When is aggregation enough? The Master Equation and Projection for Heterogeneous Agent models with Aggregate Risk

PRELIMINARY – WORK IN PROGRESS

Most recent version accessible here

Thomas Bourany*
COLUMBIA UNIVERSITY / EIEF

April 2025

Abstract

I propose a global approach to study heterogeneous agents (HA) models with aggregate risk. I build on the Master Equation representation, studied in the Mean Field Games literature, where the value function takes the infinite-dimensional distribution as a state variable. In that context, the projection of the distribution on a finite set of moments, as in Krusell and Smith (1998), provides analytical insights and yields finite-dimensional HJB and KFE that can be solved using standard numerical methods. I show how to preserve rational expectations to ensure that agents' forecasts are consistent with equilibrium dynamics. This method bypasses the constraints of perturbation methods, which rely on certainty equivalence, and other approaches used in the literature. I demonstrate the method's potential for studying aggregate uncertainty by applying it to the Krusell-Smith model, with substantial speed gains. I show how to implement the method using higher-order moments, investigate why this class of models exhibits "approximate aggregation", and test the robustness of the bounded-rationality assumptions used in other methods. I also illustrate how the analysis extends to richer environments, like price-setting models or portfolio choice problems.

*Thomas Bourany, *tb3219@columbia.edu*. I thank Pierre Cardaliaguet and Yves Achdou, who introduced me to Mean Field Games and the Master equation many years ago. All errors are mine

1 Introduction

Heterogeneous-agent models in macroeconomics have fundamentally changed our understanding of inequality and the distributional consequences of aggregate fluctuations. Yet the analysis of aggregate uncertainty in these environments remains constrained by a fundamental computational and conceptual difficulty. Once aggregate risk is introduced, individual decisions depend not only on idiosyncratic states and aggregate shocks but also on beliefs about the future evolution of the cross-sectional distribution of agents. Under rational expectations, this distribution becomes a state variable in the recursive dynamic programming problem. As a result, the Bellman equations are naturally defined on an infinite-dimensional space, and equilibrium requires consistency between individual forecasts, aggregate prices, and the law of motion of the distribution. This is the core difficulty behind dynamic heterogeneous-agent models with aggregate risk, and this curse of dimensionality has long limited the scope of global solution methods and our comprehension of business-cycle and stabilization policies in the presence of inequality. ï¿¼

This paper proposes a global approach to that problem by combining the Master Equation representation with a finite-dimensional projection of the cross-sectional distribution. The starting point is the observation that the Master Equation provides a natural recursive formulation of heterogeneous-agent economies under aggregate risk: the value function depends on individual states and on the distribution of agents, while that distribution evolves endogenously through individual policy rules. This representation is powerful because it makes explicit the equilibrium dependence of current decisions on the future evolution of the entire economy. At the same time, this richness is precisely what makes the problem difficult to solve in practice. The contribution of this paper is to show that, in many economically relevant settings, one can project the distribution onto a finite set of moments or sufficient statistics, and thereby transform the infinite-dimensional recursive problem into a lower-dimensional system that remains disciplined by equilibrium conditions and rational expectations.

The paper is motivated by the classic Krusell-Smith insight that aggregate dynamics in HA models can be well approximated by a small number of distributional statistics, especially aggregate capital, see . In the original approach, however, tractability relies on a parametric forecasting rule, typically linear in the first moment, and on an iterative Monte Carlo simulation procedure used to estimate its coefficients. While this strategy has been remarkably influential, it leaves several questions unanswered. It does not derive the forecasting law directly from the structure of the individual optimization problem; it imposes a functional form on beliefs rather than recovering it from equilibrium; and it makes it difficult to assess systematically when approximation by a small set of moments is justified, and when it fails. The present paper addresses these limitations by replacing the exogenous parametrization of beliefs with a projection-based solution of the Master Equation itself. In this framework, the aggregate law of motion is not guessed and estimated externally. Instead, it emerges from the aggregation of optimal policy rules obtained from the projected recursive problem.

The method developed here can thus be seen as an intermediate solution between the full-information rational expectation equilibrium and reduced-form bounded-rationality approaches. On the one hand, it preserves internal consistency: agents' decisions are derived from a recursive problem in which their forecasts are tied to equilibrium dynamics, rather than to ad hoc statistical laws. On the other hand, it acknowledges that the full cross-sectional distribution may be too rich an object to track exactly, and therefore allows agents to base their decisions on a lower-dimensional representation of the economy. In that sense, the paper formalizes further the concept of restricted yet disciplined equilibrium concept, close in spirit to restricted-perception approaches, while remaining firmly anchored in the structural logic of dynamic general equilibrium. This feature is central for the paper's broader argument: the question is not only how to solve heterogeneous-agent models faster, but also how to identify when aggregation onto a few statistics is sufficient for equilibrium analysis, and when distributional detail matters in an essential way.

Beyond tractability, the paper speaks to an important substantive issue in macroeconomics. Many of the questions for which heterogeneous-agent models are most valuable are precisely those where aggregate risk, nonlinearity, and time-varying dispersion interact: precautionary behavior, asset pricing, portfolio choice, financial amplification, and the transmission of macro shocks in environments with incomplete insurance. These applications require methods that go beyond first-order perturbation and local approximations around a stationary distribution. They also require a framework in which one can study how higher-order moments, tail behavior, or mass points in the distribution shape equilibrium outcomes. By working directly with the Master Equation and allowing the projection to include not only the mean but also higher moments and other statistics, the paper offers a systematic way to investigate the boundaries of approximate aggregation rather than taking them for granted.

The Krusell-Smith model provides the natural first laboratory for this analysis. It combines uninsurable idiosyncratic risk, borrowing constraints, and aggregate productivity shocks in a setting where the wealth distribution matters for prices and allocations. The paper first develops the relevant intuition in that benchmark environment, then recasts the model in continuous time and derives its Master Equation formulation. This allows the analysis to move from the familiar idea of moment-based approximation to a more general characterization of how distributional projections enter the recursive problem. From there, we can see how policy rules can be solved globally under aggregate uncertainty using standard numerical methods – e.g. finite difference – once the distribution has been projected onto a chosen set of statistics, and how the induced law of motion of aggregate variables follows from the aggregation of those policy functions. The resulting procedure yields substantial computational gains while preserving a transparent connection between microeconomic behavior and aggregate dynamics.

More broadly, the paper contributes to the ongoing effort to expand the set of heterogeneous-agent models that can be solved globally and interpreted structurally. Recent work has emphasized either perturbation-based representations of the Master Equation or simulation-based methods using Deep Learning or Reinforcement Learning that approximate the equilibrium laws of motion with flexible parametric tools. This paper takes a different route. Its objective is neither to linearize the problem nor to replace the recursive structure with a black-box approximation, but rather to exploit the analytical discipline of the Master Equation while reducing its state space through projection. In doing so, it provides a bridge between the mathematical theory of Mean Field Games and the computational practice of macroeconomics. The broader payoff is conceptual as much as numerical: it clarifies what is being approximated when economists speak of "aggregation" and it provides a framework for evaluating whether a given set of moments is genuinely sufficient for equilibrium analysis.

The rest of the paper develops this argument in stages. It begins with the Krusell-Smith model to illustrate the logic of the method and the role of projection in a familiar setting. It then derives the Master Equation in continuous time, studies how projections onto moments modify the recursive structure, and presents the associated numerical implementation. Finally, it shows how the same logic extends to richer environments, including applications where uncertainty, inequality, portfolio choice, and asset-pricing jointly determine aggregate dynamics. The central question throughout is the one posed in the title: when is aggregation enough? The answer proposed here is that aggregation is not a primitive shortcut, but an equilibrium approximation whose validity can be studied formally within the Master Equation itself.

2 Intuitions – Krusell-Smith Model in discrete time

We first provide an outline and intuitions for the method, using the workhorse model of Krusell-Smith, before turning to a general continuous-time framework.

Households solve a consumption-saving model, choosing consumption c, to maximize lifetime utility $\sum_t \beta^t u(c_t)$ subject to their budget constraint, with asset state a. They face (i) idiosyncratic income risk on their labor productivity z, (ii) incomplete market as this risk is uninsurable, (iii) credit constraint $a \geq \underline{a}$ preventing them from borrowing. Moreover, the economy is hit by (iv) aggregate shocks on aggregate TFP Z. This model yields expost heterogeneity, characterized by a distribution of households $g(a, z)$ over wealth and income.

These households rent their savings to a representative firm, which produces with a Cobb-Douglas production function, which yields interest rates r and wage w:

$$Y = ZK^{\alpha} \qquad \Rightarrow \qquad r = \alpha K^{\alpha-1} - \delta \qquad\qquad w = (1-\alpha)K^{\alpha}$$

As a result, the household problem solves the following Bellman equation, where they need to forecast the evolution of the distribution of other agents $g(\cdot)$, as well as their states (a, z) and aggregate condition Z:

$$
\begin{aligned}
v(a, z, g, Z) &= \max_{c, a'} \; u(c) + \beta \mathbb{E}^{z', Z'} \left[v(a', z', g', Z') \,\middle|\, z, Z \right] \\
s.t. \qquad c + a' &= zw + (1{+}r)a \\
g' &= H(g, Z, Z')
\end{aligned}
$$

where the last equation represents the law of motion of the distribution. In general equilibrium, this model yields the market clearing for assets:

$$
K = \int_{a,z} a \, dg(a, z) \qquad \forall \; Z
$$

Challenge: This model is untractable computationally as agents need to forecast the whole distribution g. With recursive methods, the value function $V(a, z, g, Z)$ depends on this infinite-dimensional object g: if one would approximate it with an histogram $\{p_1, \dots, p_M\}$, this would "add" M state variables to the Bellman equation, involving a curse of dimensionality. With sequential methods, one would need to follow the distribution $\{g_t\}_t$ on *every potential path* of the stochastic process $\{Z_t\}_t$. Such a brute force approach, explored in Achdou, Bourany (2018), is computationally intensive.

In their original article, Krusell-Smith's solution relies on two assumptions related to *bounded-rationality.* They assume that:

1. The Household approximates the dynamics of the distribution with its first moment, which is aggregate capital: $K = \iint a \, g(da, z)$.
2. They approximates its dynamics with a *linear* forecasting rule for future capital

$$
\log K' = a_1^Z \log K + a_2^Z
$$

To simulate the model, they would choose parameters (a_1^Z, a_2^Z) to match the *realized / simulated* path of capital $\{K_t\}_t$, resulting from the agents' consumption-saving choice, using Monte Carlo simulations of a series of shocks $\{z_t, Z_t\}_t$.

The goal of the method developed here is to bypass assumption 2 to go beyond the linearity – or even the parametrization – of the forecasting rule. To that end, we rely on the structure of the agent's decision problem and develop a general methodology based on the Master equation. This mathematical formalism allows us to formalize the infinite-dimensional Bellman equation for $V(\cdot, g)$. Moreover, one can do a *projection* of this distribution g and *aggregation* onto sufficient statistics, e.g. the mean $K = \int ag(da, z)$. This allows us to appropriately transform the Bellman equation from $V(\cdot, g)$ into $V(\cdot, K)$, and obtain agents' policy functions globally, e.g. optimal consumption $c(a, z, K, Z)$ and saving $a'(a, z, K, Z)$ under aggregate uncertainty. Aggregating such policies a' will provide us with the *realized* law of motion:

$$K' = h(K, Z)$$

which is a dynamic equation that *results* from the agents' optimization. This differs from other methods like Fernández-Villaverde et al. (2023) or Yang et al. (2026), which uses deep-learning or reinforcement learning such that $K' = h(K, Z, \theta)$, or $r' = \tilde{h}(r, Z, \theta)$ are approximated with a parametric function using (Monte Carlo) simulations of shocks and agents decisions.

The introduced method remains consistent with the Lucas critique, as the policies are never parameterized by statistical laws. Because the distribution is projected onto lower-dimensional moments, the model is misspecified, or restricted. However, the agents are still internally rational given those beliefs, and this could be considered as "restricted perception equilibrium". Additionally, the general approach allows testing the robustness of both assumptions 1 and 2, since the distribution can be projected onto any higher-order moments and sufficient statistics.

