + \rho U(t,x,m) - \mathcal{H}(t,x,m,D_x U) - \sigma^{idio}\partial^2_{x_2} U(t,x,m) - \sigma^{com}\partial^2_{x_3} U(t,x,m) \\
\quad -(\sigma^{idio} + \sigma^{com})\displaystyle\int_{\mathbb{X}} \operatorname{div}_y [D_m U(t,x,m;y)]\, m(dy) + \int_{\mathbb{X}} D_m U(t,x,m;y)\cdot D_p\mathcal{H}\big(t,y,m,D_x U(\cdot,y,\cdot)\big) m(dy) \\
\quad -2\sigma^{com}\displaystyle\int_{\mathbb{X}} \operatorname{div}_x [D_m U(t,x,m;y)]\, m(dy) - \sigma^{com}\int_{\mathbb{X}\times\mathbb{X}} \operatorname{tr}\left[D^2_{mm} U(t,x,m;y,y')\right] m_t^{\otimes 2}(dy,dy') = 0 \qquad \text{in } [0,T]\times\mathbb{X}\times\mathcal{P} \\
U(t,x,m) = v^{\infty}(x,m) \qquad\qquad \text{in } \mathbb{X}\times\mathcal{P}(\mathbb{X})
\end{cases}
\tag{9}
$$

In the model considered, the idiosyncratic and common noise affect respectively the second and third states and the agents controls interact only through the measure of the first state. Using these observations and using the formulation of eq. (5) – since the master equation is obtained through the characteristics of the MFG system – we use the method developed in Cardaliaguet et al. (2017). In our case, the first order and second order derivatives simplify and the resulting master equation writes as:

$$
\begin{cases}
-\partial_t U + \rho U - \mathcal{H}(\cdot, D_x U) - \sigma^{idio}\partial^2_{x_2} U - \sigma^{com}\partial^2_{x_3} U - \displaystyle\int_{\mathbb{X}} \sigma^{idio}\partial_{y_2}\partial_{m_2} U(\cdot;y) + \sigma^{com}\partial_{y_3}\partial_{m_3} U(\cdot;y)\, m(dy) \\
\qquad + \displaystyle\int_{\mathbb{X}} \partial_{m_1} U(\cdot;y)\, \partial_{p_1}\mathcal{H}\big(\cdot,y,\cdot,D_x U(\cdot,y,\cdot)\big) m(dy) - 2\sigma^{com}\int_{\mathbb{X}} \partial_{x_3} D_{m_3} U(\cdot,x,\cdot;y) m(dy) \\
\qquad -\sigma^{com}\displaystyle\int_{\mathbb{X}\times\mathbb{X}} \partial^2_{m_3} U(\cdot,x,\cdot;y,y')\, m_t^{\otimes 2}(dy,dy') = 0 \qquad \text{in } [0,T]\times\mathbb{X}\times\mathcal{P}(\mathbb{X}) \\
U(t,x,m) = v^{\infty}(x,m) \qquad\qquad \text{in } \mathbb{X}\times\mathcal{P}(\mathbb{X})
\end{cases}
\tag{10}
$$

where U and $\mathcal{H}$ are taken in (t,x,m) and $U(\cdot,x,\cdot;y) = U(t,x,m;y)$ where there could be ambiguities. Note that to transform the MFG system of SPDE into the master equation, we identified the random field Z^0_t with :

$$
Z^0_t(x) = \int_{\mathbb{X}} D_m U(t,x,m;y) m_t(dy)
$$

Let us also do some remarks on this second order master equation eq. (10) and the effects of the common noise:

- This equation is of second order, since the derivative w.r.t. the measure m_t is of second order both with the terms $\operatorname{div}(D_m U)$ and $D^2_{mm} U$.
- This equation is analogous to the infinite-dimensional HJB expressed in equation (43) of the

Appendix A.1. of Ahn, Kaplan, Moll, Winberry, and Wolf (2018)

- As in the case without common noise, the first five terms correspond to the control dynamics – involved in the HJB – and represent the anticipation of an agent of his own future states.
- The last five terms are non-local terms and describe the particles evolution – as in the FP equation. To be more specific, among these terms:

(i-ii) the first & second terms show how value changes for a distortion of the measure after a variation in states x_2 and x_3: these second order derivatives are the direct effects of a Brownian changes in idiosyncratic or common noise – hence the variance terms σ^{idio} and σ^{com}.

(iii) the third terms show the agents control will change the state $x_1 \equiv a$ (wealth) and how the wealth distribution will evolve. Since the agents observe ("expect") the other agents move, the expectation is drawn on all $\mathbb{X}$.

(iv) this additional term in $\partial_{x_3} D_{m_3} U(x)$ is the effect of uncertainty of $x_3 \equiv A_t$ on the value of the agents. Measured on the state x – with ∂_{x_3} – this term appears with the random field $Z_t^0(x)$ and represents how an agent of state x anticipate the common noise.

(v) the last term shows how an agent expect the anticipations of the other agents (!). This diffusion term is thus integrated over the whole space twice: the agent x will account on how the agents y will expects the moves of the players y' (!!).

Aggregation with common noise

We will reiterate the procedure we used above to derive the master equation when the dependence of the measure can be summarized to a finite set of moments :

$$K^{h,m} = \langle h, m_t \rangle = \int_{\mathbb{X}} h(x) m_t(dx) \qquad U(t,x,m) = \overline{U}(t,x,K^{h,m}) = \overline{U}(t,x,\langle h, m\rangle)$$

$$D_m U(t,x,m;y) := (\partial_{m_1} U, \partial_{m_2} U, \partial_{m_3} U)^T \qquad \partial_{m_i} U(t,x,m;y) = \partial_{K^h} \overline{U}(t,x,K^{h,m})\, \partial_{y_i} h(y)$$

The important point is that the functions we consider $g(m) = \langle h, m\rangle$ are linear functions of m. Indeed, the derivative is no longer a function of m (only of K^h and y). Therefore, the second order derivative $D_{mm} g(m)(y,y') = 0$. Consequently, the (last) term of type (v) is cancelled, simplifying substantially the equation.

Applying the above techniques and some calculus, the master equation with common noise, subject to aggregation $K^h = \langle h, m\rangle$ becomes:

$$
\begin{cases}
-\partial_t \overline{U}(t,x,K^h) + \rho\, \overline{U}(t,x,K^h) - \overline{\mathcal{H}}(t,x,K^h,D_x\overline{U}) - \sigma^{idio}\partial^2_{x_2}\overline{U}(t,x,K^h) - \sigma^{com}\partial^2_{x_3}\overline{U}(t,x,K^h) \\
\qquad - \partial_K \overline{U}(t,x,K^h)\Big[\int_{\widetilde{\mathbb{X}}} \sigma^{idio}\partial^2_{y_2}h(y) \ + \sigma^{com}\partial^2_{y_3}h(y)\, m(dy)\Big] \\
\qquad + \partial_K \overline{U}(t,x,K^h)\int_{\widetilde{\mathbb{X}}} \nabla h(y)\cdot D_p\overline{\mathcal{H}}(t,y,K^h,D_x\overline{U})\, m(dy) \\
\qquad -2\sigma^{com}\partial_{x_3}\partial_K\overline{U}(t,x,K^h)\int_{\mathbb{X}}\partial_{y_3}h(y)m(dy) = 0 \qquad \text{in } [0,T]\times\mathbb{X}\times\mathbb{R} \\
\overline{U}(t,x,K^{h,m}) = v^\infty(x,m) \qquad \text{in } \mathbb{X}\times\mathbb{R} \quad \text{with } \langle h,m\rangle = K^{h,m}
\end{cases}
\tag{11}
$$

Result: Aggregation with common noise

1. In presence of common noise, $\sigma^{com} > 0$, if the coupling of the MFG system is reduced to the dependence in some moments of the distribution, i.e. $K^{h,m} = \langle h, m_t\rangle$ and $U(t,x,m) = \overline{U}(t,x,K^{h,m})$, then the value function of the MFG can be represented by the function $\overline{U}$ solving the master equation eq. (11)
2. If one can aggregate the economy, through the first moment of the distribution: $K = \bar{m}_t = \langle a, m_t\rangle$, then[20] the master equation reduces to a standard Hamilton Jacobi Bellman equation, on extended space $\widetilde{\mathbb{X}}\times\mathbb{R}$ and is given by :

$$
\begin{cases}
-\partial_t \overline{U}(t,x,K) + \rho\, \overline{U}(t,x,K) - \overline{\mathcal{H}}(t,x,K,D_x\overline{U}) - \sigma^{idio}\partial^2_{x_2}\overline{U}(t,x,K) - \sigma^{com}\partial^2_{x_3}\overline{U}(t,x,K) \\
\qquad + \partial_K \overline{U}(t,x,K)\int_{\widetilde{\mathbb{X}}} \partial_{p_1}\overline{\mathcal{H}}(t,y,K,D_xU(\cdot,y,\cdot))\, m(dy) = 0 \qquad \text{in } [0,T]\times\mathbb{X}\times\mathbb{R} \\
\overline{U}(t,x,K) = v^\infty(x,m) \qquad \text{in } \mathbb{X}\times\mathbb{R} \quad \text{with } \langle a,m\rangle = K
\end{cases}
\tag{12}
$$

In this case, when aggregation occurs, there is no more feedback loop of the expectations of other agents: the terms (i-ii) and (iv-v) disappear since the measure of states x_2 and x_3 has no influence on states and controls of the other players. However, as before, the term (iii) remains, showing how the flow (in state $x_1 \equiv a$) will change wealth distribution and prices r_t and w_t. However, when the measure is unknown, the integral term:

$$
\int_{\widetilde{\mathbb{X}}} \partial_{p_1}\overline{\mathcal{H}}(t,y,K,D_xU(\cdot,y,\cdot))\, m(dy)
$$

is unknown. The main idea behind aggregation result in economics is to be able to express the optimal drift $\partial_{p_1}\overline{\mathcal{H}}(t,y,K,D_xU(\cdot,y,\cdot))$ as a linear function of the state – or a linear function of the function h. Only in such case, the master equation would be an equation of K (or K^h) and would not feature the measure m.

[20] Since again $h(x) = x_1$, we thus have $\nabla_x h(x) = (1,0,0)^T$ and $\Delta h(x) = 0$

3.1.3 Weak equilibria of MFG with common noise

The strategy developed in Carmona, Delarue, and Lacker (2016) to handle MFG with common noise is completely different as the one described in the two previous sections. Nevertheless, this method could be seen as a rationale for the numerical scheme we develop in the next section.

The approach relies on pure probabilistic arguments and a procedure to discretize the common noise. Focusing on the case where the common noise has a finite number of outcomes, one can use the "usual" procedure for MFG existence, this time conditioning on a finitely-supported approximation of the aggregate noise. The authors use Kakutani's fixed point theorem – a extension to multiple controls (& correspondences) of the Brouwer fixed point theorem – to prove existence.

However, beside on finite approximation, conditioning usual fail to be continuous and one can not recover the existence at the limit: the limiting solution can actually fail to be adapted to the filtration generated by the common noise. The existence when refining the discretization is thus only proved as a *weak equilibrium*, exactly as weak solutions for stochastic differential equations, where here more specifically the measure m_t of the controlled process may not be adapted to the aggregate shocks. However, under the assumption of pathwise uniqueness, the authors can prove an analogous version of the Yamada-Watanabe theorem for SDE: any weakly existing equilibria that satisfy pathwise uniqueness – i.e. indistinguishable trajectories – is in fact a strong solution of the MFG.

3.2 The discretization procedure

The previous section was an introduction to the mathematical formalism of MFG with common noise. Starting from the framework described in section 3.1, we present here one numerical technique to tackle this problem.

The main idea is to approximate the stochastic process of common noise. In our setting, the only state affected by this aggregate shock is the productivity level A_t. We use a tree structure to approximate this process in time with a finite number M of simple shocks, and in space with a finite number K of possible states. With this method, one can approximate any process using a finite number K^M of trajectories. When we consider a simple Brownian Motion dW^0 or other diffusion processes, we can use different tree structures: (i) binomial, trinomial or "K-nomial" trees and (ii) optimal quantization trees. The different between these two approximations will be explained in the following sections.

Building trees

More precisely, in our case, the process for productivity A_t, will start at $\bar{A}_0$. We first discretize the time: $[0,T]$ is divided $M+1$ periods $[t_m, t_{m+1}]$ with $m \in \{1,\dots,M\}$, say uniformly, every $\Delta T = \frac{T}{M+1}$. At each date t_m, the process A_t will switch between K deterministic trajectories $(A_t^k)_{t\in[t_m,t_{m+1}]}$ with $k \in \{1,\dots,K\}$. The probability of transition from one state A_{t_m} to one of these K trajectories is $\pi_{k_{m+1}|k_m} := \mathbb{P}(A_t = A_t^k\ ,\ \forall t \in [t_m, t_{m+1}]\,\big|A_{t_m})$

This way, we can "build" a *tree* of different trajectories of common noise, and the discretized process is still stochastic, with the probability of transition given by the coefficient of the hypermatrix $\pi_{k_{m+1}|k_m}$ – whose coefficients should be carefully chosen as we will see. When taking $\Delta T \to 0$, you can approximate any process.

Solving the MFG system – grafting branches

A way to 'solve' the MFG with common noise will be to compute the evolution of the MFG system – HJB and FP equations – on *each branch*, i.e. on each deterministic trajectories $(A_t^k)_{t\in[t_m,t_{m+1}]}$

For a given trajectory, the value function at time t_m will depends in the state A_{t_m} and will writes:

$$
\begin{aligned}
v(t_m, x, A_{t_m}) &:= \sup_{c_t} \mathbb{E}\Big[\int_{t_m}^{T} e^{-\rho t} u(c_t)dt \;\Big|\; A_{t_m}\Big] \\
&= \sup_{c_t} \mathbb{E}\Big[\int_{t_m}^{t_{m+1}} e^{-\rho t} u(c_t)dt + e^{-\rho t_{m+1}} v(t_{m+1}, X_{t_{m+1}}, A_{t_{m+1}}) \;\Big|\; A_{t_m}\Big]
\end{aligned}
$$

by applying standard dynamic programming arguments. Therefore, by denoting t_m^- the time before revelation of the shock, and t_m^+ when the future trajectory is revealed, the HJB equation with random coefficients $A_t^{k_m}$ becomes:

$$\begin{cases} -\partial_t v(t,x,A_t^{k_m}) + \rho\, v(t,x,A_t^{k_m}) = \mathcal{H}\big(t,x,m,D_x v\big) + \sigma^{idio}\partial_{x_2}^2 v(t,x,A_t^{k_m}) & \text{on } [t_m,t_{m+1}]\times\mathbb{X} \\ v(t_{m+1}^-,x,A_t^{k_m}) = \mathbb{E}\Big[v(t_{m+1}^+,x,A_{t_{m+1}}) \;\Big|\; A_{t_m}^{k_m}\Big] & \text{on } \mathbb{X} \end{cases}$$

The dynamic of A_t will thus matter a lot on how the agents will consider this terminal condition, which can thus be written, when the common noise is discretized, as follow:

$$v(t_{m+1}^-,x,A^{k_m}) = \sum_{k=1}^{K} \pi_{k|k_m}\, v(t_{m+1}^+,x,A_{t_{m+1}}^k)$$

In particular, since agents are forward looking (and rational!), they form expectations over the different set of future branches, and that will be accounted in the value function. The value function will 'jump' between t_m^- and t_m^+ to correct the (past) expectations after the shock of information. Also, the presence of this conditional expectation $\mathbb{E}[\cdot\,|\,A_{t_m}]$ will allow the value function to be adapted to the filtration generated by the common noise A_t.

Note that, at the last period $[t_M,T]$, there is no more uncertainty and the terminal condition becomes:

$$v(T,\cdot,A_T^{k_M}) = v^{\infty,k_M}$$

where v^{∞,k_M} is the stationary value function of one of the K^M terminal equilibria. To compute these value function, the algorithm will therefore the HJBs backward, starting from the terminal condition and correcting the value function at each node by taking the conditional expectation w.r.t. the common noise.

Consider now the associated Fokker Planck equation, on each interval $[t_m,t_{m+1}]$, when the controlled drift is given by $D_p\mathcal{H}$. Again, the measure depends on the random coefficient A^{k_m} through the controls:

$$\begin{cases} \partial_t m(t,x,A^{k_m}) - \operatorname{div}\Big[D_p\mathcal{H}\big(t,x,m,D_x v(\cdot,A^{k_m})\big)\, m(t,x,A^{k_m})\Big] - \sigma^{idio}\partial_{x_2}^2 m(t,x,A^{k_m}) = 0 & \text{on } [t_m,t_{m+1}]\times\mathbb{X} \\ m(t_m^+,\cdot,A_{t_m^+}^{k_m}) = m(t_m^-,\cdot,A_{t_m^-}^{k_{m-1}}) & \text{on } \mathbb{X} \end{cases}$$

where $m(t_1^+,\cdot) = m^0$ for the first period starting at $t_1 := 0$. However, in contrast of the value function, the measure is backward-looking and naturally accounts for the past drift values. It is therefore continuous in time t. To compute the measure, the F.P. equations will be solved forward, starting from the initial condition for each trajectory of the common noise.

3.2.1 Binomial/trinomial tree structure

When we consider a simple Brownian Motion dW^0 or other diffusion processes[21], an obvious approximation would be to consider a simple random walk, a process that rises or drops with probability 1/2. Moreover, one could also think about the "normal"-random-walks:

$$S_n = \sum_k^n \varepsilon_k \qquad \text{with } \varepsilon_k \sim \mathcal{N}(0,1) \qquad \text{and } W_t^{(n)} := \frac{S_{\lfloor nt \rfloor}}{\sqrt{n}}$$

The Donsker's theorem provides a strong justification of this approximation: the process $W_t^{(n)}$ as a random variable on the Skorokhod space $\mathcal{D}([0,T])$ will converge in law toward a standard Brownian Motion, when time increment goes to zero, i.e. when $n \to \infty$.

Parametrization

How to choose the trajectories $(A_t^{k_m})_{t\in[t_m,t_{m+1}]}$ and the transition probability $\pi_{k_{m+1}|k_m}$? Here, we take advantage of the fact that the increment of the random walk follows a normal distribution. Hence, one should refer to the "Gaussian case" in the literature on optimal quantization, cf. Pagès and Printems (2003) and Pagès (2017).

Concretely, thanks to the database on the website http://quantize.maths-fi.com, the procedure looks similar to clustering methods. One would choose the Voronoï quantization grid $\Gamma = \{a_1, \dots, a_K\}$ in order to minimize the L^2-mean quantization error $e(A_t, \Gamma) = \big|\big| \min_{\{a_t\}\in\Gamma} |A_t - a_t| \big|\big|_{L^2}$ between the random variable A_t and the nearest neighbor a_t in Γ. The problem consists thus in the minimization over all grids Γ of size K:

$$e_{K,L^2}(A_t) := \inf \left\{ e_{L^2}(A_t, \Gamma) \;\middle|\; \Gamma \subset \mathbb{R}, |\Gamma| \leq K \right\}$$

Now, considering this grid Γ, with the Voronoï partition $\big(C_k(\Gamma)\big)_{1\leq k\leq K}$ such that

$$C_k(\Gamma) \subset \left\{ z \in \mathbb{R} : |z - a_k| \leq \min_{1\leq j\leq K} |z - a_j| \right\}$$

one could compute the probability weight p_k by measuring the probability of the random variable falling into this particular Voronoi cell

$$p_k := \mathbb{P}\big(A_t \in C_k(\Gamma)\big)$$

Moreover, (a) the Brownian motion has stationary increment, one could apply the same method

[21] If one would approximate diffusions, e.g. Geometric Brownian motion or Ornstein-Uhlenbeck processes with bi/trinomial trees, one could also refer to the simulation by T. Montes, cf. http://simulations.lpsm.paris/trinomial_trees/

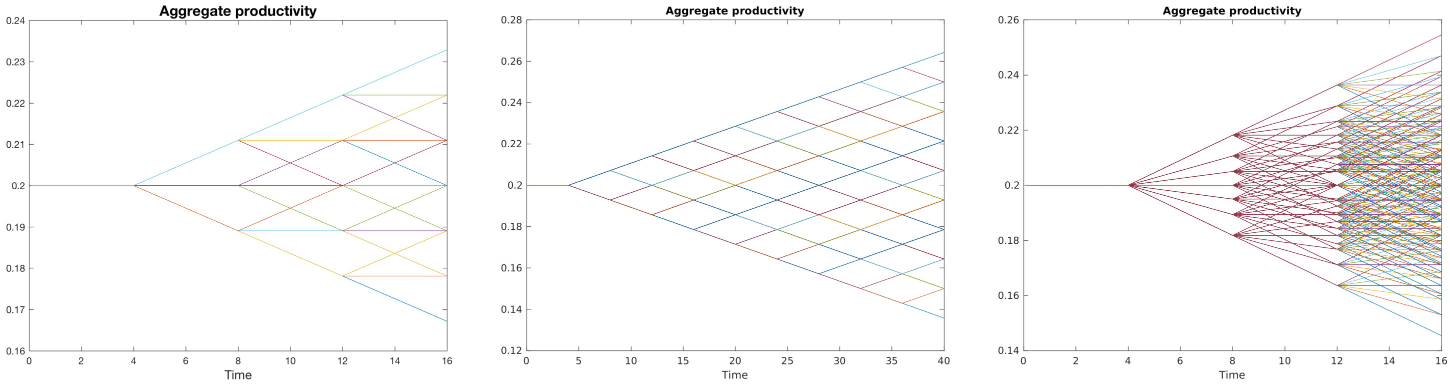
for every node and every level $1 \leq m \leq M$ of the tree and (b) the increments being independent, the probability of transition is simply $\pi_{k_{m+1}|k_m} = p_{k_{m+1}}$.

In the following, we will use several concrete examples:

(i) A simple tree, where $K = 3$ branches per nodes and $M = 3$ "waves" of shocks,

(ii) A "long" tree, where $K = 2$ and $M = 9$, where the time discretization is finer

(iii) a "large tree", where $K = 7$ and $M = 3$ the (space) quantization grid is finer



3.2.2 Quantization tree

While in the previous section we used optimal quantization for every increment of the tree, which were gaussian, we now consider quantization trees for the *whole* process. One could refer to Bally, Pagès, and Printems (2005). The main idea is approximate the process using a skeleton of the distribution, supported by a tree whose branches are not identical anymore.

We again consider grids with a finite number of points, say $N = K \times M$ made again of M "waves"/"layers" and K branches per layer. The objective is to find a grid such that we can approximate the process A_t by a collection of grids $\widehat{A}_{t_m} := \text{Proj}_{\Gamma_m}(A_{t_m})$ with $\Gamma_m = \{a_1^m, \ldots, a_K^m\}$ a grid of size K. The main difference is that the grid Γ_m is chosen in order to approximate optimally the stochastic process A_t at time t_m and thus account for the underlying structure of A_t.

Parametrization

When A_t is Brownian, the scaling property tells us that $A_t \sim \frac{1}{\sqrt{t}} B_1$ with B_1 a gaussian r.v. with variance 1. Therefore, at time t_m one could chose the grid Γ_m using the exact same method as above, this time rescaling appropriately by $\sqrt{t_m}$ at each layer. The Voronoï partition is thus a collection of sets $\left(C_k(\Gamma_m)\right)_{1 \leq m \leq M}$.