3 Master equation in the Krusell Smith model

We derive the Master equation in the concrete case of the Krusell-Smith model. We first recast the Aiyagari model in continuous time, following Achdou et al. (2022), and show that, in the absence of aggregate risk, we can still express the model with a Master equation. We then introduce aggregate risk in the firm's TFP. In the next sections, we will study how to tackle those challenges using projection methods, both analytically and numerically.

3.1 Primer on MFG and the Master Equation without aggregate risk

Without aggregate risk, this model boils down to the standard Aiyagari framework. One can write it as a system of coupled Partial Differential Equations, also called the Mean Field Games system, between the agents' decision and the law of motion of the distribution.

As before, households solve a consumption-saving problem, maximizing lifetime utility:

$$v(0, a, z) = \max_{\{c_t\}_{t=0}^{\infty}} \mathbb{E}_0 \int_0^{\infty} e^{-\rho t} u(c_t)\, dt \quad ,$$

choosing the consumption "control" c_t over time. They own assets a_t, and they face a credit constraint $a \geq \underline{a}$ and uninsurable idiosyncratic Markov risk on their labor productivity z_t, which takes discrete values $\{z_1, \dots z_{n_z}\}$. We consider two states $\{z_1, z_2\}$ for exposition purposes, with intensity of that jump process λ_j and λ_{-j} for $j = 1, 2$.

The distribution of households is a measure g over states $(a_t, z_t) \in \mathbb{X}=[\underline{a}, \infty) \times \{z_1 \dots z_{n_z}\}$, which can have mass points. We consider that $g(t, \cdot) \in \mathcal{P}_2(\mathbb{X})$ the space of measures on $\mathbb{X}$, indexed by time, that admit second moments: $\iint_{a,z} |a|^2 g(t, da, dz) < \infty, \forall t, z$, and similarly for z. We denote $K_t := \langle a, g_t \rangle := \iint_{a,z} a\, g(t, da, dz) := \sum_j \int_{\mathbb{X}} a g(t, da, z_j)$ the first moment of this distribution with respect to wealth a_t. We normalize labor supply $L_t = \langle z, g_t \rangle$ to a constant $\bar{L} = 1$ to alleviate the notations. With the budget constraint, the agents' state dynamics for saving follow, starting from an initial distribution g_0 on $[\underline{a}, \infty)$:

$$da_t = [z_t w_t + r_t\, a_t - c_t]dt \qquad\qquad a_t \geq \underline{a}$$

On the firm side, a representative firm uses the agents' savings and inelastic labor supply to produce the final good, using Cobb-Douglas technology $Y_t = Z_t K_t^{\alpha}$. For now, the aggregate TFP is deterministic (and constant) $Z_t = \bar{Z} = 1$, but we will allow that variable to be stochastic in a later section. The firm's optimality conditions imply the asset prices, interest rate, and wages (r_t, w_t):

$$r_t = \alpha \bar{Z} K_t^{\alpha-1} - \delta \qquad\qquad w_t = (1-\alpha)\bar{Z} K_t^{\alpha}\ .$$

As a result, the households, for their optimal consumption-saving decision, get a value $v(t,a,z)$ that solves a Hamilton-Jacobi-Bellman (HJB) Equation:

$$-\partial_t v(t,a,z) + \rho v(t,a,z) = \max_c\, u(c) + \mathcal{L}[\, v \,|\, c\,](t,a,z)\ , \tag{1}$$

for $t \in [0,\infty)$, $(a,z) \in \mathbb{X} = [\underline{a},\infty) \times \{z_1, z_2\}$, subject to the agents dynamics:

$$\begin{aligned} da &= s(a, z_j, g, c^\star)dt := \Big(z_j\, \underbrace{\bar{Z}\,(1-\alpha)\,\langle a,g\rangle^{\alpha}}_{=w} + \underbrace{(\bar{Z}\,\alpha \langle a,g\rangle^{\,\alpha-1} - \delta)}_{=r}\, a - c^\star \Big) dt \qquad a \geq \underline{a} \qquad \text{(wealth)} \\ dz_j &= dJ_j \qquad \text{with intensity} \quad \lambda(z_j, z_{-j}) = \lambda_j \qquad \text{(labor productivity)} \end{aligned} \tag{2}$$

where the transport/jump-operator of that HJB comes from agents' decision $c^\star = c^\star(t,a,z_j,g,v)$ and shocks z_j:

$$\mathcal{L}[\, v \,| c^\star](t,a,z_j) = \partial_a v(t,a,z_j)\, \underbrace{[z_j w + r\, a - c^\star]}_{\text{change in saving}} + \underbrace{\lambda_j (v(t,a,z_{-j}) - v(t,a,z_j))}_{\text{change in labor income}} \tag{3}$$

This "infinitesimal generator" is effectively a transition "matrix" in infinite dimensions and characterizes the movement of agents in the state-space $(a,z) \in \mathbb{X}$. We assume that at some long horizon, absent any shocks or deterministic dynamics, the value function stabilizes at a terminal condition $v_\infty(a,z_j) = \lim_{t\to\infty} v(t,a,z_j)$.

The driver of wealth dynamics is the optimal consumption $c^\star$. It maximizes the Hamiltonian $\mathcal{H}(t,a,z_j,g,v,c) = u(c) + \mathcal{L}[\, v\,|c](t,a,z_j)$, subject to the constraint, and as a result depends on the value function $v(\cdot)$. This yields:

$$c^\star(t,a,z_j,g,v) = \begin{cases} u'^{-1}\big(\partial_a v(t,a,z_j)\big) & a > \underline{a} \\ z_j\,(1-\alpha)\bar{Z}\,\langle a,g\rangle^{\alpha} + (\alpha \bar{Z} \langle a,g\rangle^{\alpha-1} - \delta)\, a & \text{if } a \leq \underline{a} \end{cases} \tag{4}$$

Second, the distribution of agents $g \in \mathcal{P}_2(\mathbb{X})$ reacts to agents' dynamics, and thus follows a Kolmogorov Forward Equation (KFE).[1]. Starting from an initial distribution g_0, the dynamics are given by:

$$\partial_t g(t,a,z) = \mathcal{L}^*[\,g\,|c^\star](t,a,z) \quad , \tag{5}$$

for $(a,z) \in \mathbb{X} = [\underline{a},\infty) \times \{z_1,z_2\}$. There is a duality between the HJB and the KFE: since the distribution dynamics come from agents' decisions, the operator $\mathcal{L}^*$ is the adjoint of the HJB operator $\mathcal{L}$. More intuitively, the asset distribution shifts and changes shape as agents save and consume.

In General Equilibrium, we obtain the market-clearing between agents' savings and the representative firm's capital demand:

$$\iint_{z,a\geq\underline{a}} a\, g(t,da,z_j) = K_t$$

We therefore obtain a system of PDEs – the Mean Field Games system – the HJB and the KFE, that are coupled thanks to this market clearing.

Primer on the Master Equation. The previous approach led to a system of PDEs. An alternative approach developed by Lions and Cardaliaguet, Delarue, Lasry and Lions (2019) relates to the *Master equation*. This combines, in *one single equation*, both the HJB and the KFE in the agents' optimization problem. A general treatment of this framework can be found in Cardaliaguet et al. (2019), and economic applications Schaab (2020), Bilal (2023), and Gu, Lauriere, Merkel and Payne (2024).

The general intuition is as follows. The value function $V(\cdot,g)$ depends on the state g. To write a more "general" HJB equation for $V(\cdot,g)$ using recursive methods, it would need to "follow" the dynamics of the states.[2] In this case, for the space of distribution g, one would follow the evolution of g using the KFE: $\partial_t g(t,a,z) = \mathcal{L}^*[\,g\,|c^\star](t,a,z)$.

The Master Equation is hence a combination of two parts: (i) a continuation value following the dynamics of the states $(a,z) \in \mathbb{X}$ and (ii) a continuation value for the dynamics of the distribution $g(t,a,z)$.

$$-\partial_t V(t,a,z,g) + \rho V(t,a,z,g) = \overbrace{\max_c\, u(c) + \mathcal{L}[\,V\,|c^\star](t,a,z)}^{\text{standard HJB continuation value}} + \underbrace{\iint_{z,a} \frac{\delta V(t,a,z,g)}{\delta g}[\tilde{a},\tilde{z}]\ \mathcal{L}^*[\,g\,|c^\star](t,d\tilde{a},\tilde{z})}_{\text{change in } V \text{ due to the distribution dynamics}} \tag{6}$$

[1] It is called the Fokker-Planck equation in the Mean Field Games literature, due to its origin in Mean Field theory and statistical physics

[2] In a simpler example without control, we can see clearly that if a state $dx = \mu(x)$, the backward equation is $-\partial_t v(t,x) + \rho v(t,x) = u(x) + v_x(x)\mu(x)$. One could use the same logic when x is a distribution g.

for $t \in [0, \infty)$, $(a, z) \in \mathbb{X}$ and $g \in \mathcal{P}_2(\mathbb{X})$ and hence $\mathcal{L}^*[\, g \,|c^\star] \in \mathcal{P}_2(\mathbb{X})$ too.

The novelty – compared to more "standard" HJBs – is tracking how the distribution g affects the value V. With Full Information Rational Expectations, agents with states $(a, z) \in \mathbb{X}$ forecast how the distribution changes, namely the distribution of all the *other agents* with states $(\tilde{a}, \tilde{z})$. For this reason, agents know that the operator $\mathcal{L}^*[\, g \,|c^\star]$ characterizes how the agents' decision $c^\star$ changes the distribution g over time.

Derivative with respect to distributions. To see how that change in g at point $\tilde{x}$ affects the value V, one needs to define the derivative of functions in the space of distributions $\mathcal{P}_2(\mathbb{X})$. We say that the function $V : \mathcal{P}_2(\mathbb{X}) \to \mathbb{R}$ admits a *flat-derivative* or *derivative in the* L^2*-sense*, if the operator $\frac{\delta V(g)[\tilde{x}]}{\delta g} : \mathcal{P}_2(\mathbb{X}) \times \mathbb{X} \to \mathbb{R}$ is identified as an element of L^2, and is thus a measure on elements $\tilde{x}$. Formally, the definition of this L^2 derivative is, for $g \in \mathcal{P}_2(\mathbb{X})$:

$$\frac{\delta V(g)}{\delta g}[\tilde{x}] := \lim_{\varepsilon \to 0+} \frac{V\left((1-\varepsilon)g+\varepsilon\delta_{\tilde{x}}\right) - V(g)}{\varepsilon} \qquad \text{or} \qquad V(\tilde{g}) = V(g) + \int_0^1 \int_{\mathbb{X}} \frac{\delta V\left((1-\varepsilon)g+\varepsilon\tilde{g}\right)}{\delta g}[\tilde{x}](\tilde{g}-g)(d\tilde{x})\, d\varepsilon$$

which is a function from $\mathcal{P}_2(\mathbb{X}) \times \mathbb{X}$ to $\mathbb{R}$.[3] Note that this *coincides to the usual Fréchet derivative* if we *restrict* to the measure that have density in L^2.

Intuitively, $\frac{\delta V(g)}{\delta g}[\tilde{x}]$ represents how the value changes when changing the mass at $\tilde{x}$, which is affected by the change in the distribution caused by the agents' drift-jump-diffusion $\mathcal{L}^\star[g|c^\star]$. If agents save more $\dot{a} = s(\cdot)$ and drift up, the mass on g will be shifted to a higher $\tilde{a}$, changing the value for all the other agents. In the following, we will also consider a different notion of derivative in $\mathcal{P}_2(\mathbb{X})$, namely the *Lions-derivative* or $L-$derivative $\frac{dV(g)}{dg}[\tilde{x}] : \mathcal{P}_2(\mathbb{X}) \times \mathbb{X} \to \mathbb{R}$, defined as:

$$\frac{dV(g)}{dg}[\tilde{x}] := \frac{d}{d\tilde{x}} \frac{\delta V(g)}{\delta g}[\tilde{x}] \; .$$

This definition will be useful in section 4, both analytically and in our approximation method.