However, computing the transition probability requires to use the Bayes formula.

$$\pi_{k_{m+1}|k_m} = \frac{\mathbb{P}\Big(A_{t_{m+1}} \in C_k(\Gamma_{m+1}) \ \& \ A_{t_m} \in C_k(\Gamma_m)\Big)}{\mathbb{P}(A_{t_m} \in C_k(\Gamma_m))}$$

A first order approximation would be to consider $A_{t_m} = a_{k_m}$ instead of $A_{t_m} \in C_k(\Gamma_m)$. This way the transition probability becomes :

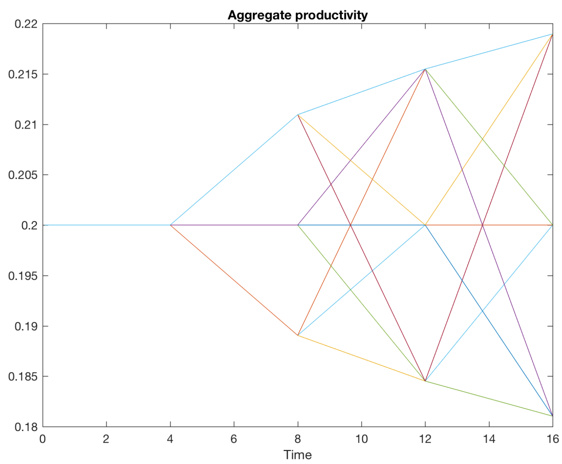
$$\pi_{k_{m+1}|k_m} \approx \mathbb{P}\Big(\mathcal{N}(a_{k_m}, t_{m+1} - t_m) \in [a_{k_m - \frac{1}{2}}; a_{k_m + \frac{1}{2}}]\Big)$$

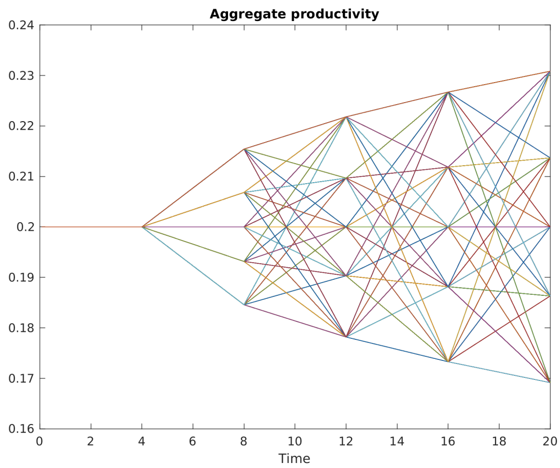
where $C_{k_m}(\Gamma_m) := [a_{k_{m+1} - \frac{1}{2}}; a_{k_{m+1} + \frac{1}{2}}]$ when working in one dimension.

Therefore, we will again display two examples:

(iv) A simple tree, where $K = 3$ branches per nodes and $M = 3$ "waves" of shocks. Note that the variance slows down with time since we know rescale to account for the low probability of extreme values

(v) A larger tree, where $K = 5$ and $M = 4$, where the time and space discretization is slightly finer.





3.2.3 Relative advantages and computational challenge

The main difference between the two types of trees, "K"-nomial vs. quantization trees, is the way to approximate the random variable in space. The quantization tree will keep the same grid for all the trajectories: the tree will be *recombining*. However, some trajectories will have very low probability – e.g. when the brownian reaches the worst state on the 3rd wave after having peaked at the best level on the 2nd wave.

In contrast, the "K"-nomial tree will not be recombining and thus, the space discretization will be made finer with time. However, the space of trajectories is restricted to process that have low probability of extreme values: the discretized common noise will have a lower variance than its equivalent quantization tree. This drawback can be compensated by increasing M the discretization in time and K in space. The main issue will be that increasing both value M and K will be computationally difficult as we explain below.

In these two cases, the interesting experiment will be to see that the MFG equilibrium strongly depends on the past values of the measure: two equilibria with the same value A_{t_m} might differ if the past values $A_{t_{m-1}}$ were different. This past value would indeed affect the evolution of the measure – which is backward looking and thus path-dependent.

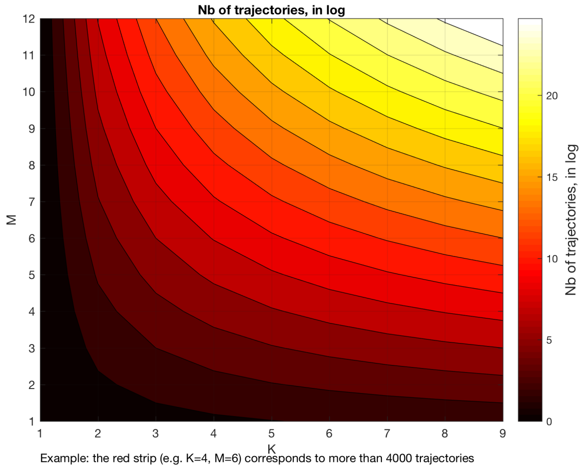
This is a concrete sign that the MFG equilibria is not Markovian in $x \in \mathbb{X}$. However, it will Markovian when enlarging the space to include the infinite-dimensional object m, and thus on $\mathbb{X} \times \mathcal{P}(\mathbb{X})$. This observation justify to focus on the MFG equilibrium through the Master equation in section 3.1.2

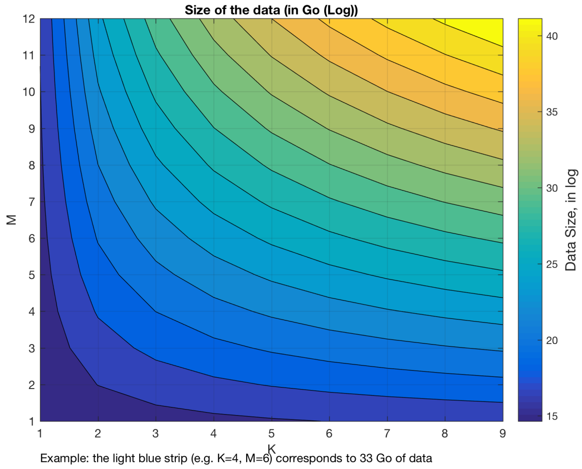
Computational challenge

The main common point between these two types of trees is the computational difficulty of simulating the stochastic system of forward-backward PDEs.

Indeed, the tree structure complexity increases exponentially due to the size K^M of the discretization grid. This has two effects: (i) it increases the number of trajectories K^M one has to simulate, the number of both HJB and FP equations to solve, and thus the time of computation, and (ii) it increases the size of the data, since the value function and the measure are array in four dimension: $[0, T] \times \mathbb{X}$ and Ω the space supporting the underling common noise. This memory therefore increases when the discretization is made finer.

The following pictures shows the explosion (since it is displayed in log value) of number of trajectories (LHS) and the size of the memory (RHS).





These two difficulties can been tackles by different arrangements:
(i) The length of computation can be shortened by parallelizing the algorithm. Indeed, the tree structure and the backward-forward nature of the algorithm allow to compute different trajectories on different cores of computer.
(ii) The memory used by the data can be reduced by computing the (stochastic) infinitesimal generator of the controlled diffusion "à la volée" (on the run). First one would solve the value function on the different trajectories. Without saving the heavier set of (still sparse) matrices representing the operator of HJB –needed to compute the F.P. equation – recover the controls and the drift "on the run" at each step of the Fokker Planck equation[22].

3.3 Potential extensions

We will now use some intuitions gained from the section section 3.1 and section 3.2 to propose two novel methods that could bypass some difficulties inherent to the method developed above. A first extension would be to use recent quantization methods to handle the Forward-Backward system, and the second would be to reduce the problem to finite-dimension, using projection methods.

3.3.1 Forward-Backward system and functional quantization

In mathematical finance, the common way to solve stochastic problems – e.g. computing option prices, optimal portfolio allocation etc. – is to use methods such as Monte-Carlo, Quasi Monte Carlo or quantization. According to researchers of the field[23] optimal quantization is

[22] However, this method reveals more difficult than it seems due to the specific upwind scheme we use in the HJB equation as we will explain in appendix.
[23] Cf. Pagès (2017)

one of the most efficient method for such problems that can be solved backward. Indeed, when the problem is Markovian, dynamic programming yields a formulation with a HJB equation or an expectation $\mathbb{E}_t[f(X_T)]$ where X_t solves a SDE. The idea is always to start from the future and compute the value backward in time, updating the conditional expectations.

However, the mean field game system with common noise is no longer Markovian in the state $X_t \in \mathbb{X}$. The future transition of the system (t, X_t, m_t) will depend on past shocks and is no longer memoryless. This is due to the fact that the Fokker Planck equation – which is computed forward in time – is coupled with the HJB. The equilibrium fixed point has to be found on the whole trajectory and for each trajectory (whose number can be large).

In the previous section, when handling common noise, we used "K"-nomial and quantization trees in a different way compared to the usual methods developed in numerical probability. In our problem, we used the trees *in the two directions*: for backward computation (HJB, updating conditional expectations) *and* in the forward computation (FP) where the whole path of the trajectory is needed.

A contrario, practitioners of the field in mathematical finance would use other methods to compute the price of "path-dependent" option – e.g. Asian options, or lookback options – rely on *functional quantization*. In these cases, the set of trajectories of the Brownian motion is reduced to a finite number of trajectories chosen optimally, as shown in the figure below for different example. The idea is to consider a stochastic process as a random variable in L^2 and to find the optimal grid – i.e. the optimal Voronoï cells – in this infinite-dimensional space.

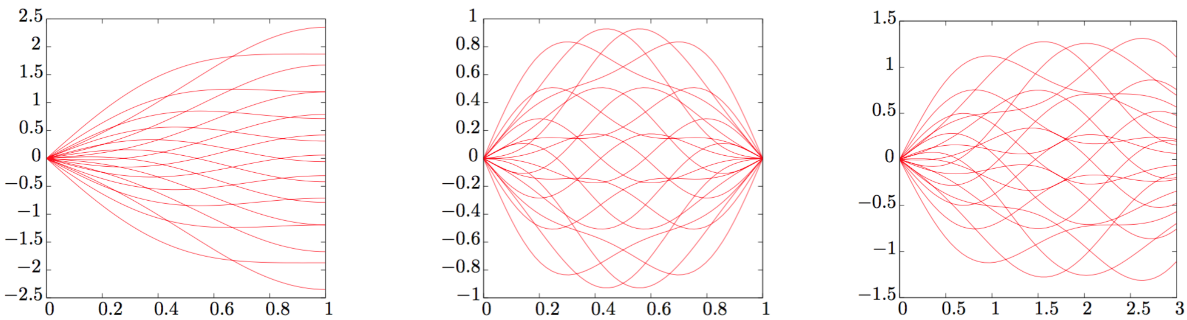


Figure 1: Functional quantization in the case of Brownian Motion (LHS) Brownian Bridge (Middle) and Ornstein-Uhlenbeck process (RHS), cf. Corlay (2011)

In the case of MFG with common noise, to avoid the use of a growing number of trajectories in the case presented before, functional quantization procedure could be used to reduce the number of trajectories for the common noise. Using a large but fixed number of such trajectories (instead of a exponentially increasing number), one could compute the evolution of the measure and the prices r_t and w_t. However, this method would be completely deterministic: for a given trajectory, the evolution of the system would not be subject to any uncertainty, i.e. change in the common noise.

A potential solution would be to associate this method with standard quantization/"K"-nomial trees for the backward computation. There, the branching of the tree would allow to compute conditional expectations and the system would be adapted to the filtration generated by the common noise.

The general idea would be is the following:

1. Build a "K"-nomial trees/ quantization tree for the common noise of dimension K^M (K branches, M layers) denoted A_k^{tree}, $1 \leq k \leq K^M$. The transition probability is expressed in the hypermatrix Π_{t_m} $\forall 1 \leq m \leq M$.
2. Choose a set of functional quantization trajectories $N < K^M$ denoted $A_j^{fq}, 1 \leq j \leq N$
3. For each layer t_m, "match" each branch of the tree A_k^{tree} to a portion of the closest quantization trajectory A^{fq}, using the L^2 norm for instance[24]:

   $A_j^{fq,\star}(k) \in \text{argmin} \left\| A^{fq} - A_k^{tree} \right\|_{L^2([t_m, t_{m+1}])}$
4. Compute the HJB backward on each portion of trajectory to get the value function v_t. At each time t_m, "update" the terminal condition of the HJB to account for the probability of "changing" trajectory using Π_{t_m}
5. Compute the FP forward on the whole trajectory to obtain the measure m_t.
6. Update the fixed point[25].

In practice, to effectively quantize a Gaussian Processes as the ones proposed above, one need to compute the Karhunen-Loeve basis of the process cf. Luschgy and Pagès (2002) or Corlay and Pagès (2015). Using the method to approximate any of these process and with any number N we believe this method could provide substantial advantages to limit the computational challenge posed by this model.

3.3.2 Master equation and projection method

In this section we will use the logic developed in the section 3.1.2, concerning the master equation with aggregation. The idea would be – as in Krusell-Smith – to describe the measure m_t on $\mathbb{X}$ by a finite set of moments I, as $K^i = \langle h_i, m \rangle, \forall 1 \leq i \leq I$, and where h_i is an arbitrary function, but could be $h_i(x) = x, x^2, x^3$ etc. Note that one could also consider a projection on a basis : $\mathcal{P}m_t(x) = \sum_i \langle h_i, m \rangle h_i(x)$ and obtain an analogous result.

[24] If the average distance between the tree and the quantized trajectories is too large, increase N or change the norm used in the Nearest Neighbor method.

[25] Note that the fixed point procedure may not converge completely right after the revelation the shocks at t_{m^+} since there might a difference between the original tree A_k^{tree} which is continuous and the new "patchwork" $A_j^{fq,\star,k}$ obtained by Nearest Neighbor that hence be discontinuous

Therefore, extending the logic seen before, we could write the solution of the master equation as a finite-dimensional object:

$$U(t,x,m) = \overline{U}(t,x,K^1,\dots,K^i,\dots K^J) = \quad \overline{U}(t,x,\langle h_1,m\rangle,\dots,\langle h_J,m\rangle)$$

Yielding, as before: $$D_m U(t,x,m;y) := (\partial_{m_1}U, \partial_{m_2}U, \partial_{m_3}U)^T$$

$$\partial_{m_j}U(t,x,m;y) = \sum_{i=1}^{I} \partial_{K^i}\overline{U}(t,x,\{K^i\}_i)\partial_{y_j}h_i(y)$$

However, we now have a method to approximate the integral terms of the type

$$\int_{\mathbb{X}} G(y)\, m(dy) \equiv \sum_i \langle h_i, m\rangle \int_{\mathbb{X}} G(y)h_i(y)dy$$

Applying the above techniques and calculus, the master equation with common noise, with projection on the basis $\{h_j\}_j$ and $K^i = \langle h_i, m\rangle$ can be written:

$$\begin{cases} -\partial_t\overline{U} + \rho\,\overline{U} - \overline{\mathcal{H}}(\cdot, D_x\overline{U}) - \sigma^{idio}\partial^2_{x_2}\overline{U} - \sigma^{com}\partial^2_{x_3}\overline{U} \\ \qquad\qquad - \displaystyle\sum_{i=1}^{I}\partial_{K^i}\overline{U}\ \sum_{j=1}^{I}\int_{\mathbb{X}}\sigma^{idio}\partial^2_{y_2}h_i(y)\ + \sigma^{com}\partial^2_{y_3}h_i(y)\,K^j h_j(y)dy \\ \qquad\qquad + \displaystyle\sum_{i=1}^{I}\partial_{K^i}\overline{U}\ \sum_{j=1}^{I}\int_{\mathbb{X}}\nabla h_i(y)\cdot D_p\overline{\mathcal{H}}(\cdot,y,\cdot,D_x\overline{U}(\cdot,y,\cdot))\,K^j h_j(y)dy \\ \qquad\qquad -2\sigma^{com}\displaystyle\sum_{i=1}^{I}\partial_{x_3}\partial_{K^i}\overline{U}\ \sum_{j=1}^{I}\int_{\mathbb{X}}\partial_{y_3}h_i(y)\,K^j h_j(y)dy = 0 \qquad\qquad \text{in } [0,T]\times\mathbb{X}\times\mathbb{R}^I \\ \overline{U}(t,x,\{K^i\}) = v^\infty(x,m) \qquad\qquad\quad \text{in } \mathbb{X}\times\mathbb{R} \quad \text{with } \ \langle h,m\rangle = K^{h,m} \end{cases}$$

where $\overline{U}$ is taken in $(t,x,\{K^i\}_i)$ and recalling that the Hamiltonian is

$$\mathcal{H}(t,x,m,D_xv) = \sup_\alpha u(\alpha) + b(x,m,\alpha)\cdot D_xv \equiv \sup_c u(c) + s(a,z,A,m,c)\partial_a v + b^{idio}(z)\partial_z v + b^{com}(A)\partial_A v$$

using the notation of the economic problem.

We therefore see that this problem can be reduced to finite dimension if we expect some form of dependence with respect to the moments of the measure. If one search to increase the number of moment/the size of the basis $\{h_i\}_i$ the PDE can still be difficult to compute. However, one could rely on Monte Carlo method for PDE[26]

[26] One reference among many could be Fahim, Touzi, and Warin (Fahim et al.) for instance.

4 Applications and Economic results

In this section, we will now cover four different models:

(1) the Krusell-Smith model

(2) an extension with endogenous labor supply

(3) the one-asset H.A. New Keynesian model (HANK)

(4) the two-asset H.A. model à la Kaplan-Moll-Violante.

In the following, after a brief summary of the main equations and equilibrium conditions, we cover the numerical scheme. An overview is summarized in the first Krusell-Smith model, while the complete description is given in appendix section B. This explanation is partly based on the description made in Achdou, Han, Lasry, Lions, and Moll (2017) and its "Numerical Appendix", for the case without common noise. However, in the text, we consider different way to handle the state-constraint for the extensions (2)-(4).

Afterward, we provide results of the model :

- with a Brownian common noise $A_t = \bar{A}_0 + B_t$.
- with a common noise following a "Jump-Drift" process $d\widetilde{A_t} = -\theta(\widetilde{A_t} - \widetilde{A_0})dt + \ \varepsilon dN_t$ where dN_t is a jump process ($dN = 1$ with intensity λ) and $\varepsilon \sim \mathcal{N}(0, \sigma^2)$. We consider such process to provide economic intuitions. It allows to plot the Impulse Response Functions (IRF) after a one-time temporary deviation from steady-state.
- we then compare the previous results with standard IRF of the representative agent model (i.e. single player in the game), simulated using Dynare

4.1 The Krusell-Smith model

Recall the household control problem at the basis of the Aiyagari-Bewley model:

$$\max_{c_t} \mathbb{E}_0 \int_0^\infty e^{-\rho t} u(c_t) dt$$

$$\text{subject to :} \qquad da_t = (z_t w_t + r_t a_t - c_t)dt \qquad \text{(Budget constraint)}$$

$$\text{and} \qquad a_t \geq \underline{a} \qquad \text{(Credit constraint)}$$

The supply-side rewrites:

$$K_t(r_t) := \left(\frac{\alpha A_t}{r_t + \delta}\right)^{\frac{1}{1-\alpha}} z_{av} \qquad \text{(Capital demand)}$$

$$w_t = (1-\alpha)\, A_t\, K_t^{\alpha}\, z_{av}^{-\alpha} \qquad \text{(Wage)}$$

$$r_t = \alpha\, A_t\, K_t^{\alpha-1} z_{av}^{1-\alpha} - \delta \qquad \text{(Interest rate)}$$

The MFG formulation is given by:

$$\rho\, v_j(t,a) = \partial_t v_j(t,a) + \max_c u(c) + \partial_a v_j(t,a)\, s_j(t,a) + \lambda_j(v_{-j}(t,a) - v_j(t,a)) \quad \text{on} \quad [0,T]\times[\underline{a},\infty)\times\{j_1,j_2\} \qquad \text{[HJB]}$$

$$0 = \partial_t g_j(t,a) + \frac{d}{da}[s_j(t,a)\, g_j(t,a)] + \lambda_j g_j(t,a) - \lambda_{-j} g_{-j}(t,a) \quad \text{on} \quad [0,T]\times[\underline{a},\infty)\times\{j_1,j_2\} \qquad \text{[FP]}$$

$$S(t,r) := \int_{\underline{a}}^{\infty} a\, g_1(t,a)da + \int_{\underline{a}}^{\infty} a\, g_2(t,a)da = K(t,r) \qquad \text{[Market clearing]}$$

$$s_j(t,a) = z_j w_t + r_t\, a - c_j(t,a) \qquad c_j(t,a) = (u')^{-1}(\partial_a v_j(t,a)) \qquad \text{[FOC]}$$

$$v_j(T,\cdot) = v_j^{\infty} \qquad g_j(0,\cdot) = g_j^0 \qquad \partial_a v_j(\underline{a}) \geq u'(z^j w + r\underline{a}) \qquad \text{[Boundary conditions]}$$

4.1.1 Algorithm: an overview

The objective is to find equilibrium of the MFG i.e. the value function v_j and the measure g_j $(j = 1, 2)$ solving the two PDEs and the interest rate r clearing the market.

A complete description can be found in appendix, section B. We summarize here the general method to find the equilibrium fixed point, iterating on ℓ:

1. Guess interest rate r^ℓ, compute capital demand $K(r^\ell)$ & wages $w(K)$
2. Solve the HJB using finite differences (semi-implicit method): obtain the controlled drift $s_j^\ell(a)$ and then the value function v_j^ℓ, by solving a system of sort:

$$-\frac{\mathbf{v}^{n+1} - \mathbf{v}^n}{\tau} + \rho\, \mathbf{v}^{n+1} = \mathbf{u}(\mathbf{v}^n) + \mathbf{A}(\mathbf{v}^n; r)\mathbf{v}^{n+1}$$

   - In the stationary equilibrium, solve for v_∞. In the transition case, compute the path of v^n, starting from the *terminal* condition $v^N = v_\infty$ and iterating *backward.*

3. Using $\mathbf{A}^T$, solve the FP equation, via the finite diff. system: $\frac{\mathbf{g}^{n+1} - \mathbf{g}^n}{\tau} + \mathbf{A}(\mathbf{v}^n; r)^T \mathbf{g}^{n+1} = 0$, and obtain g_j

   - In the stationary equilibrium, solve for g_∞. In the transition case, compute the path of g^n, starting from the *initial* condition $g^0 = g_0$ and iterating *forward.*

4. Compute the capital supply $S(\mathbf{g}, r) = \int_{\underline{a}}^{\infty} a\, g_1(a)da + \int_{\underline{a}}^{\infty} a\, g_2(a)da$
5. If $S(r) > K(r)$, decrease $r^{\ell+1}$, updating (using bisection method, and conversely, and come back to step 2.
6. Stop if $S(r) \approx K(r)$

Duality: Solving the FP using the HJB

In this problem, the FP is the adjoint equation of the HJB equation. As a result, the operator matrix in the Fokker Plank equation is the transpose (A^T) of the HJB operator matrix A – *taking optimal consumption as given.* Therefore, the most important would be to solve the HJB equation numerically and then to solve FP by "simply" transposing the matrix A.

Common noise

The algorithm is rationalized and explained above in section 3.2: it consists in choosing an appropriate tree structure to approximate the common noise A_t. On this tree representing K^M trajectories – M waves of shocks and K branches for each shock – one would solve the time-varying MFG equilibria, using the right boundary conditions:

- A terminal condition for the HJB: $\mathbf{v}^{k_N} = \sum_k \pi_{k_{N+1}|k_N} \mathbf{v}^{k_{N+1}}$
- An initial condition for the FP: $g^0 = g_0$ and solve the FP on the full trajectory on $[0, T]$

The equilibrium on one branch of the tree, with a deterministic trajectory of A_t, is exactly the one described above.

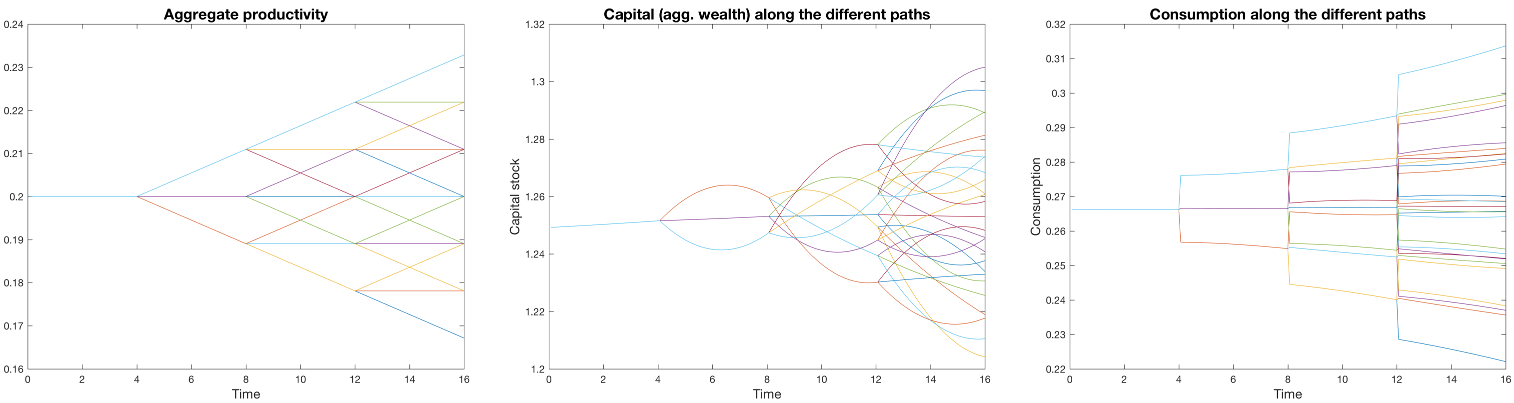
4.1.2 Results 1 – Discretized Brownian common noise

We presents the simulation of the MFG system with common noise of the "simple tree" (i) in the following, to provide intuitions and economic results. The trees (ii)-(v) described in section 3.2 are displayed in appendix section C.

The evolution of aggregate variables

The main interest of macroeconomists is to know whether microeconomic heterogeneity can have a impact on aggregate variables. In the following graph we plot the evolution of the main variables – the capital supply, consumption and saving, and prices: interest rate and wages – along each branch of the tree of productivity.

Note that the colors of the trajectories are linked between the different graphs: for instance the sky blue line corresponds to the "best scenario" of productivity on all the plots.



As a result of the continuity of the wealth distribution g in time, the capital stock adjust continuously, unlike the consumption and saving that feature jumps at each nodes of the tree. In

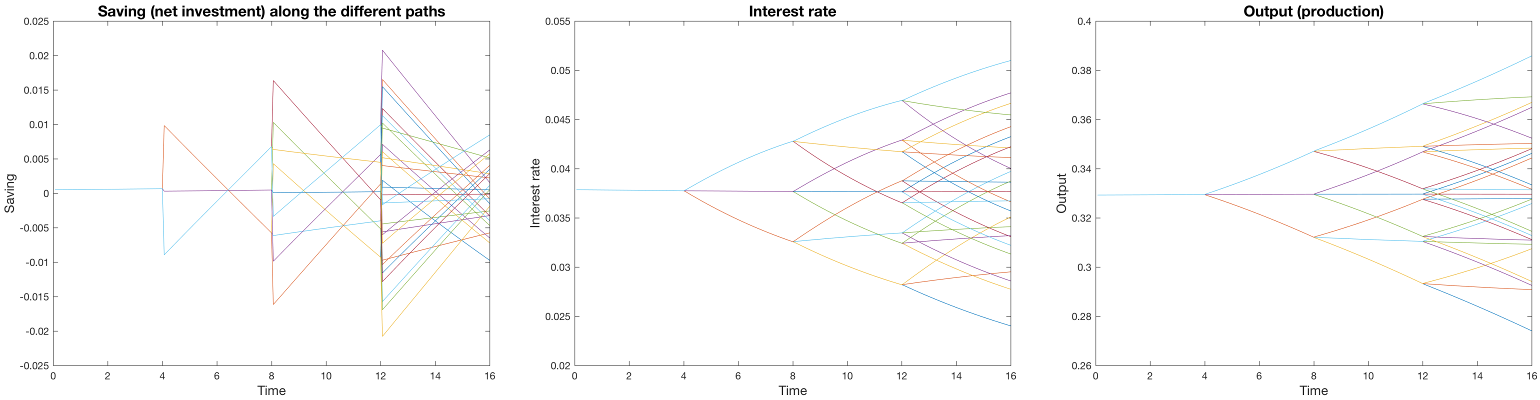


Figure 2: Evolution of aggregate variables with a Brownian TFP in the Krusell-Smith model

particular, saving will adjust instantaneously at each node of the tree, and this for two reasons: (i) at the revelation of the common noise, there will be a shift in expectations of the households, and during the boom, they will reduce their precautionary saving, decreasing the demand for asset/capital and raising the interest rate.
(ii) since the productivity rises over time, the capital will be more productive and the demand for capital from firms will slowly increase as well, raising interest rate and wages. This will mechanically increase the disposable income, saving and investment.

The first motive is instantaneous but typical of heterogenous agents model while the second is revealed with the change in productivity and prices and is similar to standard Brock-Mirman dynamics. The shift from the first to the second will change the trend of capital evolution, from a drop to a surge during boom periods.

The value function and wealth distribution as a function of time

As explained above both the value function and the measure are now random variables, which should be adapted to the filtration generated by the common noise $\sigma(B_t^0)$. Since these objects are infinite dimensional and "functions" of the common noise (i.e. adapted to the "alea" generated by this common noise), the task to display the evolution might not be obvious.

With our discretization procedure, we choose to use the tree (ii) as described in section 3.2, with $K = 2$ and $M = 9$ to make the result of the evolution of v_t and m_t clearer. Note that the simulation for this tree (ii) and the evolution of the aggregate variables is displayed in appendix in section C.

The probability distribution of the value or measure would be $\mathbb{E}_t\big(v_t|\sigma(B_t^0)\big) \approx h^v(B_t^0)$ and $\mathbb{E}_t\big(v_t|\sigma(B_t^0)\big) \approx h^m(B_t^0)$, where $h^v : \Omega_t \mapsto \mathcal{C}(\mathbb{X})$ and $h^m : \Omega_t \mapsto \mathcal{P}(\mathbb{X})$, where Ω_t is heuristically the probability space supporting the common noise at time t.

These random variables have value in $\mathcal{C}(\mathbb{X})$ or $\mathcal{P}(\mathbb{X})$, where $\mathbb{X}$ is in the 3 dimensions. Instead of plotting one dimension at a time, we choose to display two trajectories ω_t :

- The "best case" scenario, where aggregate productivity only increases at each tree node
- The "worst case" scenario, where the economy is only subject to TFP contraction

First of all, let us show the value function and the measure at the initial point of the game, i.e. at $t = 0$ and at the end of the game – when the randomness is switched off as explained in our remark at the beginning of section 3.1 – i.e. at $t = T$ in these two scenarios for productivity.

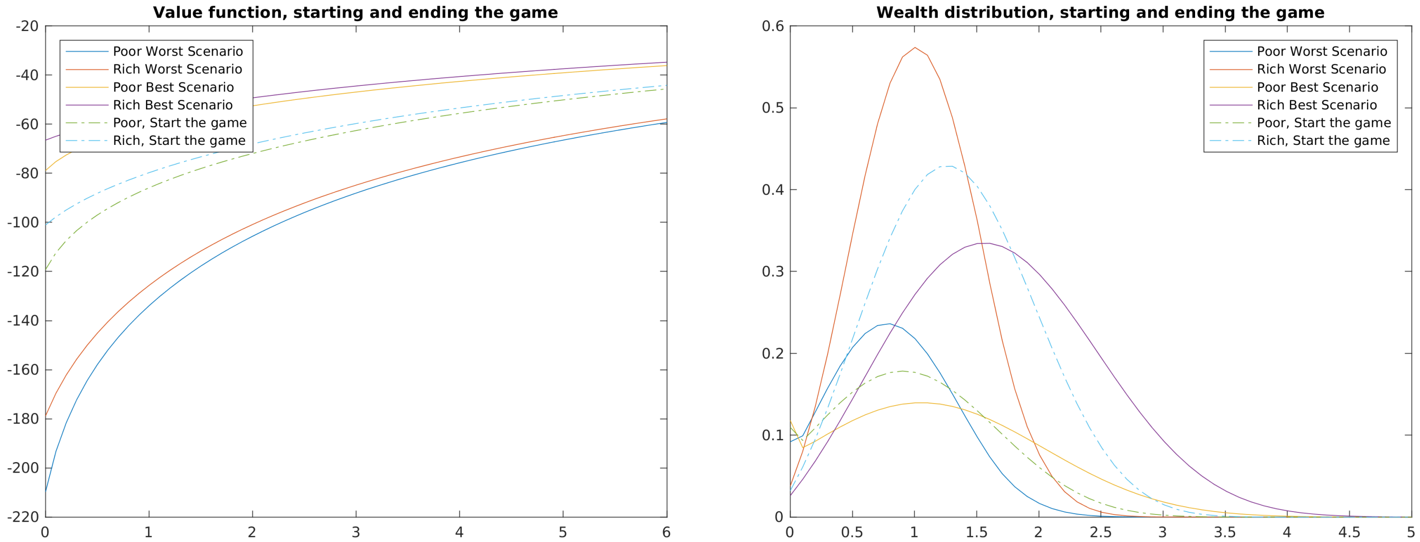


Figure 3: Value functions and distribution of both type of agents at the time boundaries

We see the aggregate productivity distorts both the measure of the agents and shift upward or downward the value function for both high and low income agents. The mechanism is intuitive: the value increases in boom and decreases in contraction, and the measure shifts to the right – people become richer on average – in boom, and shift to the left – concentrate more on low wealth – when economy is depressed. We now describe more precisely the underlying mechanisms for these two trajectories.

As described before, for each 'wave' of shock $m \in \{1, \ldots, M\}$, there will be a change of information about the productivity A_t between t_m^- and t_m^+. Therefore, the value function will be discontinuous at this point t_m to account for the revelation of the shock: this jump in the value function is linked to the jump in consumption and saving since the agents will evaluate differently the need of precautionary saving – depending of the probability of future shocks, thus changing the conditional expectation. For the two scenarios of productivity, the next two graphs display the evolution of the value functions respectively for poor agents (low income $j = 1$) on the LHS and rich agents $j = 2$ on the RHS:

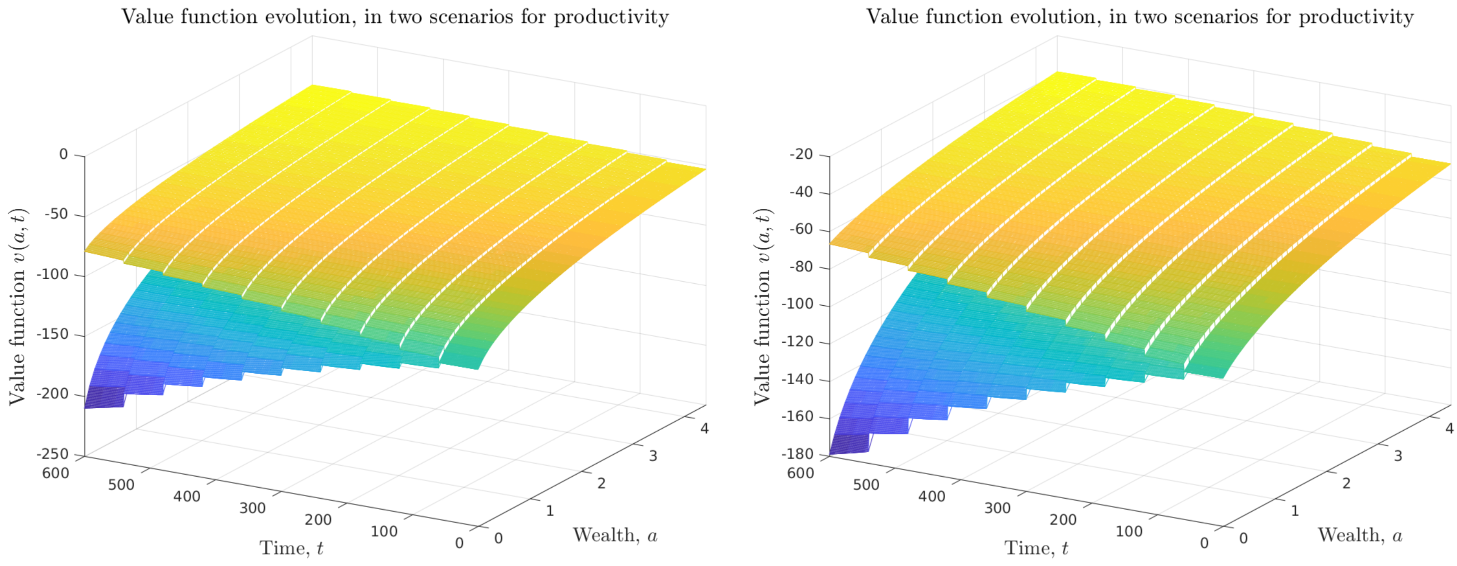


Figure 4: Value functions both type of agents in two trajectories of agg. productivity over time

Note that the scale may change the shape of the graphs: despite having different state z_j the poor and rich agents have almost similar value function, as displayed in the fig. 3

On the contrary, the wealth distribution is continuous over time. The control and the controlled drift, $b(t,x,m,\alpha^\star) = -D_p\mathcal{H}(t,x,m,\nabla_x v)$, will also jump – as function of the value function – but the measure will have a smooth evolution as only the infinitesimal variation in the Fokker Planck will distort the shape of the distribution.

In the next four graphs, the wealth distribution is represented as a function of time for the two scenario: best case (LHS) and worst-case scenario (RHS), for the poor agent (first line), or the rich (second line). Note that the wealth distribution shift leftward in situation of recession (decrease in productivity), since the income in mechanically lower: both wages and interest rate are decreasing due to the structural change in productivity A_t.

However, we claim that – due to precautionary saving – this leftward shift is attenuated by the willingness of both rich and poor agents to accumulate relatively more assets. Indeed, they are willing to self-insure against potential risk of further recessions or risk of falling (or staying) into the low income state. This mechanism changing the control will have a smoothing effect on capital evolution – the first moment of the distribution – and the evolution of consumption and saving will have general equilibrium effects on wages and interest rate, as we explain next.

Note that the effect is qualitatively reversed when the productivity increases – in the best case scenario – but the quantitative result is not symmetrical, since the MFG with common noise is a highly non linear system.

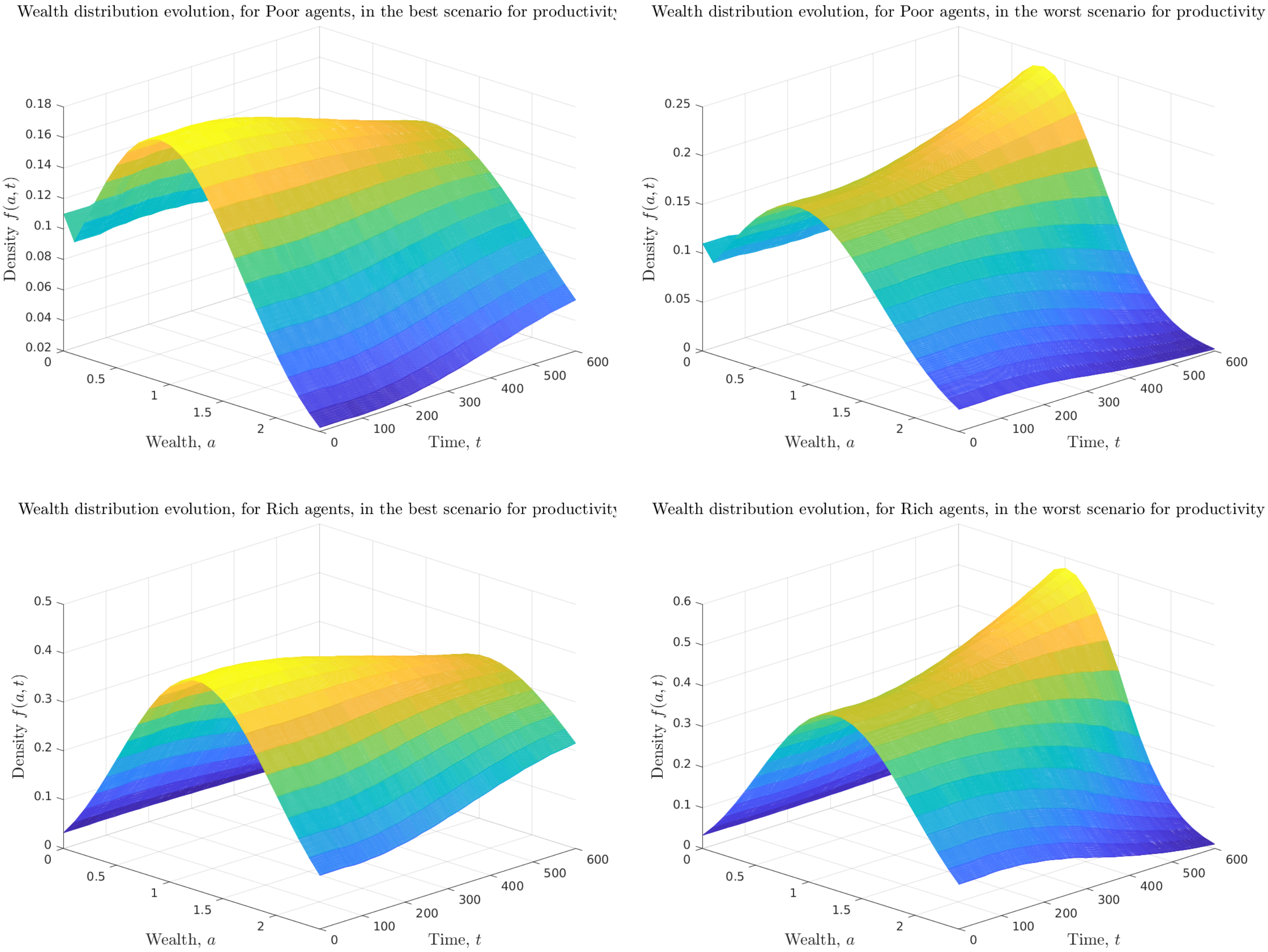


Figure 5: Wealth distribution of both type of agents, in two trajectories of agg. productivity over time

To provide explicit argument for our claim – that precautionary saving will have smoothing effect on capital – we know display the evolution of the control – $c_t^\star$ and the controlled drift $s_t = b(t, a, j, \bar{m}, c_t^\star)$ for the low income agents for the worst-case scenario.

Note that the mechanism is similar for rich agents and reversed in the case of economic growth (increase in A_t). To display the clear mechanism, we come back on our "simple" tree (i), where there is only $M = 3$ shocks, and the dynamics are more obvious.

As explained above in the description of the aggregate variables, the revelation of the common noise change abruptly the behavior of household: in recession, the saving jumps up and the consumption jumps down. This effect is strong, even though the aggregate productivity have not changed yet – due to the continuity of the Brownian motion $A_{t_m^-} = A_{t_m^+}$.

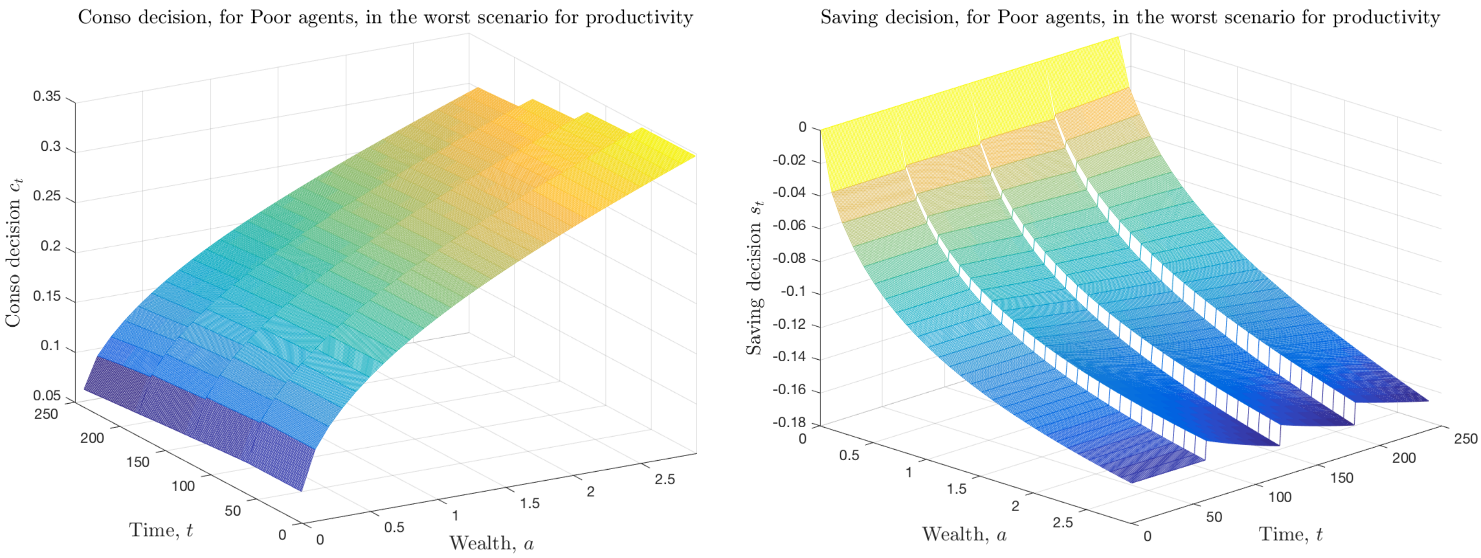


Figure 6: Control – c_t – and saving – s_t – over time for low income agents in the worst case scenario for productivity

However, the important underlying reason is the change in conditional expectation of future shocks: there is higher probability of falling is a future depression when the TFP is already low – thanks to the martingale property of the Brownian motion.

The effect of such jump will have general equilibrium effect: as displayed in fig. 2, the decrease in capital will be delayed – with capital increasing at first before decreasing subsequently – and the interest rate will drop more that what would be strictly implied by the change in productivity – almost 15% drop when the change in TFP in only 5%.

The transmission channels and the quantitative effects should be analyzed with Impulse Response Function (IRF) and that is the task we cover in the next section.

4.1.3 Results 2 – Jump-drift process and IRF

In this section, we consider the aggregate shock, affecting firm's productivity, that follows a "Jump Drift" Process. This stochastic process, resembling closely to an Ornstein Uhlenbeck process, is not a diffusion, but feature jumps and a mean-reverting behavior. Such dynamics will allow to plot the Impulse Response Functions (IRF) of the system. More precisely, the aggregate productivity and the deviation from steady-state follows:

$$A_t = \bar{A}_0 \, e^{\widetilde{A}_t} \qquad\qquad d\widetilde{A}_t = -\theta \widetilde{A}_t dt + \; \varepsilon \, dN_t$$

$$dN = 1 \quad \text{with intensity } \; \lambda \qquad\qquad \text{and } \; \varepsilon \sim \mathcal{N}(0, \sigma^2)$$

For this experiment, we plot one trajectories – among all the branches of the tree – that represents the dynamics of the economy after a one-time temporary deviation from steady-state, i.e. a jump $dN_t = 1$ with $\varepsilon = \sigma$. On the first plot (showing $\widetilde{A}_t$) we display the upper and lower bound of our approximation tree for this process. After the shock, the system is not perturbed and comes back to the steady-state value of the aggregate productivity.