The Master Equation such as the one in eq. (6) is called a "First Order Master Equation" because it contains only first-order derivatives with respect to the distribution.[4] Note that the solution of the Master equation V and the one of the "standard HJB" v in the Mean Field Games system are the same. More precisely, the solution of the Master equation $V(t, a, z, g_t)$ along the path of the distribution $\{g_t\}_t$ that solves the Kolmogorov Forward Equation eq. (5) corresponds to the solution of the Hamilton-Jacobi-Bellman eq. (1) equation $v(t, a, z) = V(t, a, z, g_t)$

[3] More precisely, $\mathcal{P}_2$ is not a vector space, and therefore, this Fréchet derivative is not a linear operator on $\mathcal{P}_2$. However, we can use the Hilbert structure in L^2 if we *restrict* g to distributions that admit densities in L^2. In that case, we can represent $\frac{\delta V(g)}{\delta g}$ as a linear functional on L^2, which is the Fréchet derivative: $\frac{\delta V(g)}{\delta g}(h) = \lim_{||h|| \to 0} \frac{V(g+h)-V(g)}{||h||}$, for perturbation $h \in L^2$. Moreover, $\frac{\delta V(g)}{\delta g}(h)$ is an element of L^2, and thus a function of $\tilde{x}$. Additional details: the perturbation measure h has to respect mass preservation, so $\int_{\mathbb{X}} h(dx) = 0$. Because of that, the derivative $\frac{\delta V(g)}{\delta g}$ is defined up to a constant, and we normalize $\int_{\mathbb{X}} \frac{\delta V(g)}{\delta g}[\tilde{x}] g(d\tilde{x}) = 0$

[4] Note that this should not be confused with the First-Order Approximation of the Master Equation (or FAME) coined by Bilal (2023)

3.2 Introducing aggregate risk

Now, let us consider aggregate risk, represented by changes in aggregate TFP Z. We can assume for exposition purposes that it follows an AR(1) – or Ornstein-Uhlenbeck – process

$$dZ_t = -\theta(Z_t - \bar{Z})dt + \widehat{\sigma} dB_t^0$$

where dB_t^0 is a Brownian motion, and starting from an initial distribution over $\mathbb{R}$. As a result, Z_t becomes an additional state variable that agents need to forecast to draw their expectations.

We can extend the Master equation with $V := V(t, a, z, g, Z)$ as follow:

$$-\partial_t V + \rho V = \overbrace{\max_c u(c) + \mathcal{L}[\,V\,|c](t,a,z)}^{\text{standard HJB continuation value}} \quad \overbrace{-\theta(Z-\bar{Z})V_Z + \tfrac{\widehat{\sigma}^2}{2} V_{ZZ}}^{\text{direct effect of risk of } Z \text{ on } v} \\ + \underbrace{\iint_{z,a} \frac{\delta V(t,a,z,g,Z)}{\delta g}[\tilde{a}, \tilde{z}]\ \mathcal{L}^*[\,g\,|\,c^\star](t,d\tilde{a},\tilde{z})}_{\text{change due to distribution dynamics}} \ . \tag{7}$$

for $(a, z) \in \mathbb{X}$ and $g \in \mathcal{P}_2(\mathbb{X})$ and $\mathcal{L}^*[\,g\,|c^\star] \in \mathcal{P}_2(\mathbb{X})$. As usual in HJB problems with drift and diffusion processes, additional terms representing the effect of aggregate risk Z on the value V. However, one can see that the structure of the coupling terms through the distribution $\delta V/\delta g$ did not change: there are no direct interactions between the individual states (a, z) and the aggregate risk on Z.

As a result, since the risk does not have *direct effects* on individual dynamics, the distribution g is not affected/deformed directly by shocks dB_t^0. Note that if it were the case, the Master equation could become *second order*, which is much more complicated and developed in the appendix XXX. In section XXX, we develop a framework for a portfolio problem in which the second-order terms are nontrivial and distort the value.

Decomposing the distributional effects. We can unpack the effects of distribution dynamics on the value. Recall that by definition, the adjoint operator gives $\langle \varphi, \mathcal{L}^*[g] \rangle = \langle \mathcal{L}\varphi, g \rangle, \forall \varphi \in \mathcal{C}_0(\mathbb{X})$ and $g \in \mathcal{P}_2(\mathbb{X})$. In that case, thanks to the definition eq. (3), and optimal saving and consumption eq. (4) and eq. (2), we obtain the value function for agent $V(a, z, g, Z)$ solving the master equation:

$$-\partial_t V + \rho V = \overbrace{\max_c u(c) + \mathcal{L}[\,V\,|c](t,a,z)}^{\text{HJB continuation value}} \quad \overbrace{-\theta(Z-\bar{Z})V_Z + \tfrac{\widehat{\sigma}^2}{2} V_{ZZ}}^{\text{direct effect of risk of } Z \text{ on } v} \\ + \underbrace{\iint_{z,a} \frac{dV(t,a,z,g,Z)}{dg}[\tilde{a}, \tilde{z}]\ s(\tilde{a},\tilde{z},g,Z,\tilde{c}^\star) g(t,d\tilde{a},\tilde{z})}_{\text{distribution dynamics through saving}} \\ + \underbrace{\iint_{z,a} \lambda_j \Big(\frac{\delta V(t,a,z,g,Z)}{\delta g}[\tilde{a}, \tilde{z}_{-j}] - \frac{\delta V(t,a,z,g,Z)}{\delta g}[\tilde{a}, \tilde{z}_j] \Big) g(t,d\tilde{a},\tilde{z}_j)}_{\text{distributional effects of income shocks}} \tag{8}$$

for $g \in \mathcal{P}_2(\mathbb{X})$ and where the derivative $\frac{dV(t,a,z,g,Z)}{dg}[\tilde{a},\tilde{z}] = \frac{d}{d\tilde{a}}\frac{dV(t,a,z,g,Z)}{dg}[\tilde{a},\tilde{z}]$ is the Lions-derivative of the value w.r.t to the distribution, and where the variable $\tilde{\cdot}$ are the values of other agents. In that case, the agents anticipate how the distribution moves in two ways: they expect the distribution to *shift* right when other agents' savings $s(\cdot)$ are higher. Indeed, the derivative $\frac{dV(t,a,z,g,Z)}{dg}[\tilde{a},\tilde{z}]$, represents how the value changes when the state of agent $\tilde{a}$ is "shifted" by $d\tilde{a}$. Note that this Lions derivative is not the same as the Fréchet derivative. More details are provided in Appendix XXX. Second, the effects of idiosyncratic risk on z of all the agents $(\tilde{a},\tilde{z})$ change the anticipation of agents (a,z) for how the distribution g of agents evolves, e.g. in particular how the value change – in a discrete way – when agents individual labor productivity switch from $\tilde{z}_j$ to $\tilde{z}_j$.

4 Aggregation and Projection

We know that in heterogeneous agent models, the requirement that agents have a well-defined view of the distribution's dynamics may be ambitious. For this reason, many economists have departed from the Full Information Rational Expectation framework. In this project, we follow the Krusell-Smith logic and assume that agents need only forecast asset prices. To make their decisions, they do not need the full shape of the distribution *per se*, but potentially only a few moments and statistics of the distribution.

As a result, we can make the conjecture that agents approximate and *project* the distribution with *sufficient statistics* and moments K^h:

$$K^h = \int_{\mathbb{X}} h(x)\, g(dx)$$

for function $h \in \mathcal{C}(\mathbb{X})$, for $\mathbb{X} = [\underline{a},\infty) \times \{z_1, z_2\}$. In the Krusell-Smith model, agent dynamics interact through the market clearing and the asset prices, i.e. the interest rate $r_t = \alpha K^{\alpha-1} - \delta$, and wages $w_t = (1-\alpha)K^{\alpha}$, which both depend on the mean K:

$$K = \iint_{a,z} a\, g(da, z)$$

and in this case $h(a) = a$. One could consider the projection on the first moment $K^1 = K$, or on a sequence of moments and statistics, e.g. $K^2 = \iint_{a,z} a^2\, g(da,z)$, $K^3 = \iint_{a,z} a^3\, g(da,z)$, etc. If one would like to follow the mass of agents at the borrowing constraint, we can also compute the function $K_{\underline{a}} = H(g) = g(\underline{a})$ to extract the size of the mass point of $g(\underline{a})$.

4.1 Projection and value

In the following, I develop a methodology to solve the Master Equation when using a *projection* of the distribution on such statistics K^h. Considering the projection $K^h = P(g)$ from

$g \in \mathcal{P}_2(\mathbb{X})$ to $K^h \in \mathbb{R}$, we approximate the value as:

$$V(a,z,g,Z) = \overline{V}(a,z,K^h,Z) + \varepsilon(a,z,g,Z) \tag{9}$$

where $\varepsilon(a,z,gZ)$ is the residual error from the Projection, since we dramatically reduce the complexity from an infinite-dimensional space $\mathcal{P}_2(\mathbb{X}) \ni g$ to $\mathbb{R} \ni K^h$. On the fiber sets $P^{-1}(k) = \{g \in \mathcal{P}_2(\mathbb{X}) \mid |P(g) = \int_{\mathbb{X}} h(x)\, g(dx) = k\}$, the value is constant $\overline{V}(\cdot,k)$ while $V(\cdot,g)$ might still vary.[5] In the next section, we mainly consider the projection on the first moment K.

$$V(a,z,g,Z) = \overline{V}(a,z,K,Z) + \varepsilon(a,z,g,Z)$$

In that case, we can provide an approximation for the residual error. Under the assumption that the value $V(\cdot,g)$ is Lipschitz continuous in the distribution g with respect to the Wasserstein distance $W_2(g,\mu) := \inf_{\pi\in\Pi(g,\mu)} \Big(\int_{\mathbb{X}} ||x-y||^2 \pi(dx,dy)\Big)^{\frac{1}{2}}$ where $\Pi(g,\mu)$ is a set of of transport plans from g to μ, we can provide a bound on the error:

$$\Big|V(a,z,g,Z) - \overline{V}(a,z,K,Z)\Big| \leq K\, W_2(g,\delta_K) \leq K\sqrt{\mathbb{V}\text{ar}_g(a)}$$

for K the Lipschitz constant of $V(\cdot,g)$. Therefore, we obtain that the approximation error is $\varepsilon(a,z,g,Z) = \mathcal{O}(\sqrt{\mathbb{V}\text{ar}_g(a)})$: The less dispersed the distribution g, the more accurate the approximation.

Naturally, we can use a longer sequence of moments, e.g. K_2, the variance over assets $K_2 = \mathbb{V}\text{ar}(a)$, or $K_{\bar{a}} = \mathbb{P}(a = \bar{a})$, the mass of agents at the borrowing constraint: $V(a,z,g,Z) \approx \overline{V}(a,z,K,K_2,\dots,K_{\bar{a}},Z)$. However, as the number of moments increases, the approximation becomes more precise, enabling agents to forecast the distribution's evolution with greater accuracy.

Such approximation can be justified for several reasons: First, using moment determinacy theorems – e.g. Hamburger and Stieltjes moment's problem, or Carleman's Theorem, one can claim that a probability distribution can be characterized by its moments $K^1, K^2, \dots$. Indeed, using moment generating functions (MGF) $M_X(t) = \mathbb{E}[e^{tX}]$ with $X \sim g$, if the MGF exists in the neighborhood of zero, then the distribution can be determined by its moments, given by $M_X^k(0) = \mathbb{E}[X^k] = \langle x^k, g\rangle$. Second, one can approximate probability densities with orthogonal polynomials – e.g. Legendre, Hermite, etc – of the form: $g(x) = \sum_n a_n P_k(x)$. The coefficients c_n in this approximation can be calibrated using the moments $K^n = \mathbb{E}[X^n]$ of the distribution.

[5] For example, the distribution could go from two extremes: (i) fully concentrated with all the mass on the mean $g(x) = \delta_K(x)$ where δ_k is the Dirac-Delta function, or (ii) fully dispersed, with maximum entropy: $g(x) = \frac{1}{K}e^{-\frac{x}{K}}$. Obviously, the value would differ between those two cases for arbitrary economic models.

4.2 Derivatives with respect to distributions

For those reasons, we consider such an approximation with projections. In our setting, such projections on the moments of the distribution have incredibly useful properties. Indeed, with this projection, the Lions' derivative of the value with respect to the distribution can be derived easily:

$$K^h = \langle h, g \rangle = \int_x h(x)\, g(dx)$$

$$\frac{\delta V(x,g)}{\delta g}[\tilde{x}] = \frac{d\overline{V}(x,K^h)}{dK^h}\, h(\tilde{x}) \qquad\qquad \frac{dV(x,g)}{dg}[\tilde{x}] \;\; = \frac{d\overline{V}(x,K^h)}{dK^h}\, h'(\tilde{x}) \tag{10}$$

where $\frac{\delta V(x,g)}{\delta g}[\tilde{x}]$ is identified to the Fréchet derivative, and $\frac{dV(x,g)}{dg}[\tilde{x}]$ is the Lions-derivative. With the latter derivative, we understand how a "shift" in other agents' states $d\tilde{x}$ affects the moment K^h and the value $\overline{V}(x, K^h)$.

Let us see some common examples that are useful in typical economic models:

- For the first moment, or aggregate capital, $K^h = K = \int_{\mathbb{X}} h(a)g(dx)$, with $h(x) = h(a) = a$, we obtain simply:

$$\frac{\delta V(a,z,g)}{\delta g}[\tilde{a},\tilde{z}] = \frac{d\overline{V}(a,z,K)}{dK}\,(\tilde{a} - K) \qquad\qquad \frac{dV(a,z,g)}{dg}[\tilde{a},\tilde{z}] = \frac{d\overline{V}(a,z,K)}{dK} \tag{11}$$

  Note that in the Master equation, the agents $x = (a, z)$ consider the integral over all the other agents $\tilde{x}$, with their movement represented by the operator $\mathcal{L}^\star[g|c^\star]$ which corresponds – for the first state a – to the saving function $s(x) := s(a,z,K,Z,c^\star) = r(K,Z)a + w(K,Z)z - c^\star$. As a result,

$$\int_{\tilde{x}\in\mathbb{X}} \frac{dV(a,z,g)}{dg}[\tilde{x}]\;s(\tilde{x})\;g(d\tilde{x}) = \frac{d\overline{V}(a,z,K)}{dK}\int_{\mathbb{X}} s(\tilde{x})g(d\tilde{x}) = \frac{d\overline{V}(a,z,K)}{dK}\frac{dK}{dt}$$

  In that case, one would only need to forecast the change in aggregate capital dK/dt to forecast future distribution changes and their impact on the value function.

- For second order moments, $K^h = K^2 = \langle x^2, g \rangle = \int_x (x-K)^2\, g(dx)$, we have similar results:

$$\frac{dV(a,z,g)}{dg}[\tilde{x}] = \frac{d\overline{V}(a,z,K^2)}{dK^2}\,2(\tilde{x} - K)$$

$$\int_{\mathbb{X}} \frac{dV(a,z,g)}{dg}[\tilde{x}]\,s(\tilde{a},\cdot)g(d\tilde{x}) = \frac{d\overline{V}(a,z,K^2)}{dK^2}\int_{\mathbb{X}} 2(\tilde{a}-K)\;s(\tilde{a},\cdot)dg(\tilde{x}) = \frac{d\overline{V}(a,z,K^2)}{dK^2} 2\,\mathbb{C}\mathrm{ov}\big(a, s(a,\cdot)\big) \tag{12}$$

  When households consider agents' dispersion K^2 in their forecast, they need to account for the covariance of savings with asset holdings. For example, if $\mathbb{C}\mathrm{ov}\big(a, s(\cdot)\big) > 0$, we have that rich agents get richer – and the poor get poorer – which increases inequality and dispersion K^2.

- For the mass at a specific point $\bar{a}$, we consider the function $h(a) = \delta_{\bar{a}}(a)$, and $K^h = K_{\underline{a}}$. One would also need some formalism using weak derivatives. In that case, the change in value write:

$$\frac{dV(a,z,g)}{dg}[\tilde{x}] = \frac{d\overline{V}(a,z,K_{\underline{a}})}{dK_{\underline{a}}} \frac{d\delta_{\underline{a}}(\tilde{x})}{d\tilde{x}}$$

$$\int_{\mathbb{X}} \frac{dV(a,z,g)}{dg}[\tilde{x}]s(\tilde{x})dg(\tilde{x}) = \frac{d\overline{V}(a,z,K_{\underline{a}})}{dK_{\underline{a}}} \int_{\mathbb{X}} \frac{d\delta_{\underline{a}}(\tilde{x})}{d\tilde{x}} s(a,\cdot)dg(\tilde{x}) = -\frac{d\overline{V}(a,z,K_{\underline{a}})}{dK_{\underline{a}}} \frac{ds(\underline{a},\cdot)}{d\tilde{x}} K_{\underline{a}}$$

  where in the last equation we used integration by parts. To forecast the change in mass at the borrowing constraint $\underline{a}$, one needs to know the net inflow of agents at that point, which is $\partial_a s(\underline{a},\cdot)$.

- Finally, we consider a simple variation, when the sufficient statistics considered is a function of a moment, i.e. $K^{\psi,h} = \psi(\langle h,g\rangle)$, which yields, in the generic case:

$$\frac{dV(x,g)}{dg}[\tilde{x}] = \frac{d\overline{V}(x,K^{\psi,h})}{dK^{\psi,h}} \psi'(\langle h,g\rangle)h'(\tilde{x}) \qquad \text{for general} \qquad K^{\psi,h} = \psi\Big(\int_{\mathbb{X}} h(x)dg(x)\Big) \tag{13}$$

  To provide intuitions, let us provide standard examples. One typical example used in many economic models is a CES index, which gives the derivative when $V(\cdot,P)$:

$$P := \Big(\int_{\mathbb{X}} p(x)^{\frac{\varepsilon-1}{\varepsilon}} dg(x)\Big)^{\frac{\varepsilon}{\varepsilon-1}} \qquad \Rightarrow \qquad \frac{dV(x,g)}{dg}[\tilde{x}] = \frac{d\overline{V}(x,P)}{dP}\Big(\frac{p(\tilde{x})}{P}\Big)^{-\varepsilon}$$

  The price index is sensitive to changes in the distribution at price $p(\tilde{x})$, and reweights the derivative of value by $(p(\tilde{x})/P)^{-\varepsilon}$, which is small for high prices and when goods are very substitutable $\varepsilon > 1$.

  Another example is when the value depends on asset prices directly, which themselves depend on a moment of the distribution:

$$r = \alpha Z\Big(\int_{\mathbb{X}} a\, g(da)\Big)^{\alpha-1} - \delta \qquad \Rightarrow \qquad \frac{dV(x,g)}{dg}[\tilde{x}] = -\frac{d\overline{V}(x,r)}{dr}(1-\alpha)\frac{(r+\delta)^{1+\frac{1}{1-\alpha}}}{(\alpha Z)^{\frac{1}{1-\alpha}}}$$

  An additional example when the value depends on an asset price that is defined implicitly by a market-clearing condition $\int_{\mathbb{X}} ag(da) = \bar{A}$. This case will be treated in Appendix XXX.

  In the generic case, and those two examples, we can use the chain rule and obtain the derivative of the value with respect to those sufficient statistics:

4.3 Projection in the Master equation

Equipped with the formalism developed in the previous section, we can now apply these techniques to the Master Equation we developed in eq. (7). We project the distribution on the first moment $K = \iint_{a,z} ag(da,z)$. We consider that agents have a restricted view of the aggregate dynamics, such that they

As a result, the value function would $V(a,z,g,Z) \approx \overline{V}(a,z,K,Z) =: \overline{V}$ with a projection, and solve the modified Master Equation:

Moreover, we also consider that there are no other time variations except for the stationary stochastic process Z, and we consider a stationary equilibrium such that $V(\cdot) = \lim_{t\to\infty} V(t,\cdot)$. As a result, the value function would $V(a,z,g,Z) \approx \overline{V}(a,z,K,Z) =: \overline{V}$ with a projection, and solve the modified Master Equation:

$$\underbrace{-\partial_t \overline{V}}_{=0} + \rho \overline{V} = \overbrace{\max_c u(c) + \mathcal{L}[\,\overline{V}\,|\,c\,]_{(a,z,K,Z)}}^{\text{standard HJB continuation value}} \; \overbrace{-\theta(Z-\bar{Z})\bar{V}_Z + \frac{\widehat{\sigma}^2}{2}\overline{V}_{ZZ}}^{\text{direct effect of } Z \text{ risk on } \overline{V}} \\ + \overline{V}_K \underbrace{\iint_{z,a} s(\tilde{a},\tilde{z},K,Z,\tilde{c}^\star) g(d\tilde{a},\tilde{z})}_{\text{distribution dynamics through } (\tilde{a},\tilde{z})\text{-savings}} \tag{14}$$

for $(a,z,K,Z) \in \mathbb{X} \times \mathbb{R}_+ \times \mathbb{R}_+$, with $K = \iint_{a,z} a g(da,z)$, and where $s(\tilde{a},\tilde{z},g,Z,\tilde{c}^\star)$ saving decision of agents $(\tilde{a},\tilde{z})$ depends on the optimal decision $c^\star$. We consider that, under full information rational expectation (FIRE), every agents (a,z) would anticipate that the decision function $c^\star(\cdot)$ of every other agent is the same as their own optimal decision $c^\star \in \operatorname{argmax}_c u(c) + \mathcal{L}[\,\overline{V}\,|\,c\,]$. As a result, the consumption choice accounts for aggregate uncertainty *and* for agents $(\tilde{a},\tilde{z})$ dynamics since it depends on the value function $\overline{V}$ that integrates those effects. We can write the consumption function as: $c^\star(a,z_j,K,Z) = c^\star(a,z_j,K,Z,\overline{V}_a(t,a,z_j,K,Z))$

$$c^\star(a,z_j,K,Z) = \begin{cases} u'^{-1}\big(\overline{V}_{a(a,z_j,K,Z)}\big) & a > \underline{a} \\ z_j \underbrace{(1-\alpha)ZK^\alpha}_{=w(K,Z)} + \underbrace{(\alpha Z K^{\alpha-1} - \delta)}_{=r(K,Z)} a & \text{if } a \leq \underline{a} \end{cases} \tag{15}$$

As a result, (a,z)-agents with FIRE internalize these decisions for others $(\tilde{a},\tilde{z})$ and thus for aggregate savings :

$$\begin{aligned} dK = \mathcal{S}(K,Z|g)dt &:= \iint_{z,a} s(\tilde{a},\tilde{z},K,Z,\tilde{c}^\star) g(d\tilde{a},\tilde{z}) \\ &= r(K,Z) \underbrace{\iint_{a,z} \tilde{a}\, g(d\tilde{a},\tilde{z})}_{=K} + w(K,Z) \underbrace{\iint_{a,z} z_j g(d\tilde{a},\tilde{z})}_{=\bar{L}} - \iint_{a,z} c^\star(\tilde{a},\tilde{z}_j,K,Z)\, g(d\tilde{a},\tilde{z}) \\ &= r(K,Z)K + w(K,Z)\bar{L} - \mathcal{C}(K,Z|g) = ZK^\alpha - \delta K - \mathcal{C}(K,Z|g) \end{aligned}$$

where $\mathcal{C}(K,Z|g)$ is the aggregate consumption, when integrating over all the agents $\tilde{a}.\tilde{z}$ according to the distribution g.

Rewriting and simplifying, the value $\overline{V} = \overline{V}(a, z_j, K, Z)$ solve the HJB-Master equation:

$$\begin{aligned}\rho \overline{V} = \max_c u(c) + [w(K,Z) z_j + r(K,Z) a - c] \overline{V}_a + \lambda(\overline{V}(a, z_{-j}, K, Z) - \overline{V}(a, z_j, K, Z)) \\ - \theta(Z - \bar{Z}) \overline{V}_Z + \tfrac{\widehat{\sigma}^2}{2} \overline{V}_{ZZ} + \underbrace{[Z K^\alpha - \delta K - \mathcal{C}(K, Z|g)]}_{=dK} \overline{V}_K\end{aligned} \tag{16}$$

for $(a, z, K, Z) \in \mathbb{X} \times \mathbb{R}_+ \times \mathbb{R}_+$. We can use a compact notation using the operator $\mathcal{A}_m[\overline{V}|c^\star]$, which depends implicitely on optimal decisions $c^\star$ for the evolution of states $x = (a, z, K, Z)$, with the Master HJB equation writing $\rho wb V = \max_c u(c) + \mathcal{A}_m[\overline{V}|c]$.