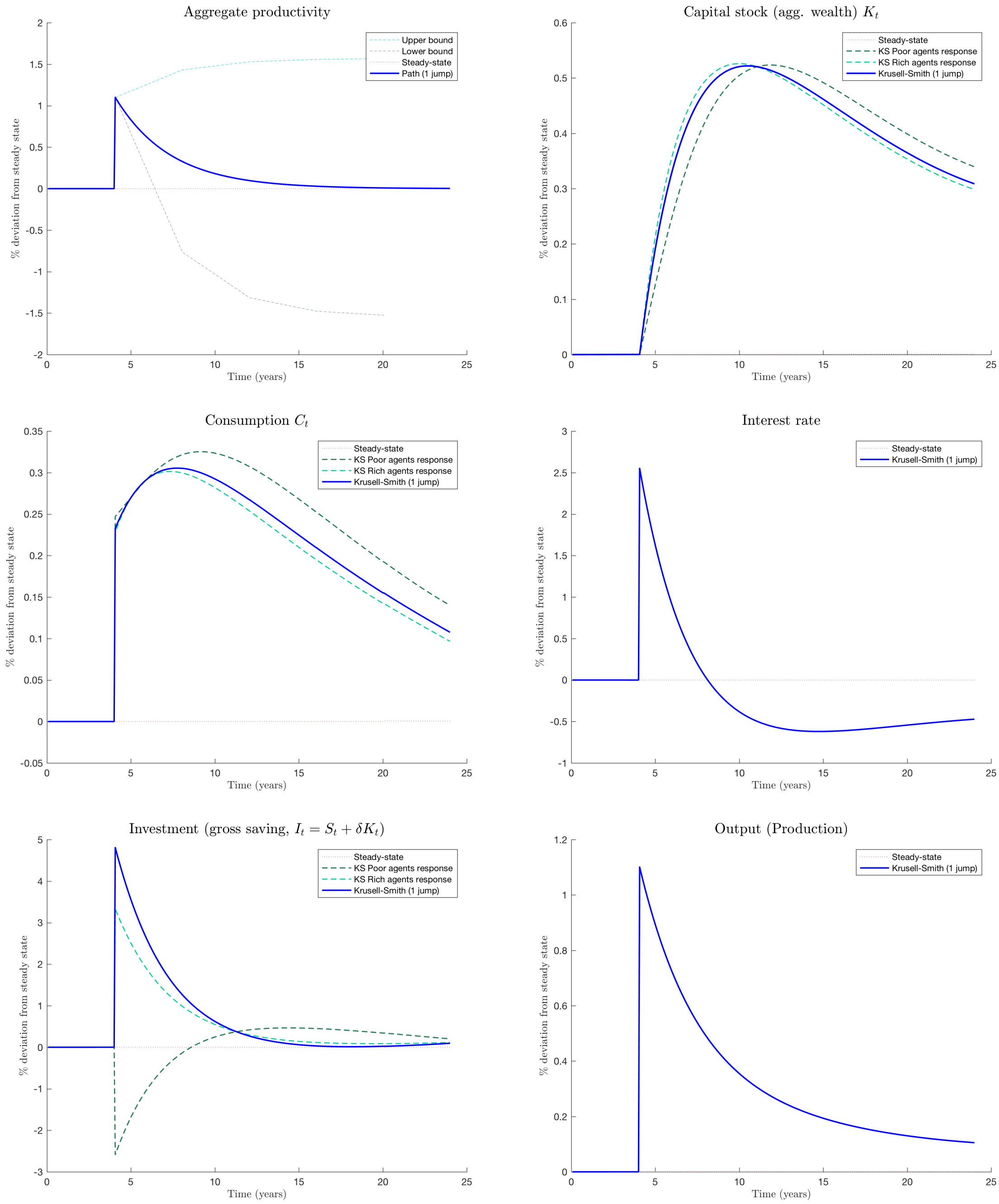


Figure 7: Impulse Responses to a one-time TFP shock in the Krusell-Smith model

An interesting feature, is the fact that poor agents, who are more likely to be credit-constraint and hand-to-mouth, are more willing to reduce their precautionary saving when the positive shock – while it is the reverse for rich agents – and increases relatively more their consumption than the rich (high-income) agents.

Moreover, the increase in capital decreases its marginal productivity over time, causing the interest rate to drop below its pre-shock level. This feature is standard in the Brock Mirman model. We now want to now if the two models can provide quantitative differences.

4.1.4 Comparison with representative agent model

In this section, we compare the previous model to the standard Brock Mirman model. We simulate the model using DYNARE (cf. the CEPREMAP working paper Adjemian et al. (2011)), a software solving rational expectations models using perturbation methods.

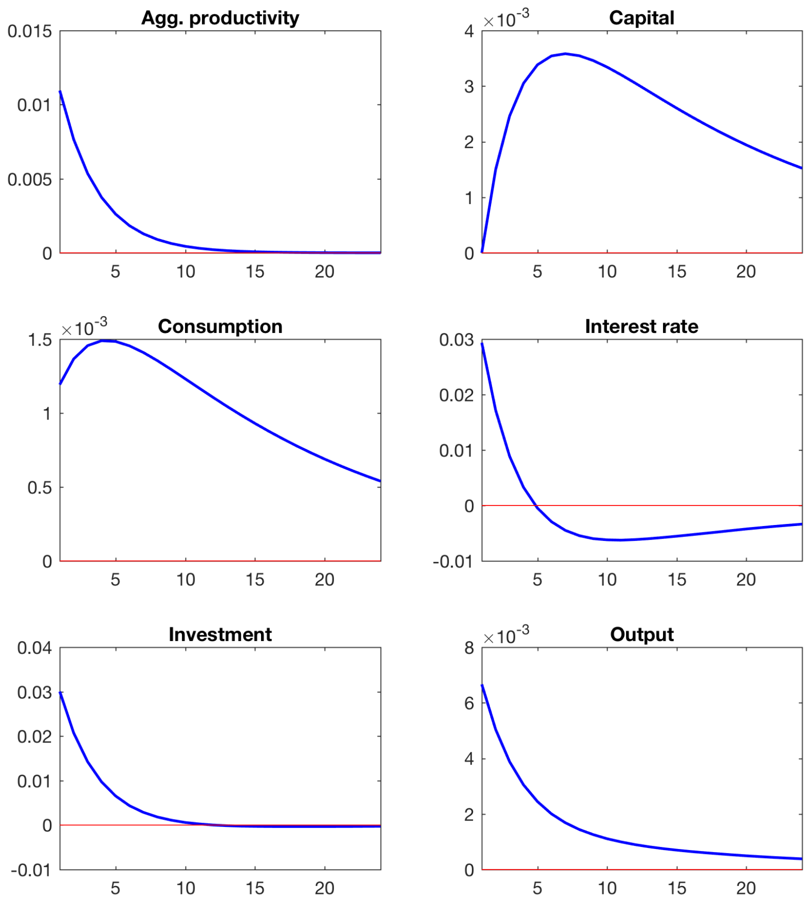


Figure 8: Impulse to a TFP shock in the Brock Mirman model

If the quantitative difference may not be obvious, note that the capital stock increase by 0.5 % in the Krusell Smith model compared to 0.3 % in the Brock Mirman. This is due to the larger increase in investment – caused by precautionary saving of the rich agents – that is almost twice as large, 5% compared to 3 % and that imply a stronger amplification effect of the productivity shock on output. Indeed the production reacts one to one to the 1.1% productivity shock, while in the Brock Mirman the effect is dampened by consumption smoothing (c.f. Euler Equation). Due to the presence of Hand-to-Mouth consumers, the reaction of consumption is also much stronger in the Krusell-Smith model, with 0.25% compared to only 0.125 % in the Brock Mirman model.

4.2 Endogenous labor supply

We now turn to the model with the Krusell Smith model with endogenous labor supply, described in section 2.2.1, which is the "Heterogeneous agents" model analogous to the Real Business Cycle model.

$$\begin{aligned}
&\max_{\{c_t\}_{t_0}^{\infty},\{\ell_t\}_{t_0}^{\infty}} \mathbb{E}_0 \int_0^{\infty} e^{-\rho t} u(c_t, \ell_t) dt \\
\text{subject to :}\quad & da_t = (z_t w_t \ell_t + r_t a_t - c_t) dt \qquad \text{(Budget constraint)} \\
\text{and}\quad & a_t \geq \underline{a} \qquad \text{(Credit constraint)}
\end{aligned}$$

Note that in this problem, there are two controls c_t and ℓ_t for two states variables a and z_j. Therefore the Hamiltonian rewrite:

$$\begin{aligned}
\mathcal{H}(p) &= \sup_{c,\ell} u(c,\ell) + (z_j w \ell - c)p \\
& \max_c[\bar{u}(c) - cp] + \max_\ell[-\tilde{u}(\ell) + z_j w\, \ell] \\
&= \mathcal{H}^c(p) + \mathcal{H}^\ell(p) \\
\text{with}\ \ \mathcal{H}^c(p) &= \max_c \bar{u}(c) - pc \qquad \text{and} \qquad \mathcal{H}^\ell(p) = \max_\ell -\tilde{u}(\ell) + w z_j \ell\, p
\end{aligned}$$

since $u(c,\ell) = \bar{u}(c) - \tilde{u}(\ell)$ is separable and maximisation problems in the Hamiltonian are both concave, thank to the preferences assumed (CRRA for consumption and Constant Frisch elasticity for labor).

$$\begin{aligned}
\partial_c \bar{u}(c) &= p &&\Rightarrow && c^\star = (\partial_c \bar{u})^{-1}\big(p\big) \\
\partial_\ell \tilde{u}(\ell) &= w z_j p &&\Rightarrow && \ell^\star = (\partial_\ell \tilde{u})^{-1}\big(w z_j\, p\big)
\end{aligned}$$

Despite that, the model is similar to the Krusell-Smith model in section 4.1.

Therefore, the HJB and FP equations are given by:

$$-\partial_t v_j + \rho v_j = \mathcal{H}^c(\partial_a v_j) + \mathcal{H}^\ell(\partial_a v_j) + r_t a\, \partial_a v_j + \lambda_j (v_{-j} - v_j) \quad \text{on} \quad [0,T]\times[\underline{a},\infty)\times\{j_1,j_2\} \qquad \text{[HJB]}$$

$$0 = \partial_t g_j + \frac{d}{da}[s_j\, g_j] + \lambda_j g_j - \lambda_{-j} g_{-j} \quad \text{on} \quad [0,T]\times[\underline{a},\infty)\times\{j_1,j_2\} \qquad \text{[FP]}$$

$$S(t,r) := \sum_j \int_{\underline{a}}^{\infty} a\, g(t,a,z_j)\, da = K(t,r) \qquad \text{[Mkt clearing, Capital]}$$

$$N(t,w) := \sum_j \int_{\underline{a}}^{\infty} \ell^\star(t,a,z_j,w)\, g(t,a,z_j) da = L(t,w) \qquad \text{[Mkt clearing, Labor]}$$

$$c_j^\star(t,a) = (\partial_c \bar{u})^{-1}(\partial_a v_j(t,a)) \qquad \ell_t^\star = (\partial_\ell \tilde{u})^{-1}\big(w_t z_j\, \partial_a v_j(t,a)\big) \qquad \text{[Optimality conditions]}$$

$$s_j(t,a) = z_j w_t \ell_t^\star + r_t\, a - c_j^\star(t,a) \qquad \partial_a v_j(\underline{a}) \geq u'(z^j w + r\underline{a})$$

$$v_j(T,\cdot) = v_j^\infty \qquad g_j(0,\cdot) = g_j^0 \qquad \text{[Boundary conditions]}$$

The supply side is exactly the same but now labor supply is endogenous:

$$Y_t = A_t K_t^\alpha L_t^{1-\alpha} \qquad \text{Production fct}$$

$$\frac{K_t}{L_t}(r_t) := \Big(\frac{\alpha A_t}{r_t+\delta}\Big)^{\frac{1}{1-\alpha}} = \Big(\frac{(1-\alpha)A_t}{w_t}\Big)^{-\frac{1}{\alpha}} \qquad \text{Capital/Labor ratio}$$

$$w_t = (1-\alpha)\, A_t\, (K_t/L_t)^\alpha \qquad r_t = \alpha\, A_t\, (K_t/L_t)^{\alpha-1} - \delta \qquad \text{Wage/Interest}$$

4.2.1 Numerical scheme – differentiation and state-constraint

There is a slight difference between the Finite-difference scheme we use in this model and the one we consider in the standard upwind scheme in the previous section. Without repeating the description developed above and in section B of appendix, we emphasize here a particular treatment of the Hamilton-Jacobi Bellman equation on two points: (i) the Hamiltonian is separable and the differentiation can be asymmetric in the two controls and (ii) the state-constraint should be handled differently in presence of two controls.

Asymmetric differentiation

We take advantage of the separability of the utility function: $u(c,\ell) = \bar{u}(c) - \tilde{u}(\ell)$ and hence of the Hamiltonian.

$$\mathcal{H}(p) = \max_{c,\ell} \bar{u}(c) - \tilde{u}(\ell) + (z_j w\, \ell - c)p = \mathcal{H}^c(p) + \mathcal{H}^\ell(p)$$

This will allow us to use an asymmetric upwind scheme for the two controls for the Hamilton-Jacobi-Bellman equation. As explain in the appendix in the case with one control, the upwind scheme intend to choose the direction of the differentiation (resp. Forward or Backward) to match the direction of the drift (resp. positive or negative). In our case, the drift for each control will have a different direction:

$$-c^\star = \frac{\partial \mathcal{H}^c(p)}{\partial p} \qquad\qquad z_j w\, \ell^\star = \frac{\partial \mathcal{H}^\ell(p)}{\partial p}$$

The consumption drift on wealth will be negative while the labor income drift will be positive. The choice of the first-order term $p \equiv \partial_a v$ will thus be backward for the first $\mathcal{H}^c$ (i.e. with $v_{i,j,B}$) and forward (i.e. $v_{i,j,F}$) for the second one $\mathcal{H}^\ell$. The underlying reason is to insure the monotonicity of the scheme, i.e. by keeping the discretized operator with positive diagonal term and negative upper/lower diagonal terms.

A good way of observing this is when linearizing the HJB equation around the solution v in the direction w:

$$-\partial_t v + \rho v - \mathcal{H}^c(\partial_a v) - \mathcal{H}^\ell(\partial_a v) - r_t a\, \partial_a v - \lambda_j(v_{-j} - v) = 0 \qquad \text{[HJB in v]}$$

$$-\partial_t w + \rho w \;\underbrace{-\frac{\partial \mathcal{H}^c(\partial_a v)}{\partial p}}_{\geq 0}\,\partial_a w \;\underbrace{-\frac{\partial \mathcal{H}^\ell(\partial_a v_j)}{\partial p}}_{\leq 0}\,\partial_a w \;- r_t a\, \partial_a w - \lambda_j(w_{-j} - w_j) = 0 \qquad \text{[linearized]}$$

Thus, it implies the use of $\partial_a w = \frac{w(a)-w(a-h)}{h}$ in the first case and $\partial_a w = \frac{w(a+h)-w(a)}{h}$ in the second case, to keep diagonal terms positive, for the operator to be a M-matrix.

Hence, we will always choose

$$\partial_a v \equiv v_{i,j,B} \qquad\qquad \text{in } c^\star = -\frac{\partial \mathcal{H}^c(\partial_a v)}{\partial p} = (\partial_c \bar{u})^{-1}(\partial_a v)$$

and

$$\partial_a v \equiv v_{i,j,F} \qquad\qquad \text{in } \ell^\star = \frac{1}{z_j w}\frac{\partial \mathcal{H}^\ell(\partial_a v)}{\partial p} = (\partial_\ell \tilde{u})^{-1}(z_j w \partial_a v)$$

State constraint

The state-constraint will impose an indirect constraint on the two controls, through the first-order of the value function. Let us restate the control problem of the Hamiltonian:

$$\mathcal{H}(p) = \max_{c,\ell} \bar{u}(c) - \tilde{u}(\ell) + (z_j w\, \ell - c)p$$

$$\text{At } a = \underline{a} \qquad da_t = (z_t w_t \ell_t + r_t a_t - c_t)dt \;\geq 0 \qquad [\lambda]$$

By Karush-Kuhn-Tucker, since it is a concave maximization problem on well-behaved functions, we pose the Lagrangian function and derive the optimality conditions:

$$
\begin{aligned}
&\mathcal{L}(c,\ell,\lambda) = \bar{u}(c) - \tilde{u}(\ell) + (z_j w\,\ell - c)p + \lambda\ (z_t w_t \ell + r_t a - c) \\
&\qquad \frac{\partial \mathcal{L}}{\partial c} = 0 \qquad \frac{\partial \mathcal{L}}{\partial \ell} = 0 \qquad \frac{\partial \mathcal{L}}{\partial \lambda} = 0 \qquad \lambda \geq 0 \qquad \lambda \frac{\partial \mathcal{L}}{\partial \lambda} = 0 && \text{[KKT conditions]} \\
&\Rightarrow \quad \partial_c \bar{u}(c) = (p+\lambda) \qquad \partial_\ell \tilde{u}(\ell) = z_j w (p+\lambda) && \text{[FOC]} \\
&\Rightarrow \quad \lambda \geq 0 \qquad (z_t w_t \ell + r_t a - c) \geq 0 && \text{[Feasibility cond.]} \\
&\Rightarrow \quad \lambda\ (z_t w_t \ell + r_t a - c) = 0 && \text{[Complementary Slackness]}
\end{aligned}
$$

Therefore, using the two FOC together, we recover the well known equality $MRS = MRT$[27], and rewriting the Complementary slackness, we obtain

$$
\begin{aligned}
\partial_\ell \tilde{u}(\ell) &= z_j w \partial_c \bar{u}(c) \\
(\partial_c \bar{u}(c) - p) \cdot\ (z_t w_t \ell + r_t a - c) &= 0
\end{aligned}
$$

Since, we did not specify the general method will be :

1. Implement the optimal allocation, using the Forward difference for $\partial_a V$
2. Check the primal feasibility conditions: $s > 0$ at $\underline{a}$
3. If the primal feasibility fails (i.e. $s^a < 0$ at $\underline{a}$), enforce the constraint: the Lagrange multipliers are no longer null and thus the FOCs can no longer be used (since we don't know the value of λ).
4. Since the Lagrange multiplier are strictly positive, the complementary slackness tells us to find the values of c, ℓ so that $s = 0$.

4.2.2 Results 1 – Discretized Brownian common noise

We now plot the evolution of the aggregate variables, in presence of a Brownian change in productivity. As observed previously in the Krusell Smith model, the main effect of a positive shock in productivity (e.g. the light blue line showing the best-case scenario) is to reduce precautionary saving.

When productivity increases in the first wave, the probability that A_t decreases below the initial level is indeed much lower, changing future value function, through its effect of conditional expectations. The optimal controlled drift – of the movement of the players – is thus reduced. As a consequence, both consumption increases and labor supply decreases.

[27] Marginal rate of substitution of the consumer ($\frac{\partial_\ell \tilde{u}(\ell)}{\partial_c \bar{u}(c)}$) equals the Marginal rate of Transformation of the firm (the real effective wage $z_j w_t$ here in the neoclassical model)

This effect is slightly contractionary: since output is a function of both labor and capital – that decline with the drop in savings – the output is reduced due to the change in expectations of households. This effect is only temporary since the aggregate productivity increases both wage and interest, fostering a change in the trend of labor and saving.

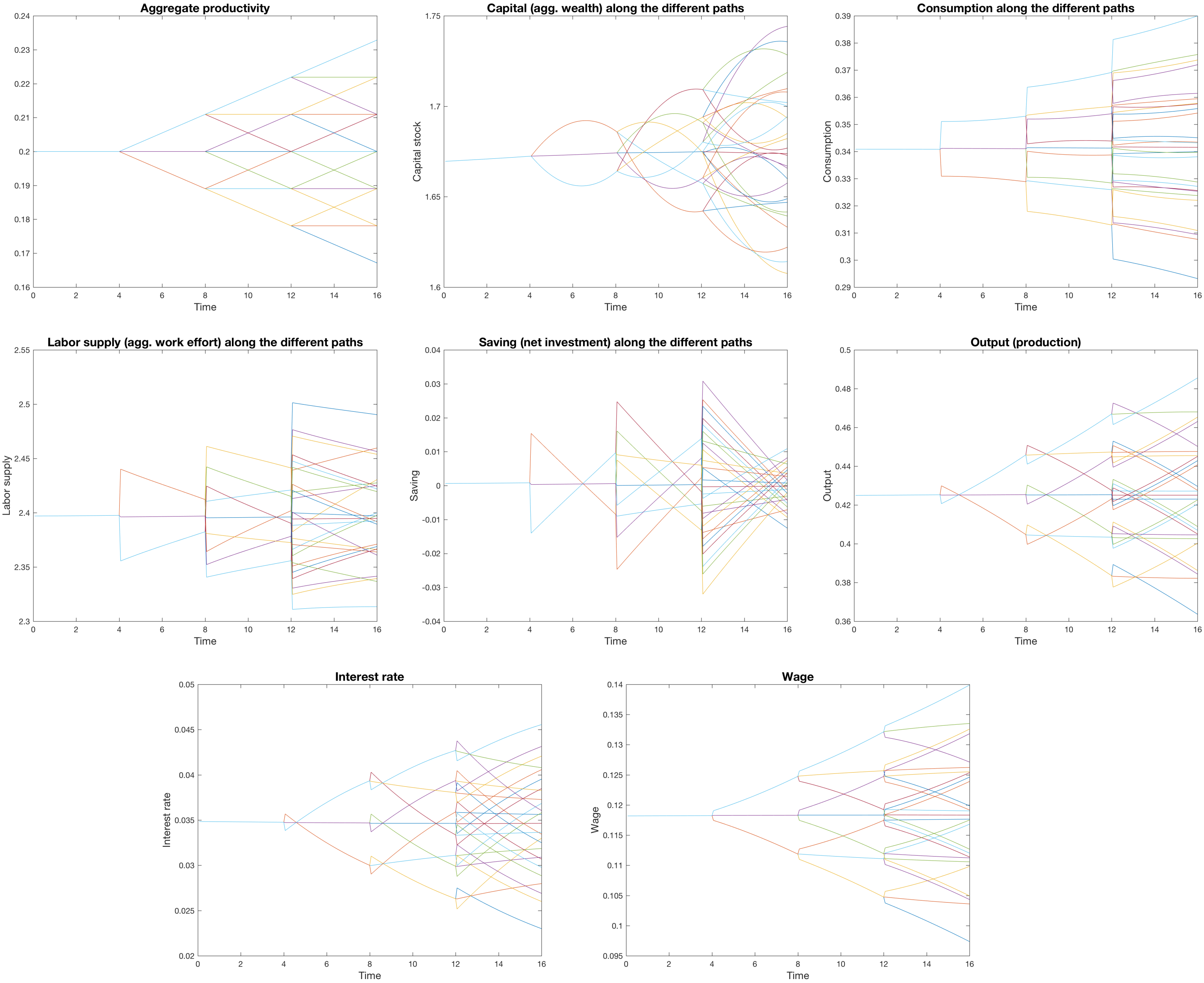


Figure 9: Evolution of aggregate variables with a Brownian TFP in the H.A. model with endogenous labor supply

4.2.3 Results 2 – Jump-drift process and IRF

In this section, we again display the transmission of a one-time shock of productivity on the other variables. A contrario to the previous section, the process for common noise follows a Jump-Drift process, and the surge in productivity is concentrated when the shock hits at $T = 4$.