We see that this HJB-Master equation resembles the fusion of two HJBs from very standard economic models:

– First, the Aiyagari model seen above, with: $V = V(a, z_j)$, which is a heterogeneous agent model without aggregate risk.

$$\rho V = \max_c u(c) + [w z_j + r a - c] V_a + \lambda_j (V(a, z_{-j}, \cdot) - V(a, z_j, \cdot))$$

– Second, the Brock-Mirman model[6], with $V = V(K, Z)$, which is a representative agent model with aggregate risk.

$$\rho V = \max_C u(C) + [Z K^\alpha - \delta K - C] V_K - \theta(Z - \bar{Z}) V_Z + \tfrac{\widehat{\sigma}^2}{2} V_{ZZ}$$

Therefore, since the agents have four states (a, z, K, Z), two idiosyncratic states, and two aggregate states, they forecast their dynamics with an HJB equation that exactly merges the dynamics in those two models. Note also that the dynamic of the idiosyncratic state z, and the terms $\iint_{a,z} \lambda_j \big(\frac{\delta V}{\delta g}[\tilde{a}, \tilde{z}_{-j}] - \frac{\delta V}{\delta g}[\tilde{a}, \tilde{z}_j]\big) g(d\tilde{a}, \tilde{z})$ does not show up in the interaction term when projecting the distribution on the any moment of the asset dimension a. Indeed, the distribution of income z is stationary and does not depend on the dynamics of wealth. In our case, with the projection on the 1st moment, $\frac{\delta V}{\delta g}[\tilde{a}, \tilde{z}_j] = \overline{V}_K \tilde{a}$, which does not depend on $\tilde{z}$, and those terms drop out of the equation.[7]

To solve the complete allocation, we notice that an issue remains: the aggregate saving $\mathcal{S}(K, Z|g)$ consumption $\mathcal{C}(K, Z|g)$ functions are still an aggregation of individual policies $c^\star(a, z, K, Z)$ and therefore still depend on the distribution of agents $g(a, z)$. The literature has faced such an issue by *parametrizing* the aggregate decision function $\mathcal{S}(K, Z|\theta)$. In the original paper, Krusell-Smith have assumed that agents would forecast a linear rule:

$$dK = \beta^Z \log(K) dt$$

[6] Alternatively, we can say that it looks like the HJB from the RBC model with exogenous labor supply

[7] Note that if the dynamics of the distribution of income would depend on consumption, saving, and wealth – for example, in a model with search – then that term would remain and there would be another integral in the equation.

where β^Z depends on Z. Other methods in the literature have approximated this capital dynamics using deep-learning with *parametric representations* of the aggregate variables, Fernández-Villaverde et al. (2023) or reinforcement learning as in Yang et al. (2026):

$$dK = \mathcal{S}(K, Z|\theta)dt$$

where $\mathcal{S}(\cdot|\theta)$ is a non-linear function – e.g. a neural network / deep-learning representation – with parameters θ or β to be guessed and updated after model-based Monte Carlo simulations of the decisions of agents (a, z).

The strategy pursued here will be to try to characterize the low of motion $dK = \mathcal{S}(K, Z|g)$ for a well-chosen distribution g. In the next section, we explain how to derive the distribution coming from the dynamics of the system of state variables (a, z).

4.4 Distribution, aggregation, and system dynamics

To find the distribution over states $g(a, z)$, we solve the dynamical system in two stages: (i) first derive the whole distribution over all states $\widetilde{g}(a, z, K, Z)$, (ii) second, from that distribution $\widetilde{g}$, we compute the marginal distribution and derive the appropriate $g(a, z)$ conditional on the two aggregate states (K, Z).

Stage 1: from the Master equation to the distribution $\widetilde{g}$ of the global system

First, from the Master equation, and the value function $\overline{V} = \overline{V}(a, z, K, Z)$, we obtain the individual decisions in consumption and saving[8] as in :

$$c^\star(a, z_j, K, Z) = \begin{cases} u'^{-1}\big(\overline{V}_{a(a,z_j,K,Z)}\big) & a > \underline{a} \\ z_j \underbrace{(1-\alpha)ZK^\alpha}_{=w(K,Z)} + \underbrace{(\alpha Z K^{\alpha-1} - \delta)}_{=r(K,Z)} a & \text{if } a \leq \underline{a} \end{cases}$$

From that, we naturally obtain the dynamical system for the complete system of $x = (a, z, K, Z) \in \widetilde{\mathbb{X}} := [\underline{a}, \infty) \times \{z_1, \ldots, z_N\} \times \mathbb{R}_+ \times [\underline{Z}, \overline{Z}]$.

$$\begin{cases} da &= s(x) = \big[z\, \overbrace{(1-\alpha)Z\, K^\alpha}^{=w(K,Z)} + \overbrace{(\alpha Z\, K^{\alpha-1} - \delta)}^{=r(K,Z)}\, a - c^\star(a, z, K, Z)\big]dt \\ dz_j &= dJ_{jt} \qquad \text{Markov, w/ intensity} \qquad \{\lambda_j\}_j \\ dK &= \big(ZK^\alpha - \delta K - \mathcal{C}(K, Z|g)\big)dt \\ dZ &= -\theta(Z - \bar{Z})dt + \widehat{\sigma} dB_t^0 \end{cases}$$

for $x \in \widetilde{\mathbb{X}}$, and for a "guess" for the distribution $g(a,z)$ and aggregate dynamics $\mathcal{C}(K, Z|g) = \iint_{a,z} c^\star(a,z,K,Z) dg(a,z)$.

[8] Similarly with saving: $s(x) = s(a, z_j, K, Z) = z_j w(K,Z) + r(K,Z)a - c^\star(a,z_j,K,Z)$.

Given that the previous system gives us a complete characterization for the states evolution and hence the distribution $\widetilde{g}(x) = \widetilde{g}(a, z, K, Z)$, which is solution of the following the Kolmogorov forward equation for the system $x = (a, z_j, K, Z) \in \widetilde{\mathbb{X}}$.

$$\begin{aligned} 0 &= -\partial_a\big[s(x)\,\widetilde{g}(x)\big] + \lambda_{-j}\,\widetilde{g}(x_{-j}) - \lambda_j \widetilde{g}(x) \\ &\qquad - \partial_K[\mathcal{S}(K, Z|g)\;\widetilde{g}(x)] + \partial_Z[\theta(Z-\bar{Z})\widetilde{g}(x)] + \widehat{\sigma}\partial^2_{ZZ}\widetilde{g}(x) = 0 \qquad \text{on} \quad \widetilde{\mathbb{X}} \\ 0 &= \mathcal{A}^*_m[\widetilde{g}](x) \qquad \text{on} \;\; \widetilde{\mathbb{X}} \end{aligned}$$

where $\mathcal{A}^*_m[\widetilde{g}]$ is the adjoint operator of $\mathcal{A}_m[\overline{V}](x) := \mathcal{A}_m[\overline{V}|c^\star(x)](x)$ with $\widetilde{g} \in \mathcal{P}_2(\widetilde{\mathbb{X}})$ distribution on the enlarged space $x \in \widetilde{\mathbb{X}}$, and the distribution over individual states $g(a, z)$ that is an input for the aggregate dynamics $dK = \mathcal{S}(K, Z|g)$.

As is usual in Heterogeneous Agent models in continuous time, once we can solve the Hamilton Jacobi Bellman – here the master equation eq. (16) in space $\widetilde{\mathbb{X}}$ – we obtain the operator $\mathcal{A}_m$, its adjoint $\mathcal{A}^*_m$ and the solution of the Kolmogorov Forward equation $\widetilde{g}$ for "free". In the next section, we will implement a numerical method to implement that in practice.

Stage 2: From the global distribution $\widetilde{g}$ to the distribution of agents g.

The connection between the distribution $\widetilde{g}$ over states $x = (a, z, K, Z)$ and the distribution g over agents (a, z) holds using the Radon-Nikodym theorem – or Bayes rules when both measure g and $\widetilde{g}$ have continuous densities – when conditioning on specific values of the aggregate states, capital K and TFP Z.[9]

$$g(da, z_j)\big|_{K,Z} = g\big(da, z_j\big|K, Z\big) = \frac{\widetilde{g}(da, z_j, dK, dZ)}{\displaystyle\int_{\mathbb{X}} \widetilde{g}(da, z_j, dK, dZ)} \tag{17}$$

where $g\big(da, z_j\big|K, Z\big)$ is the Radon-Nikodym derivative with respect to the distribution of aggregate states $g(d\widetilde{K}, d\widetilde{Z}) = \int_{\mathbb{X}} \widetilde{g}(da, z_j, d\widetilde{K}, d\widetilde{Z})$ which is the aggregation over all agents $(a, z) \in \mathbb{X}$ for each states-of-the-world (K, Z). Additionally, we also obtain the following aggregation:

$$\int_{\mathbb{X}} a\; g(da, z_j) = \int_{\widetilde{\mathbb{X}}} a\; g\big(da, z_j\big|K, Z\big) = K \qquad\qquad \forall K, Z \in \widetilde{\mathbb{X}}$$

where the aggregate stock of asset – when summing all agents over all the states $\widetilde{\mathbb{X}}$ – corresponds to that particular aggregate state K.

[9]We define $g(da, z_j)\big|_{K,Z} = g\big(da, z_j\big|K, Z\big) = P\big((a, z) \in d\widetilde{a} \times \{z_j\}\big|(K, Z){=}(\widetilde{K}, \widetilde{Z})\big)$ is the conditional probability measure of (a, z) conditionally on the aggregate states (K, Z).

Using this $g = g|_{K,Z} = g(\cdot|K,Z)$, we can thus write the aggregate consumption function:

$$\mathcal{C}(K, Z|g) = \mathcal{C}\Big(K, Z\Big|\ g|_{K,Z}\Big) = \int_{\mathbb{X}} c^\star(a, z, K, Z)\ \ g(da, z_j|\widetilde{K}, \widetilde{Z}) = \frac{\int_{\mathbb{X}} c^\star(a,z,K,Z)\,\widetilde{g}(da,z_j,dK,dZ)}{\int_{\mathbb{X}} \widetilde{g}(da,z_j,dK,dZ)}$$

and similarly for aggregate saving as $\mathcal{S}(K, Z|g) = \mathcal{S}\big(K, Z\big|\ g|_{K,Z}\big)$. This last step corresponds to the consistency condition in rational-expectation equilibria with full information, where agents anticipate that the aggregate dynamics $\mathcal{C}(K, Z|g)$ evolve following the aggregation of agents' decision accross the distribution.

4.5 Summary, general method, and discussion

In the previous section, we aggregated the economy using a representation with a single aggregate moment, the mean of the distribution K. Our methods help characterize what are the two effects of this projection of the equilibrium. (i) First, the agents have a limited understanding of how aggregate dynamics evolve; they can only observe changes in the mean K rather than the full distribution g. (ii) Second, the aggregate dynamics themselves $dK = \mathcal{C}(K, Z|g)dt$ depend on an aggregation – which yields $g(da,z_j|K,Z)$ in our methods. This also averages over all the dynamics that are more "granular" than the first moment. Indeed, when taking the Bayes rule, we are averaging $g(da, z)$ over higher-order moments of the dynamics. In principle, if agents knew the movements in the variance or skewness of consumption/savings, this could help predict changes in dK more accurately. In models with projection on the mean, this is limited.

Indeed, one needs to remember that the Krusell-Smith model, like other HA models with aggregate risk, is *a priori* not Markovian in the states (a, z, K, Z), while the system is Markovian in (a, z, g, Z) when the state is enlarged to include the infinite-dimensional distribution dynamics. As a result, in our method, the "approximation" of the dynamics of its first moment $dK = \mathcal{C}(K, Z|g)dt$ is missing these more granular movements of the distribution – and potentially path-dependence in K. This is not accounted for in our distribution $\widetilde{g}$, which we use to obtain the agents' distribution g that households take into account when aggregating the other players.

Despite these limitations, this method is a first step toward making the Master equation a computable general methodology to handle aggregate risk. Before generalizing to higher-order moments, let us summarize the algorithm we have been using so far.

General numerical methodology

Here we summarize briefly the approach we considered in the previous section:

0. Setting up the general Master equation: one equation for $v(a, z, g, Z)$

1. Projection on $K = \langle a, g\rangle$, and Master Equation with "projection": $V = \overline{V}(a, z, K, Z)$

- Start from a guess for $dK^{(n)} = \mathcal{C}(K, Z|g)dt$
- Solve the Master HJB eq. (16), for $\overline{V}(a,z,K,Z)$ using standard finite difference methods
- Obtain the individual decisions $c^\star(a,z,K,Z)$ and operator $\mathcal{A}[\overline{V}|c^\star](x)$ for $x = (a, z, K, Z)$

2. Kolmogorov forward for the global dynamical system $x = (a, z, K, Z)$

- Solve for distribution $\widetilde{g}$ over all states (a, z, K, Z) for "free" with $\mathcal{A}^*[\widetilde{g}|c^\star]$
- Using Radon-Nykodym / Bayes rules, obtain $g(a, z\big|K, Z)$
- Update $\mathcal{C}(K, Z|g)$ using $g = g(a, z\big|K, Z)$

3. Obtain aggregate – potentially non-linear – dynamics:

$$dK^{(n+1)} = ZK^\alpha - \delta K - \mathcal{C}(K, Z|g)$$

- Use this dynamics dK in steps 1 and 2
- Repeat until convergence of the dynamics $||dK^{(n+1)} - dK^{(n)}|| < \varepsilon$.

This methodology can be extended by including additional moments to characterize the equilibrium, and that is what we turn to next.

5 Master equation and Projection on higher moments

Applying the above methodology, we can also project the value function $V(\cdot, g)$ on the larger set of moments. Here we will consider the first and second moments: K, $K_2 = \mathbb{V}\text{ar}(a)$, $L_2 = \mathbb{V}\text{ar}(z)$ and $KL = \mathbb{C}\text{ov}(a, z)$, with $\overline{V} = \overline{V}(a, z, K, K_2, KL, Z) = \overline{V}(\widetilde{x})$. We do not consider adding the states for mean labor productivity $\bar{L} = \mathbb{E}[z]$ and variance $L_2 = \mathbb{V}\text{ar}(z)$, since those moments are constant in a stationary equilibrium and would not vary with saving, nor any other states.

We use the calculus developed in section 4.2 and consider the first and second order moments:

$$K = \int_{\mathbb{X}} a g(dx) \qquad K_2 = \int_{\mathbb{X}} (a-K)^2 g(dx) \qquad KL = \int_{\mathbb{X}} (a-K)(z-\bar{L}) g(dx)$$

and we can define the functions h to be: $h(x) = a$ for K, or $h(x) = (a - K)^2$ for K_2, and $h(x) = (a - K)(z - L)$ for the covariance KL. Following this, the master equation rewrites with the new states $\widetilde{x} = (a, z_j, K, K_2, KL, Z)$ and $\overline{V} = \overline{V}(\widetilde{x})$:

$$\begin{cases} \rho\overline{V} = \max_c u(c) + s(a,z,K,Z,c)\,\overline{V}_a + \lambda_j(\overline{V}(a,z_{-j},\cdot) - \overline{V}(a,z_j,\cdot)) \\ \qquad -\theta(Z-\bar{Z})\overline{V}_Z + \frac{\bar{\sigma}^2}{2}\overline{V}_{ZZ} + \overline{V}_K\big[ZK^\alpha - \delta K - \mathcal{C}(K,K_2,KL,Z|g)\big] \\ \qquad +\overline{V}_{K_2}\int_{\mathbb{X}}(\tilde{a}-K)s(\tilde{a},\tilde{z},K,Z,c^\star)\,g(d\tilde{a},\tilde{z}) \\ \qquad +\overline{V}_{KL}\int_{\mathbb{X}}\lambda_j(\tilde{z}_{-j}-\tilde{z}_j)(\tilde{a}-K)\,g(d\tilde{a},\tilde{z}) \qquad \text{on } \widetilde{\mathbb{X}}\times\mathbb{R}^3_+ \end{cases} \tag{18}$$

The second order terms in the Master equation introduce three additional effects for the agents anticipation: (i) First, the correlation between the saving rate $s(\cdot)$ and the asset owned by agents a plays a role on how they foresee the change in inequalities $K_2 = \mathbb{E}[(a-K)^2] = \mathbb{V}\text{ar}(a)$. As we saw in eq. (12), we have:

$$\int_{\mathbb{X}} (\tilde{a}-K)s(\tilde{a},\tilde{z},K,Z,c^\star))\, g(d\tilde{a},\tilde{z}) = \mathbb{C}\text{ov}^g\big(a, s(a,\cdot)\big)$$

When households consider agents' dispersion K^2 in their forecasts, they account for the covariance between savings and asset holdings. For example, if $\mathbb{C}\text{ov}\big(a,s(\cdot)\big) > 0$, rich agents get richer – and poor ones get poorer – and this increases inequality and asset dispersion K_2. Note that this covariance still depends on g, the distribution of agents over (a,z).

(ii) Second, the correlation between labor productivity changes $(z_{-j} - z_j)$ and assets a plays a role in how agents anticipate the evolution of the covariance KL. In the case of labor productivity with two states, we obtain directly that:

$$\frac{dKL}{dt} = \int_{\mathbb{X}} \lambda_j(\tilde{z}_{-j} - \tilde{z}_j)(\tilde{a}-K)\, g(d\tilde{a},\tilde{z}) = -(\lambda_1+\lambda_2)\ \mathbb{C}\text{ov}^g\big(a,z\big) = -(\lambda_1+\lambda_2)\ KL$$

As a result, when the covariance KL is positive – i.e. agents with high income are also the ones with the highest assets – agents anticipate that income shocks, that arise with intensity λ_1 and λ_2, will change their state from z_j to z_{-j}, making the new "low income" agent inherit the high asset $\mathbb{E}[a|z_j]$, and conversely. As a result, this covariance is "self-regulating", leading to a stationary distribution over KL to be degenerate at 0.

(iii) Third, including more information in the agents' set changes their consumption and saving decisions $c^\star$ and $s(\cdot,c^\star)$, which in turn affects the dynamics for aggregate capital $\mathcal{C}(K,K_2,KL,Z|g)$.

Indeed, in the Master equation eq. (18), the optimal consumption depends on all the states $\tilde{x}$ and $c^\star(\tilde{x}) = c^\star{}_{(a,z_j,K,K_2,KL,Z)} = c^\star(a,z_j,K,Z,\overline{V}_{a(t,a,z_j,K,Z)})$

$$c^\star{}_{(a,z_j,K,K_2,KL,Z)} = \begin{cases} u'^{-1}\big(\bar{V}_{a(a,z_j,K,K_2,KL,Z)}\big) & a > \underline{a} \\ z_j \underbrace{(1-\alpha)ZK^\alpha}_{=w(K,Z)} + (\underbrace{\alpha Z K^{\alpha-1} - \delta}_{=r(K,Z)})\, a & \text{if } a \leq \underline{a} \end{cases} \tag{19}$$

As a result, aggregate capital dynamics also depend on those moments through the aggregation:

$$\mathcal{C}(K,K_2,Z|g) = \int_{\mathbb{X}} c^\star{}_{(\tilde{a},\tilde{z}_j,K,K_2,KL,Z)} g(d\tilde{a},\tilde{z})$$
$$dK = \big(ZK^\alpha \bar{L}^{1-\alpha} - \delta K - \mathcal{C}(K,K_2,KL,Z|g)\big)dt$$

Hence, knowing additional moments about the distribution helps predict agents' individual decisions $c^\star(\tilde{x})$ and the dynamics of that first moment K more accurately.

General numerical methodology: As before, we follow the same the overall approach to solve the equilibrium with aggregate risk. (1.) First, we solve the Master equation projected on those moments in eq. (18), starting from a guess for the aggregate dynamics dK, dK_2 and dKL. We obtain an operator $\mathcal{A}_m[\overline{V}|c^\star]\tilde{x}$ informing the dynamics of states $\widetilde{x} = (a, z, K, K_2, KL, Z)$. (2.) Thanks to that operator, we solve the Kolmogorov Forward equation – its dual – to obtain the stationary distribution $\widetilde{g}$ over $\widetilde{x}$. Thanks to that distribution $\widetilde{g}(\tilde{x})$ we can derive – using Bayes rules / Radon-Nykodym as in eq. (17) – and the distribution of agents conditional on the aggregate states $g = g(a, z|K, K_2, KL, Z)$. (3.) With such g we can compute the aggregate dynamics for dK, dK_2 and dKL as shown before:

$$\begin{aligned} dK &= \mathcal{S}(K, K_2, KL|g) = \mathbb{E}^g\big[s(\cdot, c^\star(\cdot))\big]\, dt \\ dK_2 &= \mathbb{C}\text{ov}^g\big(a, s(a, \cdot, c^\star(\cdot))\big) \\ dKL &= -(\lambda_1{+}\lambda_2)\; \mathbb{C}\text{ov}^g\big(a, z\big) \end{aligned}$$

We can use that again in step (1.) and (2.) and iterate until convergence.

6 Numerical implementation – Krusell-Smith Model

In the next numerical experiment, we consider the following risk structure, with incomplete market. The Idiosyncratic risk is a two-state Markov process for labor-income shocks, $z \in \{z_1, z_2\}$. The Aggregate risk is a three-state Markov process for TFP $Z \in \{Z_1, Z_2, Z_3\}$, centered around its mean $\mathbb{E}(Z) = Z_2 = 1$ and $\{Z_1, Z_3\}$ such that $\sigma(Z) = 12\%$.

I compare two benchmark models with the Krusell-Smith framework: First, the Aiyagari model, a benchmark heterogeneous agents model without aggregate risk. It yields the value and distribution (v, g) for individual heterogeneity on (a, z), for constant TFP $Z = \bar{Z}$ and capital $K = \bar{K}$.

Second, the Brock-Mirman model for a benchmark representative agent model with aggregate risk (similar to an RBC model with exogenous labor supply). This implies a stationary value and distribution (v, g) over aggregate capital and TFP (K, Z).

Finally, the Krusell-Smith model will have both types of idiosyncratic and aggregate risk. As a result of our projection method, the value and distribution (v, g) will be over the four states: (a, z, K, Z). As explained in section 4, we will find the agent distribution by iterating over $dK = \mathcal{S}(K, Z|g)dt = [ZK^\alpha - \delta K - \mathcal{C}(K, Z|g)]dt$.

6.1 Recap – Aiyagari model and Brock-Mirman model

In fig. 1, we plot the value function of the Aiyagari model, which differs markedly between $v(a, z_1)$ and $v(a, z_2)$, especially for low assets. When assets run down, agents cannot self-insure against income shocks, and they are forced to reduce consumption in the low state $\{z_1\}$. The stationary distribution is plotted in fig. 2, where we observe a mass point at $a = \underline{a}$ corresponding to roughly 1% of the agents at the borrowing constraint.

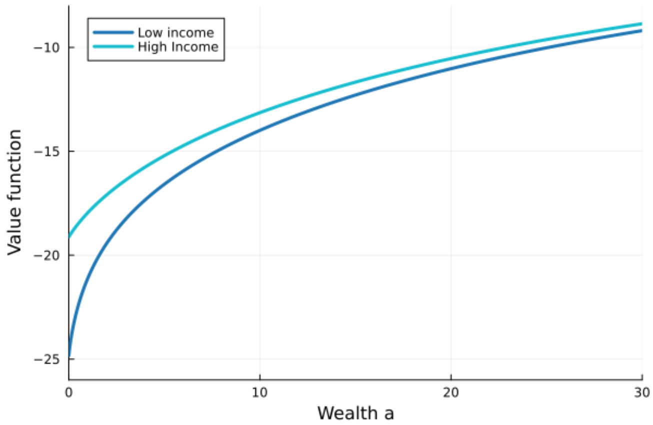

Figure 1: Value function $v(a, z)$

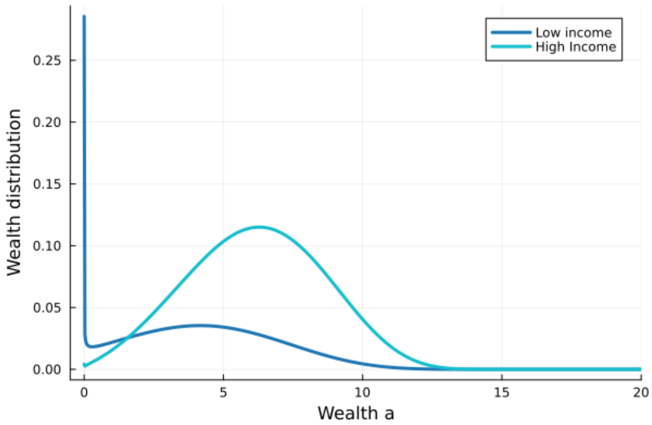

Figure 2: Distribution $g(a, z)$

In fig. 3, we plot the value function $V(K, Z)$ of the representative-agent Brock-Mirman model, which embeds production in capital K and productivity Z. Given that transitions between the three states are relatively frequent, the values only differ by $2-3\%$ between the different Z. The stationary distribution is also centered around the steady-state $K = 4.5$, and the capital is relatively higher for higher TFP Z.