The dynamics of some of the variables are similar to the responses displayed in the previous section. In particular, the investment – caused both by the interest rate rise and by precautionary motives – surges with aggregate productivity. Moreover, in presence of endogenous labor supply, work effort reacts to the change of wage. A new feature of this H.A. model is to show that poor agents – that have an effective wage $z_j w$ twice lower than the rich agents – reacts also as strongly as the rich agents. This is again due to the presence of Hand-to-mouth agents that may be close to the borrowing constraint. Indeed, instead of lowering consumption, these agents reacts by increasing their work effort when wage increases.

When capital accumulates, this decreases the marginal productivity of both capital and labor, implying a drop in interest rate. This tempers precautionary motives and revert the trend of consumption, labor and saving. Such reaction is again stronger that what one can observe in the representative agent, as we will see in the next section.

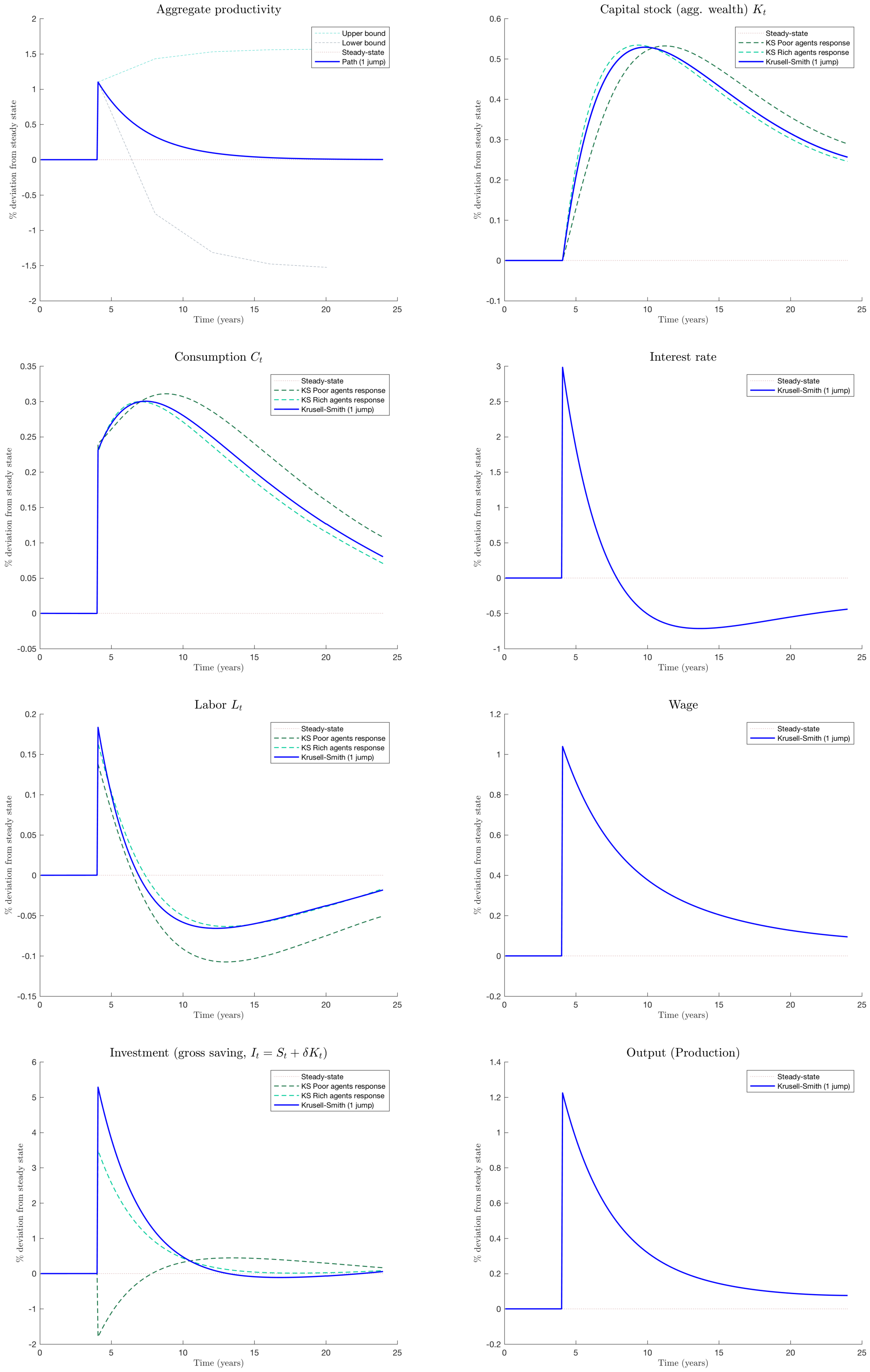


Figure 10: Impulse Responses to a one-time TFP shock in the KS model with endogenous labor supply

4.2.4 Comparison with representative agent model

In this section, we compare the previous model to the standard RBC model. We simulate again the model using DYNARE.

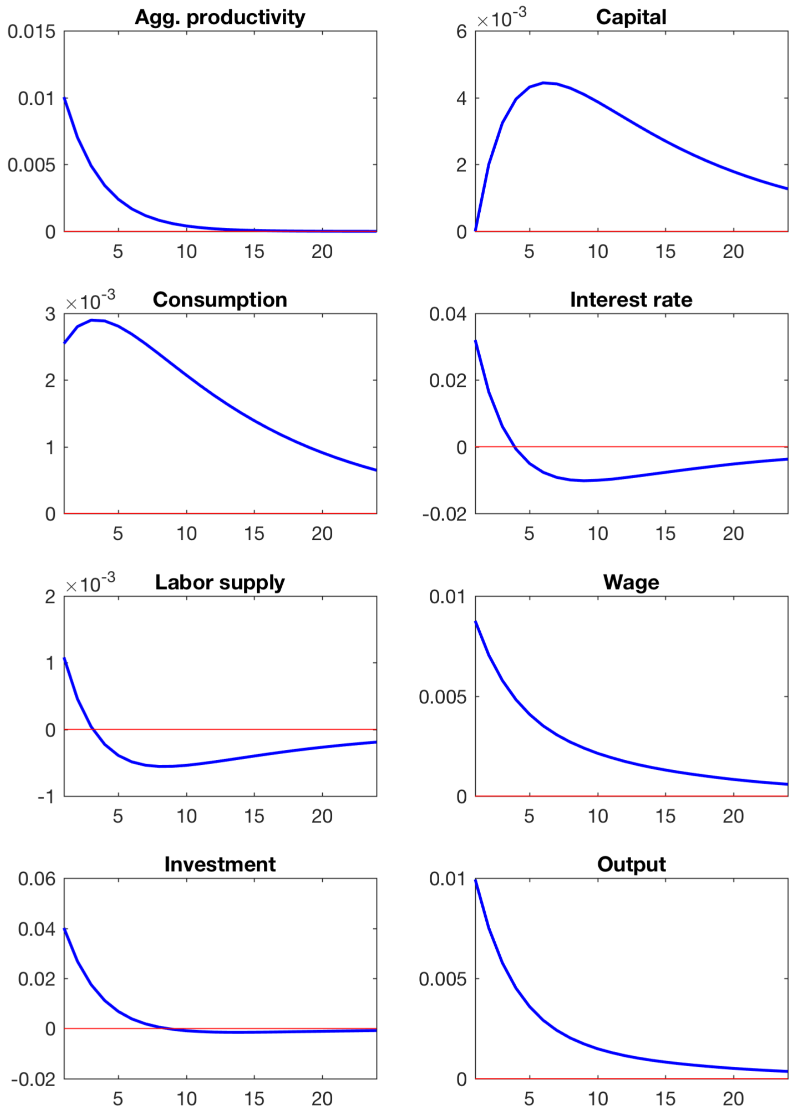


Figure 11: Impulse to a TFP shock in the RBC model

Even thought the qualitative reaction seems similar between R.A. and H.A., the quantitative effects display several differences. The reactions of saving is stronger in the H.A. model (5% instead of 4% in the RBC), as well as the labor supply, which is twice as high, with 0.2% instead of 0.1%. This is explained by respectively precautionary saving and hand-to-mouth consumers. The overall effect will imply a stronger amplification effect on output, that rises at 1.25% and last longer, compared to the RBC model.

An interesting point to note is that now the consumption reacts almost the same in both models at 0.2% of its steady state level. That can again be explained by precautionary motives, since the extra work effort will be converted into saving and capital accumulation instead of consumption in the heterogeneous agents model. Moreover,

4.3 One-asset HANK model

We now turn to the New Keynesian version of the H.A. model with labor supply, as described in section 2.2.2.

Recall the household control problem:

$$\max_{\{c_t\}_{t_0}^{\infty},\{\ell_t\}_{t_0}^{\infty}} \mathbb{E}_0 \int_0^{\infty} e^{-\rho t} u(c_t,\ell_t)dt$$

$$\text{subject to :} \qquad db_t = \big((1-\tau)z_t w_t \ell_t + r_t^b b_t + T_t + \Pi_t - c_t\big)dt \qquad \text{(Budget constraint)}$$

$$\text{and} \qquad b_t \geq \underline{b} \qquad \text{(Credit constraint)}$$

Recall the supply side and policy rules:

$$Y_t = A_t N_t \qquad w_t = A_t P_t m_t \qquad \text{(Production fct/Marg. cost)}$$

$$\Big(i_t - \pi_t - \frac{\dot{Y}_t}{Y_t}\Big)\pi_t = \frac{\varepsilon}{\theta}\big(m_t - \bar{m}\big) + \dot{\pi}_t \qquad \text{(NK Philipps Curve)}$$

$$\Pi_t := (1-m_t)P_t Y_t - \Theta_t(\pi_t) \qquad \Theta_t(\pi_t) = \frac{\theta}{2}\pi_t^2 Y_t \qquad \text{(Firms' profit, Adj. cost)}$$

$$\dot{B}_t^g + G_t + T_t = \tau_t w_t L_t + r_t B_t^g \qquad \text{(Gvt Budget)}$$

$$r_t := i_t - \pi_t \qquad i_t = \rho + \phi_\pi \pi_t \qquad \text{(Fisher relation/Taylor rule)}$$

$$B_h(r^b,t) := \sum_j \int_{\underline{b}}^{\infty} b\, g(t,b,z_j)db = -B^g \qquad \text{(Mkt clearing - Bond)}$$

$$L(w,t) := \sum_j \int_{\underline{b}}^{\infty} \ell^\star(b,z_j,t)\, g(t,b,z_j)db = N(w,t) \qquad \text{(Mkt clearing - Labor)}$$

4.3.1 Numerical scheme – some specificities of the HANK model

In this model, we use the same numerical scheme to handle the mean field games system composed of the HJB and the F.P. equations. However, we now have an additional structuring equation that drives the firm dynamics: the New Keynesian Philipps Curve now relates the evolution of inflation and output growth.

$$\Big(i_t - \pi_t - \frac{\dot{Y}_t}{Y_t}\Big)\pi_t = \frac{\varepsilon}{\theta}\big(m_t - \bar{m}\big) + \dot{\pi}_t$$

When one add common noise, this Backward differential equation becomes stochastic. If one could use the theory of BSDEs, the discretization procedure described in section 3.2 can be easily adapted to the treatment of this equation. The idea is again (i) to treat the equation on each branch of the tree in the deterministic way and (ii) to replaces the terminal condition of each branch by a conditional expectation of the endogenous variables, here π_t, m_t and Y_t.

4.3.2 Results 1 – Discretized Brownian common noise

Since the Mean Field Game system of the HANK is identical to the H.A. with labor supply displayed in the previous section, the result of this model are partly analogous.

The increase in productivity will reduce the precautionary motives and labor associated with it. This effect can also be found in standard New Keynesian model: positive productivity shocks usually imply negative output gap. Since there is no clear definition of "Natural level of output" in this H.A. model, the closest variable "symbolizing" output gap would output growth, which is indeed negative when the shock hits.

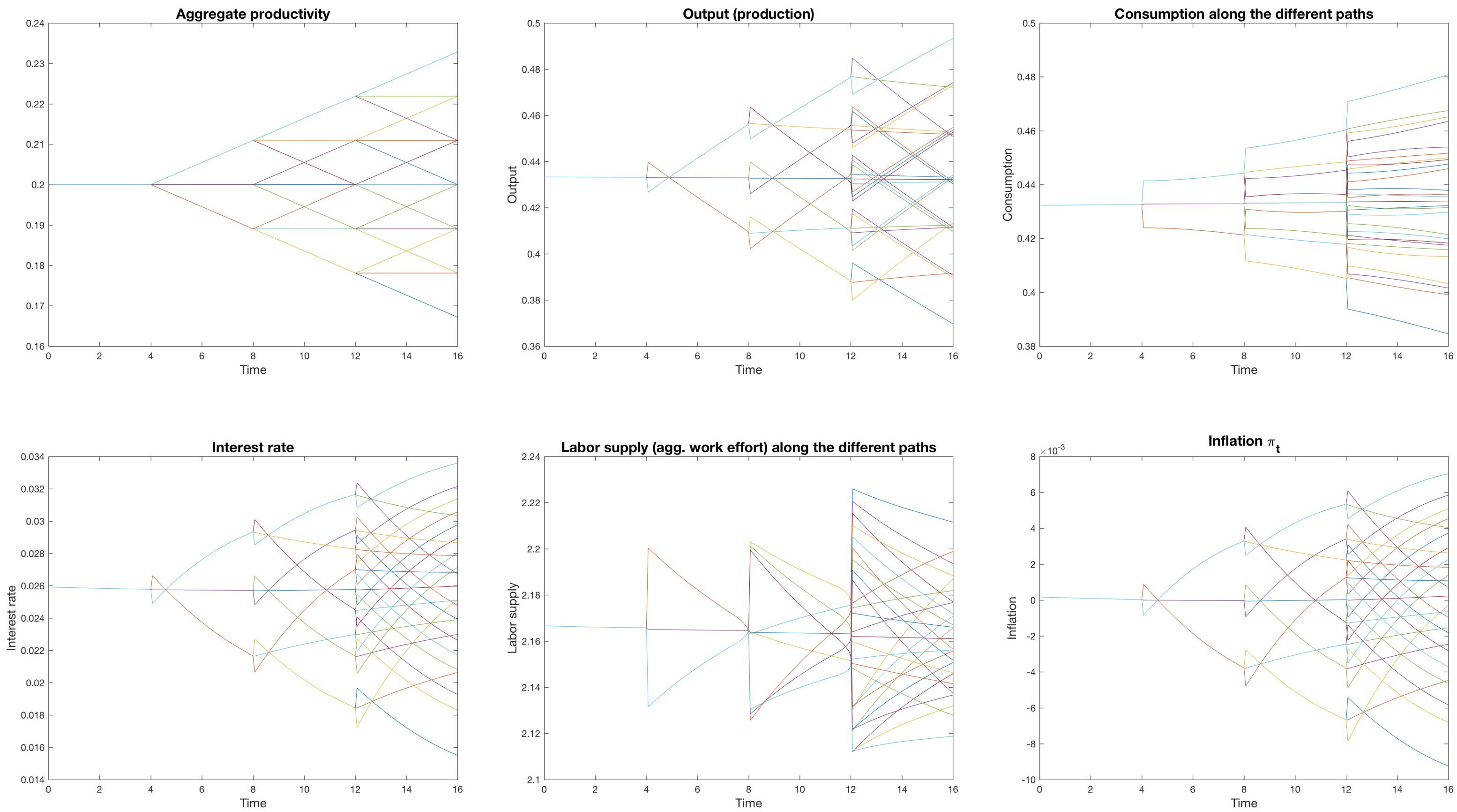


Figure 12: Impulse Responses to a one-time TFP shock in the HANK model (one asset)

The novelty of the NK model will however be (i) the rigid price setting of firms and (ii) the role for monetary policy. This small contraction will imply deflation due to the forward looking behavior of firms, as displayed in the N.K. Philipps Curve. With contraction and deflation, the central bank react by reducing its nominal interest rate (as displayed by the bottom RHS plot). However, as in the NK model, this monetary accommodation is not large enough to reduce the output gap and deflation.

The new feature of this Heterogeneous Agents model is now the change in the consumption of agents. The anticipation effect of this change in productivity – that directly affects the conditional expectation of future shocks – will incentivize the households to increase their consumption. This increase in aggregate demand is thus correlated to the future increase in output and inflation, affecting the real interest rate.

4.3.3 Results 2 – Jump-drift process and IRF

In this section, we compute the response of the economy after a shock of productivity – that follows the Jump-Drift process. The model has some similarities with the RBC model above (i.e. H.A. model with endogenous labor supply), since the MFG system is the same.

With the shock being concentrated at time $T = 4$, the productivity causes the output to surges. Moreover, the productivity will decrease the cost of labor for firms – marginal cost will drop – but the nominal wages will surge but proportionally by a larger amount. In response to this increase in wages and for precautionary motives, workers will increase their labor supply. Indeed, a large part of labor is supply by poor agents – in income and wealth – who work relatively more to get away from the credit constraint. The increase of both productivity and labor supply will imply a strong amplification effect on output, of 1.4% (for a productivity shock of less than 1.1%.

This expansionary effect will however be inflationary, as a consequence of firm's price setting – as explained by the Philipps Curve). The central bank will thus raises the nominal interest rate. This rise will inverse the trend of consumption and inflation, causing a drop in real wage and causing the strong reduction is labor supply by workers. This drop, added to the structural decrease in productivity will cause the economy to contract and experience a deflationary episode. Monetary authorities will therefore cut their interest rate until the return of inflation and output to the pre-shock level.

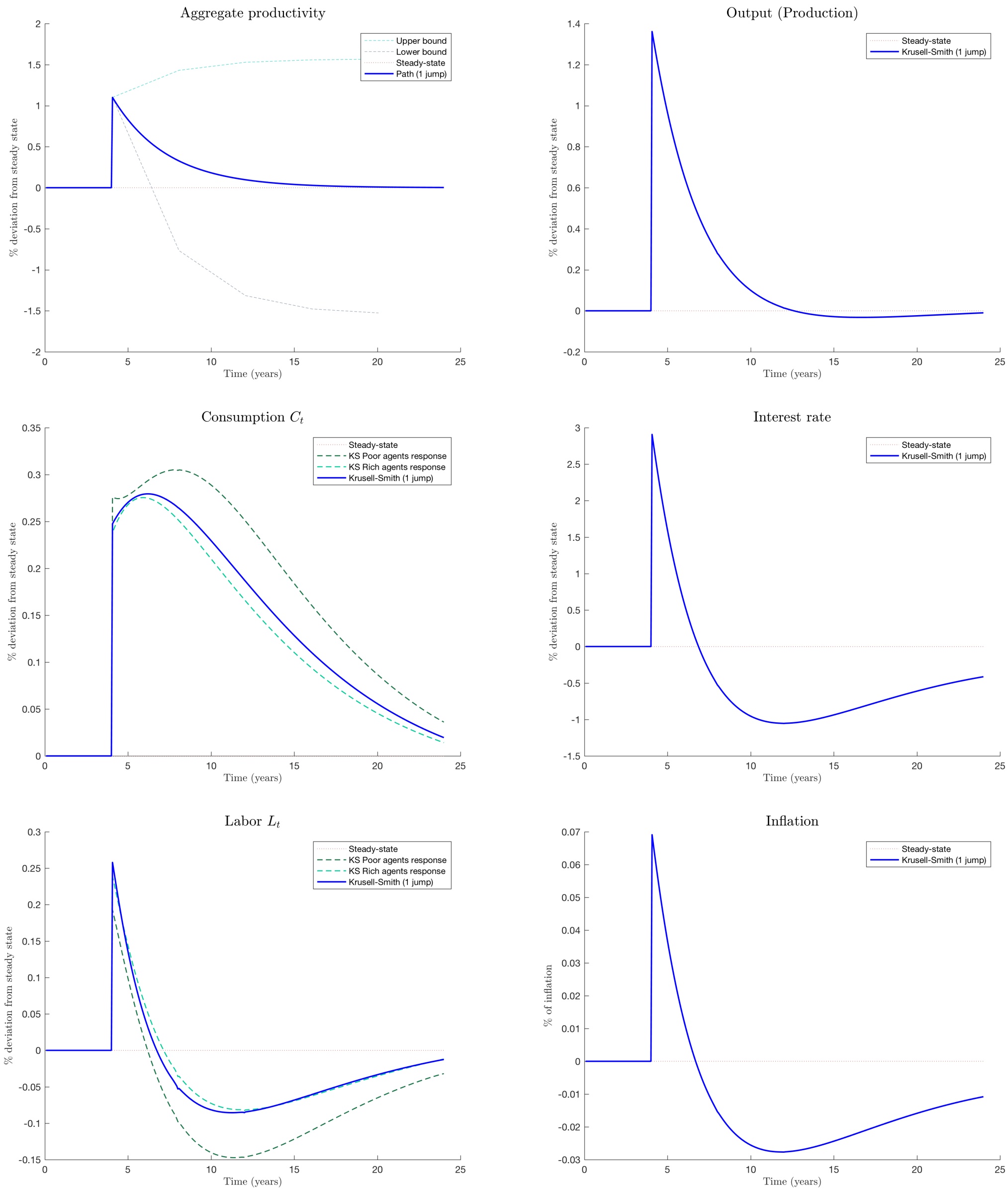


Figure 13: Impulse Responses to a one-time TFP shock in the HANK model (one asset)

4.3.4 Comparison with representative agent model

In this section, we compare the previous model to the standard New Keynesian model. We simulate again the model using DYNARE. However, our version of this framework – described in appendix section D – is the baseline version as found in Galí (2015) (or Woodford (2003)), and the relations are linearized around the zero-inflation steady state. Moreover, in the HANK model derived above, there are three deviations compared to the standard NK model:

- Price rigidity takes the form of Rotemberg pricing (instead of Calvo-Yun pricing)
- Distortive tax on labor $\tau w_t L_t$ are redistributed lump-sum T to households.
- Firms profits Π_t are redistributed to workers in function of their productivity z_j

This heterogenous agent version of the New Keynesian model will have very different predictions compared to the baseline model we plot in the next graph.

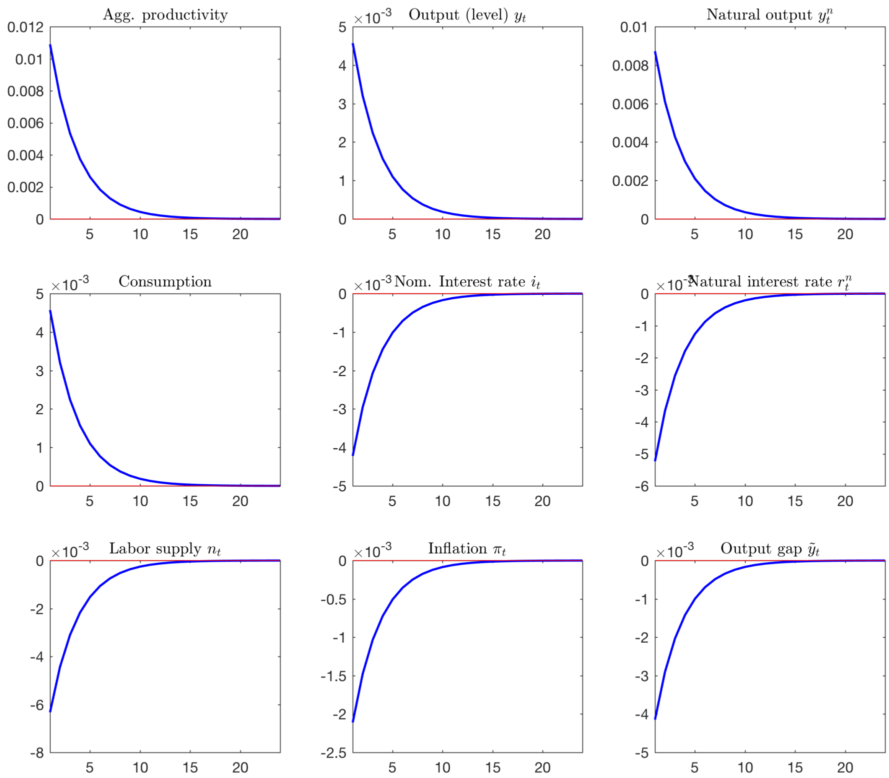


Figure 14: Impulse to a TFP shock in the standard NK model (à la Galí)

In such setting, the TFP shock increases natural level of output that would prevail under flexible prices – raising natural interest rate and decreasing firms marginal costs. However, prices are sticky and not all firms can adjust their prices because of Calvo staggered price setting. This relative price distortion reduces output of those firms, household consumption, due to monopolistic competition and "taste for diversity", and therefore employment: output gap widens. This relative contraction is partly accommodated by monetary policy, easing nominal interest rate. Nevertheless, this expansionary policy is not sufficient to close the negative output gap, yielding deflation.