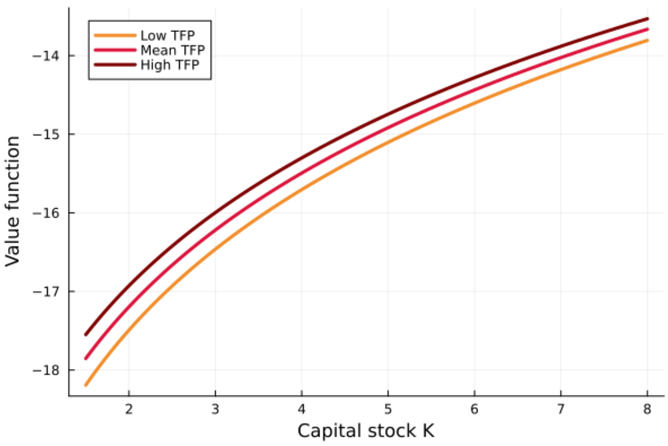

Figure 3: Value function $v(K, Z)$

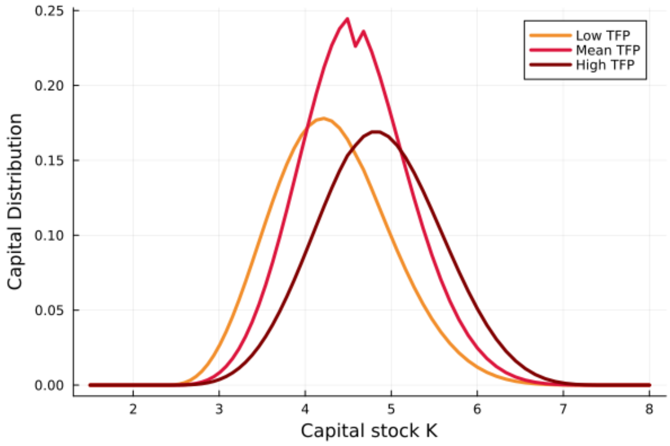

Figure 4: Distribution $g(K, Z)$

6.2 Krusell-Smith model with projection

First, we compare the individual decisions in the Krusell-Smith model in comparison to the Aiyagari model when the capital and productivity levels are the same $(K, Z) = (\bar{K}, \bar{Z})$

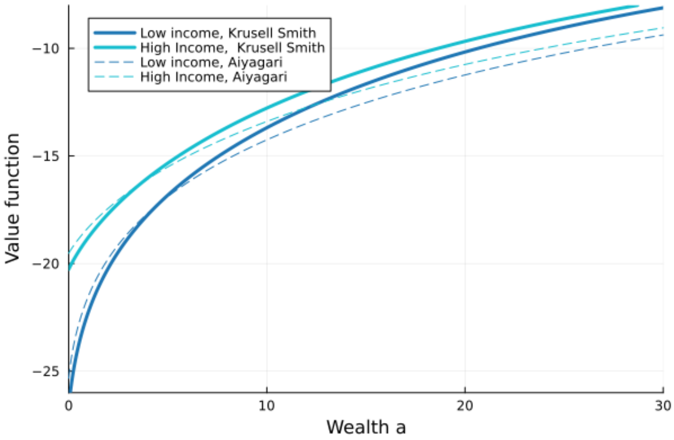

Figure 5: Value function $v(a, z, \bar{K}, \bar{Z})$

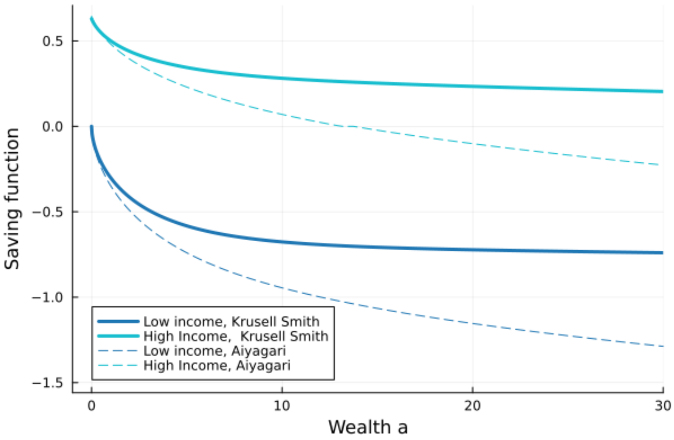

Figure 6: Saving $s(a, z, \bar{K}, \bar{Z}) = wz + ra - c^{\star}$

Second, we compare the aggregate dynamics of the HA model with those of the representative-agent RBC model.

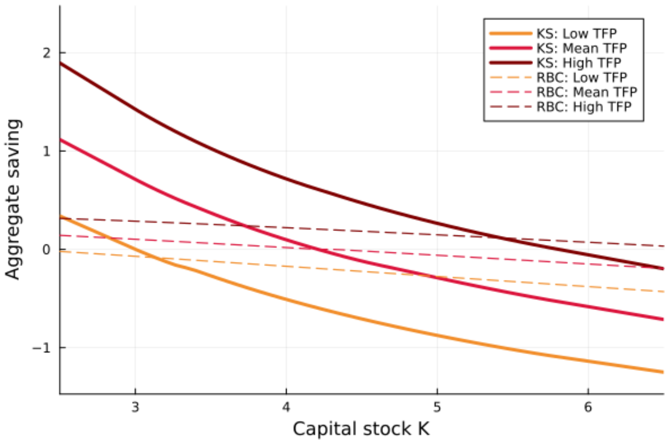

Figure 7: Aggregate dynamics $dK = ZK^{\alpha} - \delta K - \mathcal{C}(K, Z|g)$

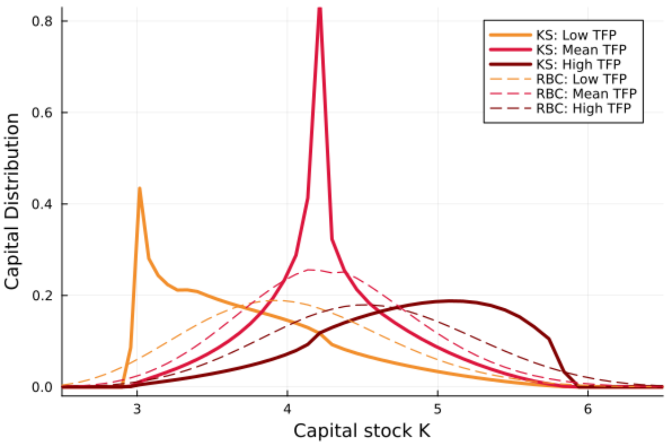

Figure 8: Distribution $\widetilde{g}(K, Z)$

We see that the effect of aggregate uncertainty with heterogeneity is threefold: (1) First, we see higher individual precautionary saving motives: Aggregate uncertainty affects the poor (reliant on labor income) more than the rich, who can hedge with $r = MPK$, which increases when capital K and income w drop. This implies much greater savings when rich and a much flatter savings function than in Aiyagari.

Second, aggregate dynamics are more "reactive" to the states – capital and productivity. Individual heterogeneity implies a steeper aggregate saving function, since rich households save a lot more when K is lower, boosting investment.

Third, it changes the distribution of economic activity over time: Output is fluctuating more in Krusell-Smith: 15.5% volatility compared to 14% for RBC. However, business cycles are less skewed – smaller right tail, more symmetric – and have less kurtosis – thinner tails of capital/output.

7 When uncertainty and inequality matter for aggregate dynamics – Applications to Asset Pricing and Portfolio choice

In this section, we consider two simple applications: (i) the first is a generalization of the Merton portfolio choice model in Partial Equilibrium, with households being affected by idiosyncratic income shocks as in the Aiyagari-Huggett model. Despite the exogeneity of asset prices, it is convenient, as it is well known and easy to compare with its representative-agent counterpart. (ii) The second is a replication of the Fernández-Villaverde et al. (2023) using our projection method. In both cases, we see how the projection method can provide analytical insight as much as a general method for simulating those economies.

7.1 Merton portfolio choice and inequality

As a proof of concept for the method developed above, we now consider a simple model as in the Merton portfolio problem. Heterogenous households can choose to invest in two assets: stocks k and bonds b.

First, we consider a firm producing a good with a Cobb-Douglas production function $Y_t = K_t^{\alpha}\bar{L}^{1-\alpha}$ where labor supply is fixed $\bar{L} = 1$ as before. With a competitive market, the rental rate of capital writes: $r_t^k = \alpha Z_t K_t^{\alpha-1}$. We also consider that the instantaneous return rate on capital is also subject to a direct shock dB_0 with volatility $\overline{\sigma}$:

$$dR_t^k = (r_t^k - \delta)dt + \overline{\sigma}\, dB_0$$

This stock is the first asset that households can hold. Moreover, the household can also borrow in a risk-free bond with a known (and constant) interest rate r. Note that this interest rate is determined exogenously as in a small-open-economy – for simplicity. One extension (WIP) is to have this interest determined in equilibrium with a zero-net-supply.

In addition, the households receive labor income w_t from a representative firm. This income is subject to idiosyncratic shocks z following a two-state Poisson process, as well as an aggregate TFP shock that follows an Ornstein-Uhlenbeck process:

$$dZ_t = -\theta(Z_t - \bar{Z})dt + \widehat{\sigma}\big(\rho\, dB_t^0 + (1-\rho)dB_t\big)$$

where ρ represents the degree of correlation between the aggregate income shock Z and the return on equity S. As a result, if $\rho = 1$, the aggregate risk dB_t^0 has both an impact on TFP/income and a *direct impact* on wealth.

The household can save in both assets, with their wealth combining bond and stocks $a = b+k$. Households can choose the stock share of their portfolio $\alpha = k/a$, depending on their income, wealth, and aggregate states. As before, markets are incomplete, and households face a borrowing constraint $a \geq \underline{a}$. Their wealth dynamics can be expressed as:

$$da_t = (ra_t + z_t w_t - c_t)dt + \alpha_t \, a_t \, (r_t^k - \delta - r)dt + \alpha_t \, a_t \, \overline{\sigma} \, dB_t^0 \tag{20}$$

The household is there to maximize utility subject to such wealth, income, and asset prices dynamics.

$$V(0, a, z, g, Z) = \max_{\{b_t, k_t, c_t\}_t} \mathbb{E}_0 \int_0^\infty e^{-\rho t} u(c_t) dt$$

where the households' states are their wealth a, their income z, and the aggregate TFP Z, as well as the distribution of agents g. Their controls are the consumption c_t, their capital holding $k_t = \alpha_t a_t$ and $b_t = (1 - \alpha_t)a_t$. Note that it would be easy to extend this framework with Epstein-Zin preferences, such that $u(c)$ can be replaced by $f(c_t, V_t)$. The reason why households need to forecast the movement of the distribution g is again that, in equilibrium, the aggregate supply of capital determines its return:

$$K_t = \int_{a,z} \alpha(\cdot) \, a \; g(t, da, z)$$

This, added with the goods market clearing $Y_t = C_t + I_t$, and the bonds market clearing $B_t = \int_{a,z} (1 - \alpha) a g(t, da, z) = 0$, implies that aggregate capital evolves as:

$$dK_t = (Y_t - \delta K_t - C_t)dt + \overline{\sigma} \, dB_t^0 = (I_t - \delta K_t)dt + \overline{\sigma} \, dB_t^0$$

Master equation Following the same approach as in the previous section, we can set up the Master equation. We consider again a stationary equilibrium where the individual value and distribution are constant over time $V(\cdot) = \lim_{t \to \infty} V(t, \cdot)$. Because aggregate risk directly affects households' budgets, the dynamics of wealth and its distribution are themselves stochastic, which implies that the Master equation now becomes ***second-order***, adding several layers of interaction between risk and wealth inequality.