The main difference with the HANK model developed above lies in the reverse prediction for labor supply, output 'gap' and inflation. In our H.A. model there is a strong role for precautionary motives and reaction of hand-to-mouth households close to the credit constraint – who are the one supplying the major part of work effort – and that can now respond positively in consumption. This positive reaction of consumption and labor imply an amplification effect absent in the standard framework, causing inflation (instead of deflation) and a monetary tightening (instead of easing). However, when the shock mean-reverts, the standard mechanisms at stake in the usual setting reappears – with drop in labor supply, output and inflation.

4.4 Two-assets model

In this section, we now turn to the two assets model. We recall the control problem:

$$\max_{\{c_t\}_{t_0}^{\infty},\{\ell_t\}_{t_0}^{\infty},\{d_t\}_{t_0}^{\infty}} \mathbb{E}_0 \int_0^{\infty} e^{-(\rho+\zeta)t} u(c_t,\ell_t)dt$$

$$\begin{aligned}
\text{subject to :}\quad & db_t = \big((1-\tau)z_t w_t \ell_t + (r_t^b(b)+\zeta)b_t + T - d_t - \chi(d_t,a_t) - c_t\big)dt && \text{(Budget – Liquid)}\\
& da_t = ((r_t^a+\zeta)a_t) + d_t)dt && \text{(Budget – Illiquid)}\\
& b_t \geq \underline{b} && \text{(State constraint – Liquid asset)}\\
\text{and}\quad & a_t \geq \underline{a} = 0 && \text{(State constraint – Illiquid asset)}\\
& r_t^b(b) = r_t^b + \mathbb{1}_{\{b<0\}}\kappa && \text{(Discrimination – lending/borrowing)}\\
& \chi(d,a) = \chi_0|d| + \frac{\chi_1}{1+\chi_2}\left|\frac{d}{a}\right|^{1+\chi_2}|a| && \text{(Illiquid asset transaction cost)}
\end{aligned}$$

The mean-field interaction, as before, lies in the Walrasian pricing that leads to market clearing in the two assets and in labor:

$$S_a(t,r^a) := \sum_j \int_{\underline{a}}^{\infty}\int_{\underline{b}}^{\infty} a\, g(t,a,b,z_j)da\, db = K(t,r^a) \qquad \text{[Illiquid asset/Capital]}$$

$$S_b(t,r^b) := \sum_j \int_{\underline{a}}^{\infty}\int_{\underline{b}}^{\infty} b\, g(t,a,b,z_j)da\, db = B^g \qquad \text{[Liquid asset/Bond]}$$

$$L(t,w) = \sum_j \int_{\underline{a}}^{\infty}\int_{\underline{b}}^{\infty} z_j\, \ell^\star(t,a,b,z_j,w)\; g(t,a,b,z_j)da\, db = N(t,w) \quad \text{[Labor supply/demand]}$$

The model, translated into a MFG formulation, yields (1) an Hamilton-Jacobi-Bellman equation :

$$\begin{aligned}
-\partial_t v(t,a,b,z_j) + (\rho+\zeta)v(t,a,b,z_j) &= \max_{c,\ell,d} u(c,\ell) + \partial_b v(t,a,b,z_j)\big[(1-\tau)z_j w\ell + (r^b(b)+\zeta)b + T - d - \chi(d,a) - c\big]\\
&+ \partial_a v(t,a,b,z_j)\big[(r^a+\zeta)a) + d\big] + \lambda_j\left(v(a,b,z_{-j},t) - v(a,b,z_j,t)\right)\\
&\text{subject to:}\qquad b \geq \underline{b} \quad \text{and} \quad a \geq \underline{a} = 0
\end{aligned}$$

The dual equation showing the evolution of the distribution is the (2) Fokker Planck equation:

$$\begin{aligned}
0 = \partial_t g(t,a,b,z_j) &+ \partial_a\left[s_t^a(a,b,z_j)\, g(t,a,b,z_j)\right] + \partial_b\left[s_t^b(a,b,z_j)\, g(t,a,b,z_j)\right]\\
&\lambda_j g(t,a,b,z_j) - \lambda_{-j} g(t,a,b,z_{-j}) - \zeta g(t,a,b,z_{-j}) + \zeta\, \delta_0(a)\delta_0(b)g^*(z_j)
\end{aligned}$$

$$\begin{aligned}
\text{with}\quad & s_t^a(a,b,z_j) = (r_t^a+\zeta)a + d_t^\star\\
\text{and}\quad & s_t^b(a,b,z_j) = (1-\tau)z_j w_t \ell_t^\star + (r_t^b(b)+\zeta)b + T - d_t^\star - \chi(d_t,a_t) - c_t^\star
\end{aligned}$$

where ζ is the rate of death (and birth) of new agents, starting with zero wealth (hence the Dirac mass $\delta_0(a)\delta_0(b)$), and in the state z_j with probability $g^*(z_j)$, with g^* the stationary distribution of z_j. Note that $c^\star, \ell^\star, d^\star$ denote the optimal controls of the agents.

4.4.1 Control problem, HJB and state-constraints

One can restate the HJB, using Hamiltonians:

$$
\begin{aligned}
-\partial_t v + (\rho+\zeta)v &= \mathcal{H}^c(\partial_b v) + \mathcal{H}^\ell(\partial_b v) + \mathcal{H}^d(\partial_b v, \partial_a v) + (\tilde{r}^b b + T)\partial_b v + \tilde{r}^a a \partial_a v \\
&\mathcal{H}^c(p) = \max_c \bar{u}(c) - pc \\
&\mathcal{H}^\ell(p) = \max_\ell -\tilde{u}(\ell) + (1-\tau)wz_j\ell p \\
&\mathcal{H}^d(p,q) = \max_d d\, q \; - (d + \chi(d,a))p \\
&\text{with} \quad u(c,\ell) = \bar{u}(c) - \tilde{u}(\ell) \qquad \text{and} \quad \tilde{r}^b = r_t^b(b) + \zeta \qquad \tilde{r}^a = r_t^a + \zeta
\end{aligned}
$$

The optimality conditions of the control problem are the following:

$$
\begin{aligned}
\partial_c \bar{u}(c) &= p & &\Rightarrow & c^\star &= (\partial_c \bar{u})^{-1}\big(p\big) \\
\partial_\ell \tilde{u}(\ell) &= (1-\tau)wz_j p & &\Rightarrow & \ell^\star &= (\partial_\ell \tilde{u})^{-1}\big((1-\tau)wz_j\, p\big) \\
(1+\partial_d\chi(d,a))p &= q & &\Rightarrow & d^\star &= (\partial_d\chi)^{-1}\left(\frac{q}{p} - 1\right)
\end{aligned}
$$

Using the analytical formulas for $\partial_c u(c,\ell) = c^{-\gamma}$, $\partial_\ell u(c,\ell) = \ell^{-1/\phi}$ and $\partial_d \chi(d,a) = \mathbb{1}_{\{d>0\}}[\chi_0 + (\frac{d}{\chi_1 a})^{\chi_2}] + \mathbb{1}_{\{d<0\}}[-\chi_0 + (\frac{-d}{\chi_1 a})^{\chi_2}]$ we obtain, for $p \equiv \partial_b v$ and $q \equiv \partial_a v$:

$$
\begin{aligned}
&c^\star = \big(p\big)^{-1/\gamma} \qquad\qquad \ell^\star = \big(\tilde{z}_j\, p\big)^{\phi} \\
&d^\star = \mathcal{D}(q,p) := \left(\chi_1 a \left[\frac{q}{p} - 1 - \chi_0\right]^+\right)^{1/\chi_2} - \left(\chi_1 a \left[-\Big(\frac{q}{p} - 1 - \chi_0\Big)\right]^-\right)^{\frac{1}{\chi_2}}
\end{aligned}
$$

with the conventional notation: $f^+ = \max(f,0)$ and $f^- = \max(-f,0) = -\min(f,0)$ so that $f = f^+ - f^-$

4.4.2 Numerical scheme

As described in appendix[28], the differentiation method should make sure that the scheme is monotonous. For that, the Upwind scheme is implemented as in previous section, where we

[28] Complete references can also be found in the numerical appendices *Note on algorithm (Kaplan-Moll-Violante)* and *Note on Two Assets and Kinked Adjustment cost (Moll)* on B. Moll website

use the separability of the Hamiltonian. This one will be split between $\mathcal{H}^c(p) + \mathcal{H}^\ell(p)$ on one side and $\mathcal{H}^d(p,q)$ on the other.

For $\mathcal{H}^c(p) + \mathcal{H}^\ell(p)$:

$$
\begin{array}{lll}
c^F_{i,j,k} = (\partial_c u)^{-1}(D_b V^F_{i,j}) \quad \& \quad \ell^F_{i,j,k} = (\partial_\ell u)^{-1}(\tilde{z}_j\, D_b V^F_{i,j,k}) & \Rightarrow & s^F_{i,j,k} = \tilde{z}_j \ell^F_{i,j,k} + \tilde{r}^n\, b_i - c^F_{i,j,k} \\
c^B_{i,j,k} = (\partial_c u)^{-1}(D_b V^B_{i,j}) \quad \& \quad \ell^B_{i,j,k} = (\partial_\ell u)^{-1}(\tilde{z}_j\, D_b V^B_{i,j,k}) & \Rightarrow & s^B_{i,j,k} = \tilde{z}_j \ell^B_{i,j,k} + \tilde{r}^n\, b_i - c^B_{i,j,k}
\end{array}
$$

For $\mathcal{H}^d(p,q)$:

Optimal drift for b : $\frac{d\mathcal{H}^d(p,q)}{dp} = -(d^\star - \chi(d^\star, a))$, when drift positive $\Rightarrow$ use Forward difference

drift negative: $\Rightarrow$ use Backward difference

Optimal drift for a : $\frac{d\mathcal{H}^d(p,q)}{dq} = d$, when the drift positive, $\Rightarrow$ use Forward difference

drift negative: $\Rightarrow$ use Backward difference

$$
d^{F,B}_{i,j,k} = \mathcal{D}(D_a V^F_{i,j,k}, D_b V^B_{i,j,k}) \qquad \& \quad d^{B,F}_{i,j,k} = \mathcal{D}(D_a V^B_{i,j,k}, D_b V^F_{i,j,k}) \qquad \& \quad d^{B,B}_{i,j,k} = \mathcal{D}(D_a V^B_{i,j,k}, D_b V^B_{i,j,k})
$$

$$
\begin{aligned}
d_{i,j,k} = \mathbb{1}\{d^{F,B} > 0 \ \& \ d^{F,B} > -\chi(d^{F,B})\}\ d^{F,B}_{i,j,k} \\
+ \mathbb{1}\big[d^{B,F} < 0 \ \& \ d^{B,F} < -\chi(d^{B,F})\big]\ d^{B,F}_{i,j,k} \\
+ \mathbb{1}\big[d^{B,B} < 0 \ \& \ d^{B,B} > -\chi(d^{B,B})\big]\ d^{B,B}_{i,j,k}
\end{aligned}
$$

and since $\chi(d) > 0\ \forall d$ we never have $d^{F,F}$ since $d > 0$ and $d < -\chi(d)$ never happen simultaneously.

4.4.3 Enforcement of the state-constraint

Let us restate the control problem at the credit constraints (there are two of them $\underline{a}$ and $\underline{b}$) when optimizing controls for the Hamilton-Jacobi-Bellman equation (verification argument):

$$
\begin{array}{lll}
H(p,q) = \max\limits_{c,\ell,d} & \bar{u}(c) - \tilde{u}(\ell) + p\big[\tilde{z}\ell - c - d - \chi(d,a)\big] + q[d] & \\
\text{At } b = \underline{b} & \dot{b} = \tilde{z}\ell - c - d - \chi(d,a) + \tilde{r}^b b + T \geq 0 & [\lambda] \\
\text{At } a = \underline{a} & \dot{a} = \tilde{r}^a a + d \geq 0 & [\mu]
\end{array}
$$

At $a = \underline{a}$, and $b = \underline{b}$, use the two constraints together.

By Karush Kuhn Tucker (appropriately posing the Lagrangian function), the optimality conditions are:

$$
\begin{array}{llcll}
[c] & \partial_c \bar{u}(c) - p - \lambda = 0 & \Rightarrow & \partial_c \bar{u}(c) = p + \lambda & (*) \\
[\ell] & -\partial_\ell \tilde{u}(\ell) + \tilde{z}p + \tilde{z}\lambda = 0 & \Rightarrow & -\partial_\ell \tilde{u}(\ell) = \tilde{z}(p+\lambda) = \tilde{z}\partial_c \bar{u}(c) & (**) \\
[d] & -(1 + \partial_d \chi(d))p + q - \lambda(1 + \partial_d \chi(d)) + \mu = 0 & \Rightarrow & (1 + \partial_d \chi(d))[p + \lambda] = q + \mu & (***)
\end{array}
$$

Adding the primal and dual feasibility conditions and complementary slackness conditions:

$$\begin{aligned}
&\tilde{z}\ell - c - d - \chi(d,a) + \tilde{r}^b b + T \geq 0 && \lambda \geq 0 \\
&\text{and} \qquad \big(\tilde{z}\ell - c - d - \chi(d,a) + \tilde{r}^b b + T\big) \cdot \lambda = 0 && \\
&\tilde{r}^a a + d \geq 0 && \mu \geq 0 \\
&\text{and} \qquad \big(\tilde{r}^a a + d\big) \cdot \mu = 0 &&
\end{aligned}$$

Using the FOC $(*)$ and $(***)$, the complementary slackness conditions become:

$$\begin{aligned}
\big(\tilde{z}\ell - c - d - \chi(d,a) + \tilde{r}^b b + T\big) \cdot \big(\partial_c \bar{u}(c) - p\big) = 0 \\
\big(\tilde{r}^a a + d\big) \cdot \big(1 + \partial_d \chi(d))[p + \lambda] - q\big)
\end{aligned}$$

Let us note that the analysis above is very generic. As we did not specify any value for a or b, the general method will be:

1. Implement the optimal allocation, using the Forward difference for $D_a V$ or $D_b V$
2. Check the primal feasibility conditions: $s^a > 0$ at $\underline{a}$, or $s^b > 0$ at $\underline{b}$ or both at $(\underline{a}, \underline{b})$
3. If the primal feasibility fails (e.g. $s^a < 0$ at $\underline{a}$), enforces the constraint: the Lagrange multipliers are no longer null and thus $(*)$ and $(***)$ are no longer possible (since we don't know the value of λ and μ.
4. Since the Lagrange multiplier are strictly positive, the complementary slackness tells us to find the values of c, ℓ, d so that $s^a = 0$ or $s^b = 0$

With this general idea in mind, let us describe each constraint in turn:

A) At $a = \underline{a}$ the illiquid asset constraint:

- Compute $\hat{d} = \mathcal{D}(D_a V^F, D_b V^{up})$, $\hat{c} = (\partial_c u)^{-1}(D_b V^{up})$ and $\hat{\ell} = (\partial_c u)^{-1}(D_b V^{up})$ with $D_b V^{up}$ given (as usual) by the right upwind scheme
- If $\hat{s}^a > 0$ keep these values
- Else if $\dot{a} = \hat{s}^a < 0$ enforces the constraints: $\hat{s}^a = 0$
- This implies

  $$\tilde{r}^a \underline{a} + d = 0$$

  i.e. $d^\star = -\tilde{r}^a \underline{a}$ $(= 0$ here).
- Feed the transition matrix appropriately with $d^\star$ (for both the diagonal and upper-diagonal).

B) At $b = \underline{b}$ the liquid asset constraint, the method is more heavy computationally:

- Compute $\hat{d} = \mathcal{D}(D_a V^{up}, D_b V^F)$, $\hat{c} = (\partial_c u)^{-1}(D_b V^F)$ and $\hat{\ell} = (\partial_c u)^{-1}(D_b V^F)$ with $D_a V^{up}$ given (as usual) by the right upwind scheme
- If $\hat{s}^b > 0$ keep these values
- Else if $\dot{b} = \hat{s}^b < 0$ enforces the constraints: $\hat{s}^b = 0$
- This implies that:

$$\widetilde{z}\ell - c - d - \chi(d,a) + \tilde{r}^b \underline{b} + T = 0$$

- Method to find the optimal c, ℓ, d s.t. this constraint is binding. The relation $(*)$, $(**)$ and $(***)$ (with $\mu = 0$) are still valid:

$$\begin{aligned}
q = (1 + \partial_d \chi(d))[p + \lambda] = (1 + \partial_d \chi(d))(\partial_c \bar{u}(c)) \quad &\Rightarrow \quad \check{d}(c) = \mathcal{D}(\partial_c u(c), D_a V^{up}) \\
-\partial_\ell \tilde{u}(\ell) = \widetilde{z} \partial_c \bar{u}(c) \quad &\Rightarrow \quad \check{\ell}(c) = -(\partial_\ell \tilde{u})^{-1}\big(\widetilde{z}\, \partial_c \bar{u}(c)\big)
\end{aligned}$$

- The only thing to find is thus to solve numerically (unique solution for positive values!) :

$$\widetilde{z}\check{\ell}(c) - c - \check{d}(c) - \chi(\check{d}(c), a) + \tilde{r}^b \underline{b} + T = 0$$

- Once this $c^\star$ found, compute $\ell^\star = \check{\ell}(c^\star)$ and $d^\star = \check{d}(c^\star)$
- Feed the transition matrix appropriately with $c^\star, \ell^\star, d^\star$ (for both diagonal & upper-diagonal).

C) At $(a, b) = (\underline{a}, \underline{b})$ the corner solution will be implemented:

- Compute $\hat{d} = \mathcal{D}(D_a V^F, D_b V^F)$, $\hat{c} = (\partial_c u)^{-1}(D_b V^F)$ and $\hat{\ell} = (\partial_c u)^{-1}(D_b V^F)$
- If $\hat{s}^a > 0$ and $\hat{s}^b > 0$, keep these values
- If $\hat{s}^a > 0$ and $\hat{s}^b < 0$, go back to step B)
- If $\hat{s}^b < 0$ and $\hat{s}^a < 0$, go back to step A)
- If both $\hat{s}^a < 0$ and $\hat{s}^b < 0$:
- 1) Impose $d^\star = -\tilde{r}^a \underline{a} =: \underline{d}$
- 2) Solve numerically: (with, again $\check{\ell} = -(\partial_\ell \tilde{u})^{-1}\big(\widetilde{z}\, \partial_c \bar{u}(c)\big)$)

$$\widetilde{z}\check{\ell}(c) - c - \underline{d} - \chi(\underline{d}, a) + \tilde{r}^b \underline{b} + T = 0$$

- Once this $c^\star$ found, compute $\ell^\star = \check{\ell}(c^\star)$
- Feed the transition matrix appropriately with $c^\star, \ell^\star, \underline{d}$ (for both diagonal & upper-diagonal).

However, the implementation of this method suffers from several numerical issues. For simulation of the model with different methods to handle state-constraint, refer to Kaplan et al. (2018).

5 Conclusion

In this project, I used novel methods to simulate the Mean Field Game system with common noise. This system is composed of two stochastic PDEs: a Hamilton-Jacobi-Bellman describing the evolution of the value function and a Fokker-Planck describing the evolution of the distribution, and both depends on the realization of the common noise. After a discretization of this process using different tree structures, these two PDEs were solved using a Finite-Difference Scheme. This method was applied to the standard Krusell-Smith model and two extension with endogenous labor supply, à la Real Business Cycle model, or sticky prices, à la New Keynesian model. We also provide some intuitions for the resolution of the stationary equilibrium of the two-assets model à Kaplan-Moll-Violante.

One of the main result is to show that precautionary saving behavior of households against aggregate fluctuations is important quantitatively and highly depends on the anticipations of the common noise.

Appendices

A MFG general setting

A Mean-Field Game can be described a game with a large number of "small" symmetric players, with a "mean-field" type of interaction: the interaction is only reflected between each agent and the "distribution" – i.e. the measure m_t – of all the other agents. This game is the limit of $N-$players differential games taking the limit as $N \to \infty$. The control problem of an agent[29] is the following:

$$\sup_{\alpha} \mathbb{E}\left[\int_0^{\infty} e^{-\rho t} f(t, X_t, m_t, \alpha_t)dt + g(X_T, m_T)\right]$$
$$dX_t = b(t, X_t, m_t, \alpha_t)dt + \sigma(t, X_t, m_t)dB_t + \sigma^0(t, X_t, m_t)dW_t^0$$

where W_t and B_t^0 refers respectively to the idiosyncratic noise and the common noise. Let us consider the case without common noise here $\sigma^0 = 0$

To control this SDE of Mean-Field type, it is useful to define the Hamiltonian:

$$\mathcal{H}(t, x, m, \nabla v, D_x^2 v) = \sup_a \left(f(t, x, m, a) + b(t, x, m, a) \cdot \nabla_x v(t, x) + \tfrac{1}{2}\operatorname{tr}(\sigma\sigma'(t, x, m, a) D_x^2 v(t, x))\right)$$

[29] Such typical agent is unfortunately called "representative agent" in the MFG literature

If v is regular enough ($\mathcal{C}^{1,2}$), the controls α are given by the feedback:

$$\alpha^\star \in \underset{a}{\operatorname{argmax}}\, \mathcal{H}(t, x, m, \nabla v, D_x^2 v)$$

In the simple case where the diffusion is uncontrolled, e.g. $\sigma(\cdot) = \sqrt{2\nu}$, the controlled drift is given by:

$$b(t, x, m, \alpha^\star) = -D_p \mathcal{H}(t, x, m, \nabla_x v)$$

For a given measure m_t (trajectory fixed!), the SDE is Markovian. We can use standard Dynamic Programming techniques. This stochastic control problem boils down to a Hamilton-Jacobi-Bellman (HJB) equation – where the agents make their choices $\alpha^\star$ taking the measure m_t of agents as given – and the transversality condition pins down the behavior at the limit:

$$\begin{aligned}
&\text{(i)} \qquad -\partial_t v + \rho\, v - \nu \Delta v - \mathcal{H}(t, x, m, \nabla_x v) = 0 \qquad \text{on } \mathbb{R}^d \times [0, T] \\
&\qquad\quad\ \lim_{t \to \infty} e^{-\rho t} \nabla_x v(t, x) \cdot x = 0 \quad \text{ on } \ \mathbb{R}^d
\end{aligned}$$

Solving the HJB equation yield the optimal control, the value function and the controlled drift of the SDE $X^{\alpha^\star}$. Taking the controls as given, this SDE allows to compute the evolution of the agents. The ideas draws from Mean-Field Theory where the movement of particles is described as a PDE on the probability measures. This PDE is called the Fokker Planck (or Kolmogorov Forward):

$$\begin{aligned}
&\text{(ii)} \qquad \partial_t m - \nu \Delta m - \operatorname{div}\big(D_p \mathcal{H}(t, x, m, \nabla_x v)\, m\big) = 0 \qquad \text{on } \ \mathbb{R}^d \times [0, T] \\
&\qquad\quad\ m(0, \cdot) = m_0(\cdot) \qquad\quad \text{on } \ \mathbb{R}^d
\end{aligned}$$

The MFG system is thus given by these two coupled partial differential equations. The coupling between the two PDE is induced by the measure m_t. The main idea to solve (and prove the existence) of MFG equilibria – without common noise – is to use a fixed-point procedure – relying on Schauder fixed-point theorem. The process is as follows:

1. Fix a measure m_t
2. Solve the optimal control problem and find $\alpha^\star$, and the controlled drift $b(t, x, m, \alpha^\star)$ and diffusion terms $\sigma(t, x, m, \alpha^\star)$.
3. Find the corresponding probability law of the controlled process $X_t^{\alpha^\star}$
4. Iterating to find a fixed point $m_t = \mathcal{L}(X_t)$.

B Numerical algorithm – a finite difference scheme

This section provides a complete description of the numerical algorithm for the transition equilibria of the Aiyagari-Bewley model, which is the basis of the Krusell-Smith model as explained in section 4.1. This appendix is inspired by the Numerical Appendix of the article by Achdou, Han, Lasry, Lions, and Moll (2017).

B.1 Solving the Hamilton Jacobi Bellman equation

We use a finite difference method, an in particular an "implicit upwind scheme". We here provide a description of the numerical algorithm. The functions v_1, v_2, g_1, g_2 are approximated at I discrete points in the space dimension, $a_i, i = 1, ..., I$. Grids are equispaced, and we denote by h the distance between grid points in space and τ distance between grid points in time. We use the short-hand notation $v_{i,j}^n \equiv v_j(a_i, t_n)$. The derivative ${v_{i,j}^n}' = \partial_a v_j(a_i, t_n)$ is approximated with either a forward or a backward difference approximation:

$$\partial_a v_j(a_i, t_n) \approx \frac{v_{i+1,j}^n - v_{i,j}^n}{h} \equiv {v_{i,j,F}^n}'$$
$$\partial_a v_j(a_i, t_n) \approx \frac{v_{i-1,j}^n - v_{i,j}^n}{h} \equiv {v_{i,j,B}^n}'$$

The aim of the upwind scheme is to use the proper scheme to approximate this first order term in the HJB equation. The $\partial_a v_j(a_i)$ is approximated with a forward difference ('décentrage à gauche') whenever the drift of the state variable $s_j^n(a_i)$ is positive and with a backward difference approximation ('décentrage à droite') whenever it is negative.

The implicit scheme will be a way to (i) iterate over the value function in the stationary case (from $n = 1$ until v^n is close enough from v_∞, to and (ii) solving for the value v^n at time t_n given the knowledge of v^{n+1} at t_{n+1}, since the HJB is running backward. The two corresponding implicit scheme will be the following :

(i) Stationary case: (13)

$$\frac{v_{i,j}^{n+1} - v_{i,j}^n}{\tau} + \rho v_{i,j}^{n+1} = u(c_{i,j}^n) + \underline{{v_{i,j}^{n+1}}'}(z_j w + r a_i - c_{i,j}^n) + \lambda_j(v_{i,-j}^{n+1} - v_{i,j}^{n+1})$$

(ii) Time-varying case (running backward): (14)

$$\frac{v_{i,j}^n - v_{i,j}^{n+1}}{\tau} + \rho v_{i,j}^n = u(c_{i,j}^{n+1}) + \underline{{v_{i,j}^n}'}(z_j w^{n+1} + r^{n+1} a_i - c_{i,j}^{n+1}) + \lambda_j(v_{i,-j}^n - v_{i,j}^n)$$

Several important remarks

Direction of time Since the second one is exactly the same scheme as the first, only by interchanging the direction of time and the iteration on n ($\tilde{t} \leftarrow T - t$ and $\tilde{n} \leftarrow N - n$), we will only focus on the first scheme for the description. All the results presented further will hold for the second case. Furthermore, it is surely more convenient to think of the numerical scheme as a search for v^{n+1} knowing v^n.

A semi-implicit scheme Concretely, this scheme is only *'semi implicit'* since the method do not approximate for the drift s^{n+1} but use the closed form solution for $s^n_{i,j} = z^j w^n + r^n\, a_i - c^n_{i,j}$ where $c^n_{i,j}$ is explicitely given by the first order condition:

$$c^n_{i,j} \approx c_j(t,a) = (u')^{-1}(\partial_a v_j(t,a)) \approx (u')^{-1}({v^n_{i,j}}') \tag{15}$$

The upwind scheme. It will aim at handling the first order underlined term above, to ensure the scheme is monotonous (cf. 'informal' theoretical results presented in the next two pages). This will imply to use a forward approximation ${v^n_{i,j,F}}'$ when the drift, i.e. the transport term, is positive $s^n_{i,j} > 0$, and conversely, use backward approximation ${v^n_{i,j,B}}'$ when the drift is negative $s^n_{i,j} < 0$. The drift will be mostly driven by consumption, and c.f. the previous remark, we will distinguish between two approximations for consumption:

$$\begin{aligned} c^n_{i,j,F} &= (u')^{-1}({v^n_{i,j,F}}') & \Rightarrow \qquad s^n_{i,j,F} &= z^j w^n + r^n\, a_i - c^n_{i,j,F} \\ c^n_{i,j,B} &= (u')^{-1}({v^n_{i,j,B}}') & \Rightarrow \qquad s^n_{i,j,B} &= z^j w^n + r^n\, a_i - c^n_{i,j,B} \end{aligned}$$

We therefore use the following approximation for the first order term:

$${v^n_{i,j}}'_{upwind} = {v^n_{i,j,B}}'\, \mathbb{1}_{\{s^n_{i,j,B}<0\}} + {v^n_{i,j,F}}'\, \mathbb{1}_{\{s^n_{i,j,F}>0\}} + {v^n_{i,j,0}}'\, \mathbb{1}_{\{s^n_{i,j,F}\leq 0 \leq s^n_{i,j,B}\}} \tag{16}$$

where some grids points fall in a set where the drift is approximately null, $s^n_{i,j,F} \leq 0 \leq s^n_{i,j,B}$. In this case, approximate the value function ${v^n_{i,j}}'$ by the utility when saving is null (i.e. consuming all income): ${v^n_{i,j,0}}' = u'(z_j w^n + r^n a_i)$.

Furthermore, the running gain $u(c^n_{i,j})$ is computed using the FOC eq. (15) and therefore $c^n_{i,j} = (u')^{-1}[{v^n_{i,j}}'_{upwind}]$ using the derivative of the value function obtained eq. (16)

The state constraint. Recall that theoretically – combining the FOC from HJB and the credit constraint – we obtained: $v^j(\underline{a}) \geq u'(z_j w + r\underline{a})$. Therefore, if the saving is negative at the boundary of the set, i.e. at a_1 in our discretized scheme, the upwind scheme imply to use a backward difference. A way to handle this state-contraint is to enforce it by setting:

$${v^n_{1,j,B}}' = u'(z_j w^n + r^n a_1)$$

This imply the state constraint is imposed whenever the backward difference approximation ${v^n_{1,j,B}}'$ would result in negative saving on the boundary. In the case the forward difference ${v^n_{1,j,F}}'$ is used – when $s_{1,j,F} > 0$ – the drift will move the agent away from the constraint and as a result the value function 'will never see the constraint'.

The initial guess. A simple initial guess for the value function is the value when savings is null and the agent never change state. The stationary HJB results in the guess:

$$v^0_{i,j} = \frac{u(z_j w + r a_i)}{\rho}$$

A matrix reformulation

With the notation $x^+ = max(x,0)$ and $x^- = min(x,0)$, the implicit scheme of the HJB eq. (13) rewrites:

$$\frac{v^{n+1}_{i,j} - v^n_{i,j}}{\tau} + \rho\, v^{n+1}_{i,j} = u(c^n_{i,j}) \ + {v^{n+1}_{i,j,F}}' \, [z_j w + r a_i - c^n_{i,j,F}]^+ \ + \ {v^{n+1}_{i,j,B}}' \, [z_j w + r a_i - c^n_{i,j,B}]^- \ + \ \lambda_j (v^{n+1}_{i,-j} - v^{n+1}_{i,j})$$

$$\frac{v^{n+1}_{i,j} - v^n_{i,j}}{\tau} \ + \ \rho\, v^{n+1}_{i,j} = u(c^n_{i,j}) \ + \ \frac{v^n_{i+1,j} - v^n_{i,j}}{h} \, [s^n_{i,j,F}]^+ \ + \ \frac{v^n_{i-1,j} - v^n_{i,j}}{h} \, [s^n_{i,j,B}]^- \ + \ \lambda_j (v^{n+1}_{i,-j} - v^{n+1}_{i,j})$$

where in the second line we only used the definition of ${v^{n+1}_{i,j,F}}'$ and ${v^{n+1}_{i,j,B}}'$. Collecting the terms with the same subscripts $(i-1, i, i+1)$, we can obtain:

$$\frac{v^{n+1}_{i,j} - v^n_{i,j}}{\tau} \ + \ \rho v^{n+1}_{i,j} = u(c^n_{i,j}) \ + \ v^{n+1}_{i-1,j}\, x_{i,j} \ + \ v^{n+1}_{i,j}\, y_{i,j} \ + \ v^{n+1}_{i+1,j}\, z_{i,j} \ + \ v^{n+1}_{i,-j}\, \lambda_j \tag{17}$$
$$x_{i,j} = -\frac{(s^n_{i,j,B})^-}{h},$$
$$y_{i,j} = -\frac{(s^n_{i,j,F})^+}{h} + \frac{(s^n_{i,j,B})^-}{h} - \lambda_j$$
$$z_{i,j} = \frac{(s^n_{i,j,F})^+}{h}$$

Note that we obtained a system of $I \times 2$ equations described in eq. (17). An important point lies at the boundary, where $x_{1,j} = z_{I,j} = 0, (j = 1, 2)$. Thus $v^{n+1}_{0,j}$ and $v^{n+1}_{I+1,j}$ will never be used.

As a consequence of the various remarks, we can write the system with matrix notations:

$$\frac{v^{n+1} - v^n}{\tau} + \rho v^{n+1} = u + \mathbf{A}^n v^{n+1}$$

This system can in turn be written as

$$\mathbf{B}^n v^{n+1} = b^n, \qquad \text{with} \qquad \mathbf{B}^n = \left(\frac{1}{\tau} + \rho\right) \mathbf{I} - \mathbf{A}^n, \quad b^n = u^n + \frac{1}{\tau} v^n$$

This system of linear equations can be solved very efficiently using sparse matrix routines, since $\mathbf{A}^n$ and thus $\mathbf{B}^n = \left(\frac{1}{\tau} + \rho\right)\mathbf{I} - \mathbf{A}^n$ are both tridiagonal by blocks.
In particular, we have the following form for matrix $\mathbf{A}$:

$$\mathbf{A}^n = \begin{bmatrix} y_{1,1} & z_{1,1} & 0 & \cdots & 0 & \lambda_1 & 0 & 0 & \cdots & 0 \\ x_{2,1} & y_{2,1} & z_{2,1} & 0 & \cdots & 0 & \lambda_1 & 0 & 0 & \cdots \\ 0 & x_{3,1} & y_{3,1} & z_{3,1} & 0 & \cdots & 0 & \lambda_1 & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & x_{I,1} & y_{I,1} & 0 & 0 & 0 & 0 & \lambda_1 \\ \lambda_2 & 0 & 0 & 0 & 0 & y_{1,2} & z_{1,2} & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 & 0 & x_{2,2} & y_{2,2} & z_{2,2} & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 & 0 & 0 & x_{3,2} & y_{3,2} & z_{3,2} & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & \cdots & \cdots & 0 & \lambda_2 & 0 & \cdots & 0 & x_{I,2} & y_{I,2} \end{bmatrix}, \quad u^n = \begin{bmatrix} u(c^n_{1,1}) \\ \vdots \\ \vdots \\ u(c^n_{I,1}) \\ u(c^n_{1,2}) \\ \vdots \\ \vdots \\ u(c^n_{I,2}) \end{bmatrix}$$

Let us make few remarks on this matrix: First, one can notice it represents the infinitesimal generator of the stochastic process with drift s^n and jumping to the other state $-j$ with intensity λ_j. It is indeed a Markovian Jump process transition matrix on the discretized state space (a_i, z_j): all rows sum to zero and the diagonal terms are non-positive and off-diagonal terms are non-negative.

This intensity matrix will thus have nice properties for the numerical algorithm. We observe that if $\mathbf{A}^n$ has rows that sums to zeros, the matrix $-\mathbf{A}^n$ will be diagonally dominant (but not strict!)[30].

Consequently, provided that $\tau < \infty, \rho > 0$ we obtain that $\mathbf{B}^n$ is a M-matrix – since $\mathbf{B}^n = \left(\frac{1}{\tau} + \rho\right)\mathbf{I} - \mathbf{A}^n$. The scheme will thus will be monotonous.

Summary of Algorithm to solve the HJB equation.

1. Guess $v^0_{i,j}, i = 1, ..., I, j = 1, 2$ and for $n = 0, 1, 2, ...$ follow:
2. Compute $(v^n_{i,j})'$ using the current guess of the value function and the upwind scheme (forward difference when drift is positive, backward difference when drift is negative)
3. Compute c^n from $c^n_{i,j} = (u')^{-1}[(v^n_{i,j})']$
4. Find v^{n+1} by solving the system of linear equations involving the matrix $\mathbf{A}$ described above (implicit scheme)
5. If v^{n+1} is close enough to v^n: stop. Otherwise, go to step 2.

[30] A matrix B is diagonally dominant if $\forall i, |B_{ii}| \geq \sum_{j \neq i} |B_{ij}|$ (strictly dominant if the inequality is strict).
A matrix $\tilde{B}$ is a M-matrix if $\tilde{B_{ii}} > 0$ and $\tilde{B_{ij}} \leq 0, \forall j \neq i$ and $\forall i, B_{ii} > -\sum_{j \neq i} B_{ij}$
A matrix $\bar{B}$ is monotonous if it is invertible and its inverse is positive.

B.2 Solving the Kolmogorov Forward/ Fokker Planck equation

Recall the Fokker Planck in the evolution case:

$$0 = \partial_t g_j(t,a) + \frac{d}{da}[s_j(a)\, g_j(t,a)] + \lambda_j g_j(t,a) - \lambda_{-j} g_{-j}(t,a)$$

Subject to the 'constraint' $\int_{\underline{a}}^{\infty} g_1(t,a)da + \int_{\underline{a}}^{\infty} g_2(t,a)da = K(t,r)$ The discretization scheme – using the same grid as before – will aim at solving the equation:

$$0 = \frac{g_{i,j}^{n+1} - g_{i,j}^{n}}{\tau} + [s_{i,j}^{n}\; g_{i,j}^{n}]' + \lambda_j g_{i,j}^{n+1} - \lambda_{-j} g_{i,-j}^{n+1}$$

Conversely to the HJB, which is fully non-linear, the Fokker Planck equation is linear, and so will be the finite-difference scheme. The main advantage is thus that the solution g can be obtained in one step.

Upwind scheme. Similarly as in the previous case, we use a upwind scheme. However, since the direction of time if *forward*, the scheme is reversed: When drift is positive, the *backward* approximation is then used, and conversely, when the drift is negative one need to use the forward approximation:

$$0 = \frac{g_{i,j}^{n+1} - g_{i,j}^{n}}{\tau} + \frac{[s_{i,j,F}^{n}]^{+}\, g_{i,j}^{n} - [s_{i-1,j,F}^{n}]^{+}\, g_{i-1,j}^{n}}{h} + \frac{[s_{i+1,j,B}^{n}]^{+}\, g_{i+1,j}^{n} - [s_{i,j,B}^{n}]^{+}\, g_{i,j}^{n}}{h} + \lambda_j g_{i,j} - \lambda_{-j} g_{i,-j}$$

Note that because $g_{0,j}$ and $g_{I+1,j}$ are outside the state-space, the measure g will be null on these points, and thus $[s_{0,j}^{n}]^{+}$ and $[s_{I+1,j}^{n}]^{-}$ will never be used.

Matrix reformulation Rewriting the scheme, collecting the terms with the same subscripts $(i-1, i, i+1)$, we can obtain:

$$\frac{g_{i,j}^{n+1} - g_{i,j}^{n}}{\tau} \;+\; g_{i+1,j}^{n+1}\, x_{i+1,j} \;+\; g_{i,j}^{n+1}\, y_{i,j} \;+\; g_{i-1,j}^{n+1}\, z_{i-1,j} \;+\; g_{i,-j}^{n+1}\, \lambda_{-j} \tag{18}$$

$$\begin{aligned} x_{i+1,j} &= -\frac{(s_{i+1,j,B}^{n})^{-}}{h}, \\ y_{i,j} &= -\frac{(s_{i,j,F}^{n})^{+}}{h} + \frac{(s_{i,j,B}^{n})^{-}}{h} - \lambda_j \\ z_{i-1,j} &= \frac{(s_{i-1,j,F}^{n})^{+}}{h} \end{aligned}$$

This discretization scheme yields the following system :

$$0 = \frac{g^{n+1} - g^{n}}{\tau} \;+\; (\mathbf{A}^{n})^{T} g^{n+1}$$

Again, it can be solved implicitly:

$$\mathbf{C}^n g^{n+1} = g^n \qquad \text{with} \qquad \mathbf{C}^n = (\mathbf{I} - \tau(\mathbf{A}^n)^T) \tag{19}$$

Stationary distribution. In the stationary case, it amounts at solving the system:

$$0 = (\mathbf{A}^n)^T g^n$$

where the matrix $\mathbf{A}^T$ is the transposed of the intensity matrix $\mathbf{A}$ found in the HJB approximation. In such situation the distribution g^n can be found in one step.

However, one issue lies in the fact that the matrix $\mathbf{A}^T$ is singular and cannot be inverted. One need to 'fix' this by applying any value to a given (i, j), say $g_{i,j} = 0.1$. Since the stationary distribution of any Markovian process is unique modulo a multiplicative constant, we only need to renormalize to obtain: $\int_{\underline{a}}^{\infty} g_1(t,a)da + \int_{\underline{a}}^{\infty} g_2(t,a)da = 1$.
In practice, one solve this issue by setting the value of the diagonal of the matrix $(A^T)_{kk} = 1$ and $b_k = 0.1$ for any k, then solve for the value $\tilde{g}^n$ and the renormalization is $g^n_{i,j} = \widetilde{g_{i,j}}^n / (\sum_{i=1}^{I} \widetilde{g_{i,1}} h + \widetilde{g_{i,2}} h)$.

Moreover, this will allow to solve for the initial distribution when this one is unknown. More specifically, this will be the case in the application in the next section.

Transition case. In the case where shocks occur and economic fundamentals change, one can solve implicitly the system as in eq. (19), with τ the time step. Given the initial distribution g^0 computed above – as in the stationary case – computing the evolution will simply implies:

$$g^{n+1} = (\mathbf{C}^n)^{-1} g^n$$

FP as the adjoint of HJB The matrix $\mathbf{A}^T$ is the transposed of the intensity matrix found in the HJB. That is, once the HJB equation is solved, we basically get the Kolmogorov Forward equation "for free" ("two birds one stone", cf. B. Moll). These methods take advantage of the property of the Fokker-Planck being the adjoint of the HJB – when optimal control is reached. Since the matrix $\mathbf{A}^n$ is the discretized version of the infinitesimal generator $\mathcal{A}$ of the jump process, the matrix $\mathbf{A}^T$ will naturally be the discretized version of the adjoint operator $\mathcal{A}^*$ associated to the Fokker Planck equation (c.f. the appendix for details).

B.3 Equilibrium on capital market

The last step of the algorithm is to update the price and quantities for the market to clear. The market clearing conditions imposes :

$$\int_{\underline{a}}^{\infty} a\, g_1(t,a)da + \int_{\underline{a}}^{\infty} a\, g_2(t,a)da =: S(t,r) = K(t,r) := \left(\frac{\alpha A_t}{r_t + \delta}\right)^{\frac{1}{1-\alpha}} z_{av}$$

First, one of the theoretical result of Achdou, Han, Lasry, Lions, and Moll (2017) provides a proof of unicity:

Under the assumption that IES $:= -\frac{u'(c)}{u''(c)c} \geq 1$ (i.e. $\frac{1}{RA} \leq 1$) and that the household is credit constrained $a \geq 0$, then we have:

(i) the optimal consumption $c_j(a,r) := (u')^{-1}(\partial_a v_j)$ is decreasing in interest rate r

(ii) the optimal saving $s(a,r) := s_j(t,a) = z_j w_t + r_t\, a - c_j(a,r)$ is increasing in r,

(iii) the stationary distribution shift rightward in r, i.e. the cdf $G_j(a;r)$ is strictly decreasing in r for all a in its support

(iv) $S(r)$ is strictly increasing and, since K(r) is clearly decreasing, the model has a unique stationary equilibrium.

In order to find iteratively this unique equilibrium, there exits several solutions:

One can iterate on quantities (B. Moll's codes). Given the supply of capital from household saving update the capital demanded by firms: $K^{new}(t,r) := \theta K^{old}(t,r) + (1-\theta) S(t,r)$ and then compute the interest rate as the marginal return of capital stock. The second option is to update the prices, i.e. interest rates, to change household saving behavior. Given the interest rate formula: $r(K) = \alpha\, A\, K^{\alpha-1} z_{av}^{1-\alpha} - \delta$, update the interest rate following the scheme:

$$r^{new} = \theta\; r(K(r^{old})) + (1-\theta)\; r(S(r^{old}))$$

The main advantage of this method is to be able to update in one step the *path* r_t for $t \in [0,T]$ given the *paths* K_t and S_t. When dealing with the stationary equilibrium however, another simple technique is the bisection method.

C Krusell-Smith model, larger binomial trees and quantization trees

In the following, we will use several concrete examples:

– Binomial/K-nomial tree:

(i) A simple tree, where $K = 3$ branches per nodes and $M = 3$ "waves" of shocks,

(ii) A "long" tree, where $K = 2$ and $M = 9$, where the time discretization is finer

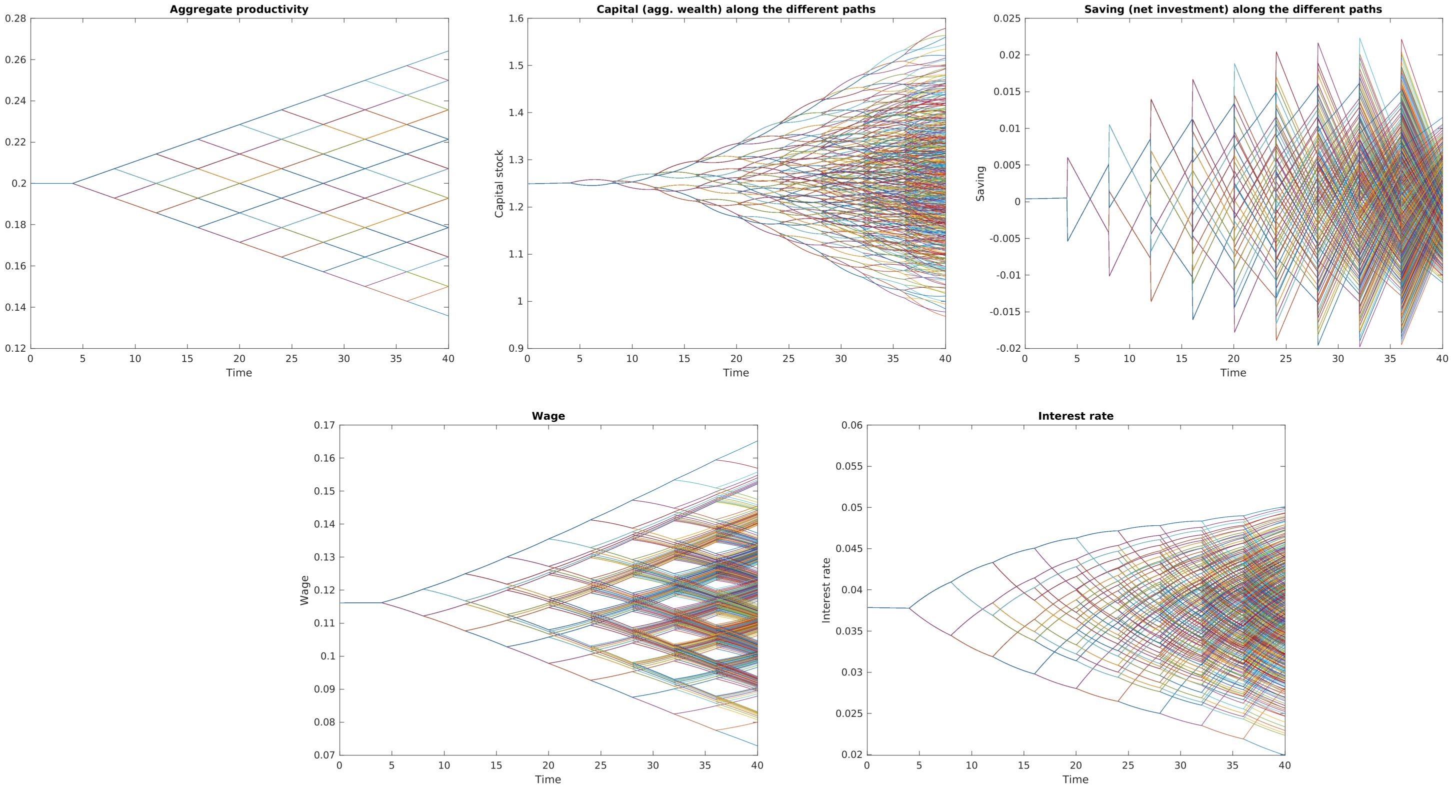
(iii) a "large tree", where $K = 7$ and $M = 3$ the (space) quantization grid is finer

– Quantization tree:

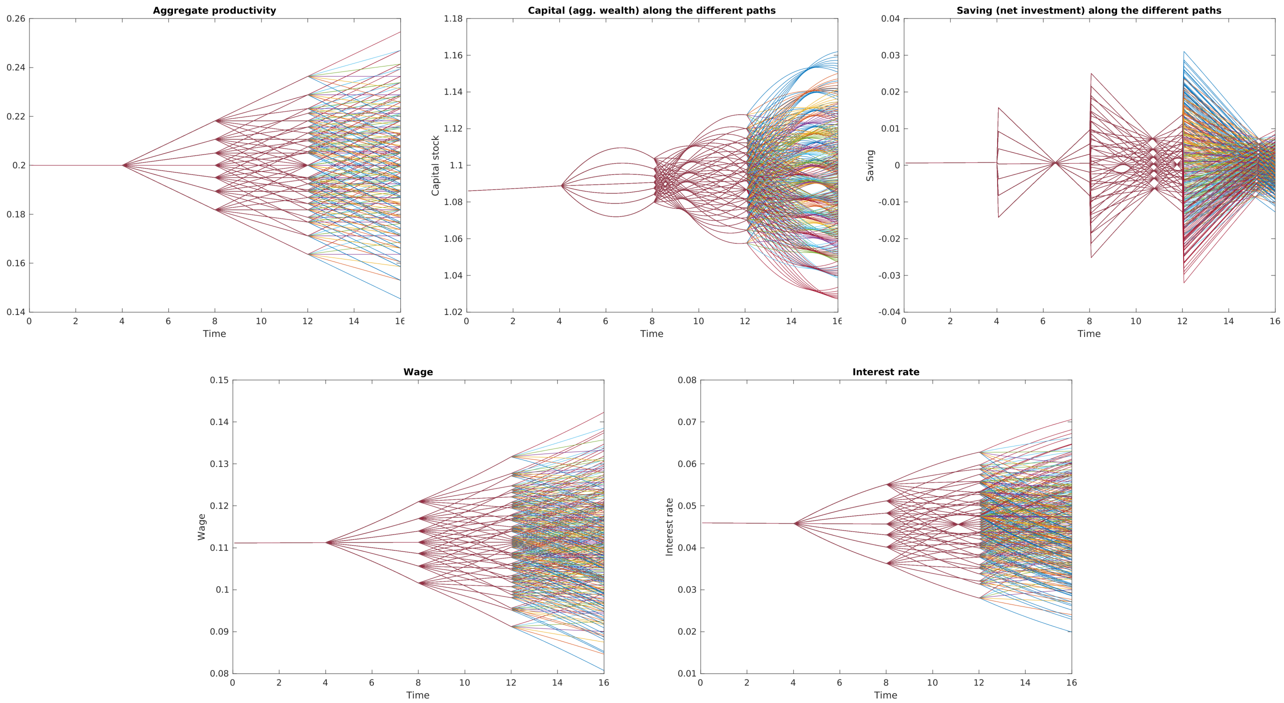
(iv) A simple tree, where $K = 3$ branches per nodes and $M = 3$ "waves" of shocks

(v) A larger tree, where $K = 5$ and $M = 4$, where time & space discretization are slightly finer.

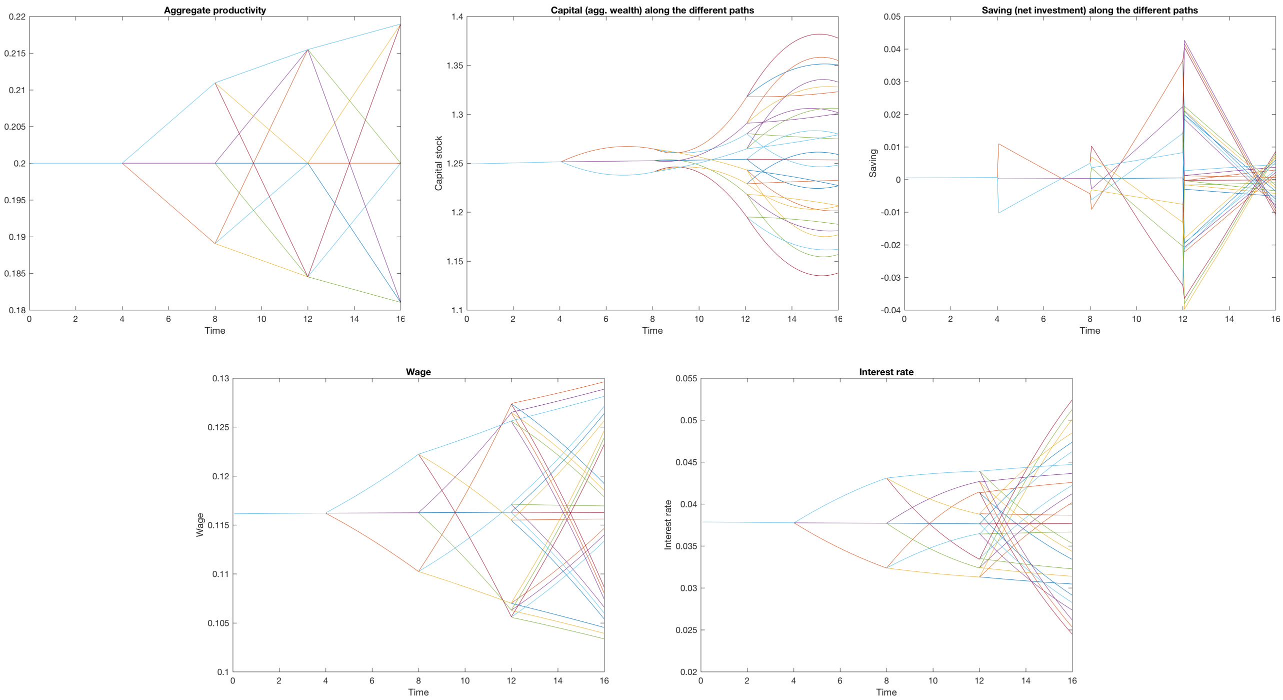
C.1 Long binomial tree, $K = 2$ and $M = 9$



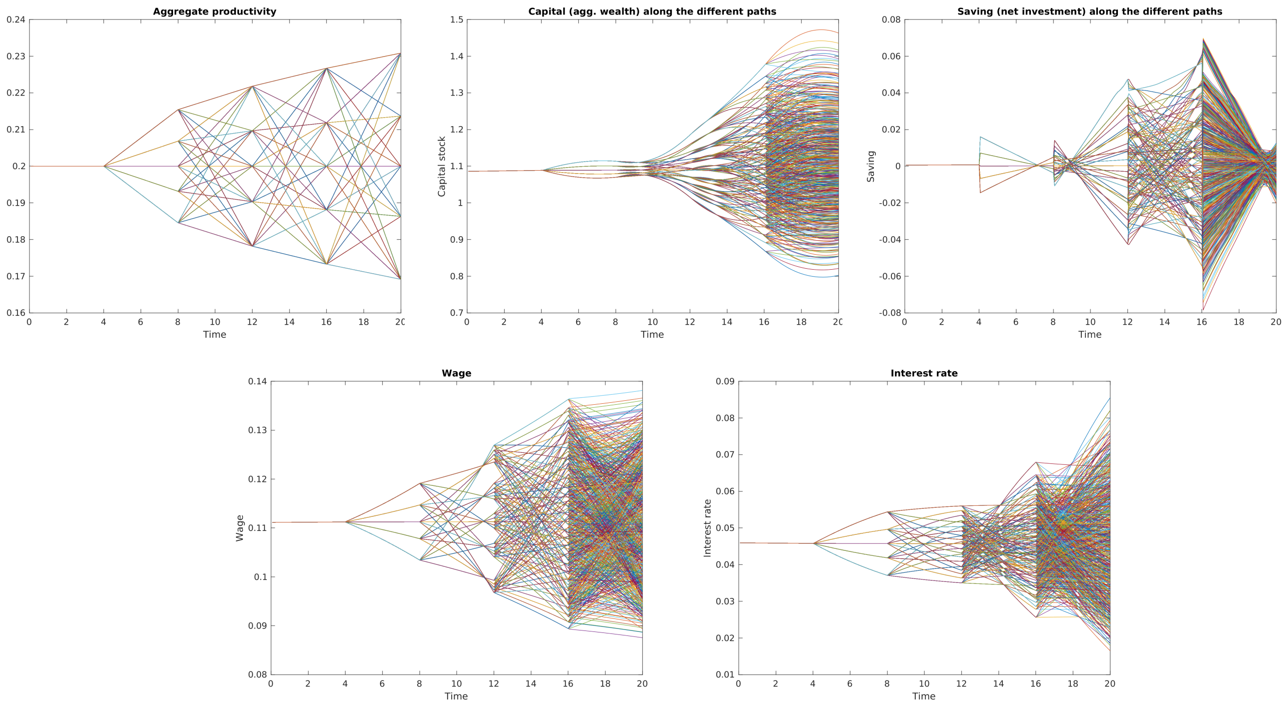
C.2 Large K-nomial tree, $K = 7$ and $M = 3$



C.3 Simple quantization tree, where $K = 3$ and $M = 3$



C.4 Larger quantization tree, where $K = 5$ and $M = 4$



D Representative agent models

The relations of the following models stem from the optimal (stochastic) control problem. Since there is only one agent, or a continuum of agents that are identical since the market is complete, we can solve the model using standard Dynamic programming techniques. The model is purely backward (the agent is forward looking) and is described in discrete time. The models and IRF are simulated using DYNARE (cf. the CEPREMAP working paper Adjemian et al. (2011)), a software solving rational expectations models using perturbation methods. In the IRF above, the calibration is the same as in the MFG with common noise.

The equilibrium relations of the Brock-Mirman model

This model is the analogous of the Krusell Smith model with a representative agent choosing consumption/saving in function of income/interest rate and expectation of future shocks.

$$\begin{aligned}
1 &= \mathbb{E}_t\big(\beta(1+r_{t+1})\frac{\partial_c u(C_{t+1})}{\partial_c u(C_t)}\big) && \text{(Euler equation)}\\
K_{t+1} &= (1-\delta)K_t + Y_t - C_t && \text{(Law of motion of Capital)}\\
Y_t &= A_t K^{\alpha} z_{av}^{1-\alpha} && \text{(Production)}\\
w_t &= (1-\alpha)\, A_t\, K_t^{\alpha}\, z_{av}^{-\alpha} && \text{(Wage)}\\
r_t &= \alpha\, A_t\, K_t^{\alpha-1} z_{av}^{1-\alpha} - \delta && \text{(Interest rate)}\\
A_t &= \bar{A}_0\, e^{\widetilde{A}_t} \qquad \widetilde{A}_{t+1} = (1-\theta)\widetilde{A}_t + \varepsilon_t && \text{(Agg. productivity)}
\end{aligned}$$

The equilibrium relations of the RBC model

This Real Business Cycle model is the analogous of the KS model with endogenous labor supply. The household now solve a control problem with labor supply – creating an intra-temporal "trade off".

$$\begin{aligned}
1 &= \mathbb{E}_t\big(\beta(1+r_{t+1})\frac{u_c(C_{t+1})}{u_c(C_{t+1})}\big) && \text{(Euler equation)}\\
z_{av} w_t &= \frac{\partial_\ell \tilde{u}(L_t)}{\partial_c \bar{u}(C_t)} && \text{(Labor-Conso trade-off)}\\
K_{t+1} &= (1-\delta)K_t + Y_t - C_t && \text{(Law of motion of Capital)}\\
Y_t &= A_t K^{\alpha}(z_{av} L_t)^{1-\alpha} && \text{(Production)}\\
w_t &= (1-\alpha)\, A_t\, K_t^{\alpha}\, L_t^{-\alpha} && \text{(Wage)}\\
r_t &= \alpha\, A_t\, K_t^{\alpha-1} z_{av} L_t^{1-\alpha} - \delta && \text{(Interest rate)}\\
A_t &= \bar{A}_0\, e^{\widetilde{A}_t} \qquad \widetilde{A}_{t+1} = (1-\theta)\widetilde{A}_t + \varepsilon_t && \text{(Agg. productivity)}
\end{aligned}$$

The equilibrium relations of the New Keynesian model

The usual NK model, found for example in reference by Galí (2008) or Woodford (2003) are usually log-linearized by hand - comparing the deviation of variables to their steady state values e.g. $z_t = \log(Z_t) - \log(Z)$. Here is the summary of the relations of this model. Since, in this case, the price-setting of firm follows the Calvo-Yun framework, there are sustantial differences between the following model and the one developed in the above MFG. In particular, in this setting there is no government (no tax, no lump-sum redistribution) and no clear profit redistributed to households. These two points imply major differences between output and consumption in our framework while the market clearing is simply $C_t = Y_t$ in the standard NK model.

However, the Euler equation and labor-consumption trade-offs are preserved. All these relations can summarize the simple 'three equations' model (IS-curve, NK-Philips curve and interest rule).

$$
\begin{aligned}
c_t &= \mathbb{E}_t(c_{t+1}) - \tfrac{1}{\sigma}(i_t - \mathbb{E}_t(\pi_{t+1}) - \rho) && \text{(Euler equation)}\\
w_t - p_t &= \phi \ell_t + \sigma c_t && \text{(Labor-Conso trade-off)}\\
y_t &= a_t + \ell_t && \text{(Production)}\\
\widetilde{y}_t &= \mathbb{E}_t(\widetilde{y}_{t+1}) - \tfrac{1}{\sigma}(i_t - \mathbb{E}_t(\pi_{t+1}) - r_t^n) && \text{(Dynamic IS equation)}\\
\pi_t &= \beta \mathbb{E}_t(\pi_{t+1}) + \kappa \widetilde{y}_t && \text{(NK Phillips Curve)}\\
r_t &= i_t - \pi_t && \text{(Real interest rate)}\\
i_t &= \rho + \phi_\pi \pi_t && \text{(Interest rate rule)}\\
r_t^n &= \rho + \sigma 1 + \phi\sigma + \phi(a_{t+1} - a_t) && \text{(Natural rate)}\\
\widetilde{A}_{t+1} &= (1-\theta)\widetilde{A}_t + \varepsilon_t && \text{(Agg. productivity shock)}
\end{aligned}
$$

References

Achdou, Y., F. Camilli, and I. Capuzzo-Dolcetta (2013). Mean field games: convergence of a finite difference method. *SIAM Journal on Numerical Analysis 51*(5), 2585–2612.

Achdou, Y. and I. Capuzzo-Dolcetta (2010). Mean field games: Numerical methods. *SIAM Journal on Numerical Analysis 48*(3), 1136–1162.

Achdou, Y., J. Han, J.-M. Lasry, P.-L. Lions, and B. Moll (2017, November). Income and wealth distribution in macroeconomics: A continuous-time approach. *R & R, Review of Economic Studies* (NBER 23732).

Adjemian, S., H. Bastani, M. Juillard, F. Mihoubi, G. Perendia, M. Ratto, and S. Villemot (2011). Dynare: Reference manual, version 4.

Ahn, S., G. Kaplan, B. Moll, T. Winberry, and C. Wolf (2018). When inequality matters for macro and macro matters for inequality. *NBER Macroeconomics annual 32*(1), 1–75.

Aiyagari, S. R. (1994). Uninsured idiosyncratic risk and aggregate saving. *The Quarterly Journal of Economics 109*(3), 659–684.

Algan, Y., O. Allais, and W. J. Den Haan (2008). Solving heterogeneous-agent models with parameterized cross-sectional distributions. *Journal of Economic Dynamics and Control 32*(3), 875–908.

Bally, V., G. Pagès, and J. Printems (2005). A quantization tree method for pricing and hedging multidimensional american options. *Mathematical Finance 15*(1), 119–168.

Barles, G. and P. E. Souganidis (1991). Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic analysis 4*(3), 271–283.

Benhabib, J., A. Bisin, and S. Zhu (2011). The distribution of wealth and fiscal policy in economies with finitely lived agents. *Econometrica 79*(1), 123–157.

Benhabib, J., A. Bisin, and S. Zhu (2015). The wealth distribution in bewley economies with capital income risk. *Journal of Economic Theory 159*, 489–515.

Bertucci, C., J. M. Lasry, and P. L. Lions (2018). Some remarks on mean field games.

Bewley, T. (1986). Stationary monetary equilibrium with a continuum of independently fluctuating consumers. *Contributions to mathematical economics in honor of Gérard Debreu 79.*

Bhandari, A., D. Evans, M. Golosov, and T. J. Sargent (2017). Inequality, business cycles and fiscal-monetary policy. Technical report, University of Minnesota working paper.

Capuzzo-Dolcetta, I. and P.-L. Lions (1990). Hamilton-jacobi equations with state constraints. *Transactions of the American Mathematical Society 318*(2), 643–683.

Cardaliaguet, P. (2013/2018). Notes on mean field games. *Lecture notes from P.L. Lions' lectures at College de France and P. Cardaliaguet at Paris Dauphine*.

Cardaliaguet, P., F. Delarue, J.-M. Lasry, and P.-L. Lions (2017). The master equation and the convergence problem in mean field games. *arXiv preprint arXiv:1509.02505*.

Carmona, R. and F. Delarue (2014). The master equation for large population equilibriums. In *Stochastic Analysis and Applications 2014*, pp. 77–128. Springer.

Carmona, R. and F. Delarue (2018). *Probabilistic Theory of Mean Field Games with Applications I-II.* Springer.

Carmona, R., F. Delarue, and A. Lachapelle (2013). Control of mckean–vlasov dynamics versus mean field games. *Mathematics and Financial Economics 7*(2), 131–166.

Carmona, R., F. Delarue, and D. Lacker (2016). Mean field games with common noise. *The Annals of Probability 44*(6), 3740–3803.

Challe, E., J. Matheron, X. Ragot, and J. F. Rubio-Ramirez (2016). Precautionary saving and aggregate demand.

Corlay, S. (2011). *Some aspects of optimal quantization and applications to finance.* Ph. D. thesis, Université Pierre et Marie Curie-Paris VI.

Corlay, S. and G. Pagès (2015). Functional quantization-based stratified sampling methods. *Monte Carlo Methods and Applications 21*(1), 1–32.

Fahim, A., N. Touzi, and X. Warin. A probabilistic numerical method for fully nonlinear parabolic pdes. *Ann. Appl. Probab.* (4), 1322–1364.

Fleming, W. H. and H. M. Soner (2006). *Controlled Markov processes and viscosity solutions*, Volume 25. Springer Science & Business Media.

Gabaix, X., J.-M. Lasry, P.-L. Lions, and B. Moll (2016). The dynamics of inequality. *Econometrica 84*(6), 2071–2111.

Galí, J. (2015). *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications.* Princeton University Press.

Hartl, R. F., S. P. Sethi, and R. G. Vickson (1995). A survey of the maximum principles for optimal control problems with state constraints. *SIAM review 37*(2), 181–218.

Heathcote, J., K. Storesletten, and G. L. Violante (2014). Consumption and labor supply with partial insurance: An analytical framework. *American Economic Review 104*(7), 2075–2126.

Huang, M., R. P. Malhamé, and P. E. Caines (2006). Large population stochastic dynamic games: closed-loop mckean-vlasov systems and the nash certainty equivalence principle. *Communications in Information & Systems 6*(3), 221–252.

Kaplan, G., B. Moll, and G. L. Violante (2018). Monetary policy according to hank. *American Economic Review 108*(3), 697–743.

Kaplan, G. and G. L. Violante (2014). A model of the consumption response to fiscal stimulus payments. *Econometrica 82*(4), 1199–1239.

Krusell, P. and A. A. Smith, Jr (1998). Income and wealth heterogeneity in the macroeconomy. *Journal of political Economy 106*(5), 867–896.

Lasry, J.-M. and P.-L. Lions (2007). Mean field games. *Japanese journal of mathematics 2*(1), 229–260.

Luschgy, H. and G. Pagès (2002). Functional quantization of gaussian processes. *Journal of Functional Analysis 196*(2), 486–531.

Pagès, G. (2017). Introduction to numerical probability for finance. *UPMC, Polycopié du 'Master 2 Probabilités et Finance'*.

Pagès, G. and J. Printems (2003). Optimal quadratic quantization for numerics: the gaussian case. *Monte Carlo Methods and Applications mcma 9*(2), 135–165.

Pham, H. and X. Wei (2017). Dynamic programming for optimal control of stochastic mckean–vlasov dynamics. *SIAM Journal on Control and Optimization 55*(2), 1069–1101.

Ragot, X. (2018). Heterogeneous agents in the macroeconomy: reduced-heterogeneity representations. In *Handbook of Computational Economics*, Volume 4, pp. 215–253. Elsevier.

Reiter, M. (2010). Solving the incomplete markets model with aggregate uncertainty by backward induction. *Journal of Economic Dynamics and Control 34*(1), 28–35.

Soner, H. M. (1986). Optimal control with state-space constraint i. *SIAM Journal on Control and Optimization 24*(3), 552–561.

Souganidis, P. E. (1985). Approximation schemes for viscosity solutions of hamilton-jacobi equations. *Journal of differential equations 59*(1), 1–43.

Winberry, T. (2016). A toolbox for solving and estimating heterogeneous agent macro models. *Forthcoming Quantitative Economics*.

Woodford, M. (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

Yong, J. and X. Y. Zhou (1999). *Stochastic controls: Hamiltonian systems and HJB equations*, Volume 43. Springer Science & Business Media.

Young, E. R. (2010). Solving the incomplete markets model with aggregate uncertainty using the krusell–smith algorithm and non-stochastic simulations. *Journal of Economic Dynamics and Control 34*(1), 36–41.