To alleviate notation, we rewrite the asset dynamics with $da = \mu(\cdot)dt + \alpha a \overline{\sigma} dB_t^0$ with $\mu(\cdot) = \mu(a, z, g, Z, c^\star, \alpha^\star)$. We also define the infinitesimal operator for the individual wealth and income coming from agents' decision in consumption $c^\star = c^\star(a, z_j, g, Z, v)$, and portfolio choice

$\alpha^\star = \alpha^\star(a,z_j,g,Z,v)$, as well as shocks z_j:

$$\mathcal{L}[V\,|c,\alpha](a,z_j,\cdot) = V_a(a,z_j,\cdot)\underbrace{[z_j w + r\,a - c + \alpha a(r^k-\delta-r)]}_{\text{deterministic change in wealth}} + \underbrace{\lambda_j(V(a,z_{-j},\cdot) - V(a,z_j,\cdot))}_{\text{change in labor income}}$$
$$+\underbrace{\tfrac{\overline{\sigma}^2}{2}\,\alpha^2\,a^2\,V_{aa}(a,z_j,\cdot)}_{\text{diffusion w/ agg. risk}} + \underbrace{\rho\,\overline{\sigma}\,\widehat{\sigma}\,\alpha\,a\,V_{aZ}(a,z_j,\cdot)}_{\text{covariance btw/ inv. \& income risk}} \tag{21}$$

where $w = w(Z,g)$, $r^k{=}r^k(Z,g)$, and $r{=}r(Z,g)$ depends on the agent's distribution g in equilibrium. In addition to familiar terms seen before, aggregate risk changes the drift by adding an additional excess return of capital $\alpha(r^k-\delta-r)$, adds a diffusion term in the wealth dynamics, scaling with the aggregate risk on return $\overline{\sigma}$, and also adds a covariance cross term between aggregate investment risk $\overline{\sigma}$ and income/TFP risk $\widehat{\sigma}$.

As a result, with the value $V = V(a,z,g,Z)$, the Master equation writes:

$$\rho V = \overbrace{\max_{c,\alpha}\, u(c) + \mathcal{L}[V|c,\alpha](a,z,g,Z)}^{\text{standard HJB continuation value}} \quad \overbrace{-\theta(Z-\bar{Z})V_Z + \tfrac{\widehat{\sigma}^2}{2}V_{ZZ}}^{\text{direct effect of risk of } Z \text{ on } V} \quad + \overbrace{\iint_{z,a} \frac{dV(a,z,g,Z)}{dg}[\tilde{a},\tilde{z}]\;\mu(\cdot,c^\star,\alpha^\star)\,g(d\tilde{a},\tilde{z})}^{\mu(\cdot)\text{ deterministic evolution of the distribution}}$$
$$+\underbrace{\frac{\overline{\sigma}^2}{2}\iint_{z,a}\alpha(\tilde{a},\tilde{z},\cdot)^2\tilde{a}^2\,\frac{d}{d\tilde{a}}\frac{dV}{dg}[\tilde{a},\tilde{z}]\,g(d\tilde{a},\tilde{z})}_{\text{diffusion of the distribution due to risk}} + \underbrace{\overline{\sigma}^2\alpha(a,z,\cdot)\,a\iint_{z,a}\alpha(\tilde{a},\tilde{z},\cdot)\,\tilde{a}\,\frac{d}{da}\frac{dV}{dg}[\tilde{a},\tilde{z}]\,g(d\tilde{a},\tilde{z})}_{\text{covariance of own state } a \text{ and distribution } \tilde{a}}$$
$$+\underbrace{\rho\,\overline{\sigma}\,\widehat{\sigma}\iint_{z,a}\alpha(\tilde{a},\tilde{z},\cdot)\,\tilde{a}\,\frac{d}{dZ}\frac{dV}{dg}[\tilde{a},\tilde{z}]\,g(d\tilde{a},\tilde{z})}_{\text{covariance of agg. state } Z \text{ and distribution } \tilde{a}} + \underbrace{\frac{\overline{\sigma}^2}{2}\iint_{(z,a)^{\otimes 2}}\alpha(\tilde{a},\tilde{z},\cdot)\tilde{a}\,\alpha(\tilde{a}',\tilde{z}',\cdot)\tilde{a}'\,\frac{d^2V}{dg^2}[\tilde{a},\tilde{z},\tilde{a}',\tilde{z}']\,g(d\tilde{a},\tilde{z})\,g(d\tilde{a}',\tilde{z}')}_{\text{covariance of distribution } \tilde{a} \text{ and } \tilde{a}'}$$

As before, the master equation's first three terms relate to (i) the gains and loss of individual decisions in consumption, portfolio holding through the operator $\mathcal{L}[V|c,\alpha]$, (ii) the direct effect of risk on income through the dynamics of Z, and (iii) the coupling (or non-local) term due to the anticipation of the deterministic evolution of wealth. The Master equation shows four additional second-order (non-local) terms in the second and third line: (iv) aggregate risk changes the curvature of the asset distribution for all the other agents $\tilde{a}$, which will affect the value through the cross derivative $V_{g\,\tilde{a}}$ for all agents α depending on their exposures $\alpha(\cdot)$. (v) Since the aggregate risk affects both the distribution and the agents' state, we also need to account for that covariance represented by the cross derivative $V_{g\,a}$, which scales now with the agent a exposure, (vi) similarly, we allowed for correlation between income with variance $\widehat{\sigma}$ and stock return $\overline{\sigma}$, which will affect value through the cross derivative V_{gZ}, and (vii) finally, we also have such shock affecting the correlation of all the agents $\tilde{a}$ and $\tilde{a}'$, which affect the value of agent a through their anticipation of the diffusion of the distribution and the second order term V_{gg}.

We implement our method with a projection, where we reduce the dimensionality of the g to its first moment $A = \int_{a,z} ag(da,z)$. Using the fact that total assets A equal aggregate capital K, since bonds b are in zero net supply, we obtain:

$$A = \int_{a,z} a g(da, z) = \int_{a,z} \underbrace{\alpha(\cdot)a}_{=k} g(da, z) + \overbrace{\int_{a,z} \underbrace{(1-\alpha(\cdot))a}_{=b} g(da, z)}^{=0} = K$$

Thanks to this property, we use A and K interchangeably in the following. Therefore, with such projection $V(a,z,g,Z) \approx \overline{V}(a,z,A,Z) = \overline{V}(a,z,K,Z)$, the master equation becomes:

$$\rho\overline{V} = \overbrace{\max_{c,\alpha} u(c) + \mathcal{L}[\overline{V}|c,\alpha](a,z,K,Z)}^{\text{standard HJB continuation value}} \overbrace{-\theta(Z-\bar{Z})\overline{V}_Z + \tfrac{\widehat{\sigma}^2}{2}\overline{V}_{ZZ}}^{\text{direct effect of risk of } Z \text{ on } \overline{V}} + \overbrace{\big(ZK^\alpha - \delta K - \mathcal{C}(K,Z|g)\big)\,\overline{V}_K}^{\text{deterministic evolution of the distribution}}$$
$$+ \underbrace{\frac{\overline{\sigma}^2}{2} K\,\overline{V}_K \times 0}_{\text{diffusion of distribution w/ risk} = 0} + \underbrace{\overline{\sigma}^2\alpha(a,z,K,Z)a\; K\,\overline{V}_{Ka}}_{\mathbb{C}\text{ov own state } a \text{ \& capital } K}$$
$$+ \underbrace{\rho\,\overline{\sigma}\,\widehat{\sigma}\; K\,\overline{V}_{KZ}}_{\mathbb{C}\text{ov agg. state } Z \text{ \& capital } K} + \underbrace{\frac{\overline{\sigma}^2}{2}K^2\;\overline{V}_{KK}}_{\mathbb{V}\text{ar agg. capital } K}$$

This Master equation would again merge a standard Aiyagari consumption saving with portfolio choice with a more standard representative agents model. For example, a Portfolio choice with aggregate income shocks – and zero aggregate bonds $B = 0$ would potentially collapse to the RBC model with investment risk:

$$\rho\overline{V} = \max_C u(C) + \overbrace{\big(ZK^\alpha - \delta K - C\big)\,\overline{V}_K}^{\text{agg. drift of capital}} \overbrace{-\theta(Z-\bar{Z})\overline{V}_Z + \tfrac{\widehat{\sigma}^2}{2}\overline{V}_{ZZ}}^{\text{direct effect of risk of } Z \text{ on } \overline{V}}.$$
$$+ \underbrace{\rho\,\overline{\sigma}\,\widehat{\sigma}\; K\,\overline{V}_{KZ}}_{\mathbb{C}\text{ov agg. state } Z \text{ \& capital } K} + \underbrace{\frac{\overline{\sigma}^2}{2}K^2\;\overline{V}_{KK}}_{\mathbb{V}\text{ar agg. capital } K}$$

We see that the additional terms that differ between the representative agent and the heterogeneous agent models relate to individuals' hedging against aggregate fluctuations in income Z (i-ii) and aggregate capital K (v), which is a form of self-insurance against systemic risk. Moreover, in the representative agents, there is no more diversification or leverage strategies.

We can now derive the optimal policies for the agents. Consumption has the same formula as it in model without portfolio choice:

$$c^\star(a, z_j, K, Z) = \begin{cases} u'^{-1}\big(\overline{V}_a(a,z_j,K,Z)\big) & a > \underline{a} \\ z_j \underbrace{(1-\alpha)ZK^\alpha}_{=w(K,Z)} + r\,a & \text{if } a \leq \underline{a} \end{cases} \tag{22}$$

Additionally, the optimal portfolio choice can also be derived:

$$\alpha^\star(a,z_j,K,Z) = -\frac{\overline{V}_{a(a,z_j,K,Z)}}{\overline{V}_{aa(a,z_j,K,Z)}a}\frac{r^k-\delta-r}{\overline{\sigma}^2} - \frac{\rho\widehat{\sigma}}{\overline{\sigma}}\frac{V_{aZ(a,z_j,K,Z)}}{\overline{V}_{aa(a,z_j,K,Z)}\,a} - \frac{\overline{V}_{Ka(a,z_j,K,Z)}\,K}{\overline{V}_{aa(a,z_j,K,Z)}\,a} \qquad \text{if} \quad a > \underline{a} \tag{23}$$

and $\alpha^\star(a,z_j,K,Z) = 0$ is $a \leq \underline{a}$ when agents hit the credit constraint. As claimed above, we see that the optimal portfolio choice changes for three reasons because of aggregate risk:
(i) it changes the value function due to risk, so the derivative V_a and curvature V_{aa} may change, (ii) it adds a cross-term for hedging against aggregate income shocks – since the asset return correlates with ρ with income risk, household would like to self insure, and (iii) agents want to hedge against systemic risk due to the aggregate dynamics in capital represented by the cross-term V_{Ka}. Depending on the sign of those last two terms, the optimal portfolio schedule in income and wealth might be dampened or amplified.

7.2 Financial frictions and the wealth distribution

In this section, we follow the approach by Fernández-Villaverde et al. (2023) and merge a problem à la Brunnermeier Sannikov with a standard Heterogeneous agents problem. This section is a work in progress.

References

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions and Benjamin Moll (2022), 'Income and wealth distribution in macroeconomics: A continuous-time approach', *The review of economic studies* **89**(1), 45–86.

Ahn, SeHyoun, Greg Kaplan, Benjamin Moll, Thomas Winberry and Christian Wolf (2018), 'When inequality matters for macro and macro matters for inequality', *NBER Macroeconomics annual* **32**(1), 1–75.

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

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

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

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

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

Bhandari, Anmol, David Evans, Mikhail Golosov and Thomas J Sargent (2021), 'Inequality, business cycles, and monetary-fiscal policy', *Econometrica* **89**(6), 2559–2599.

Bilal, Adrien (2023), Solving heterogeneous agent models with the master equation, Technical report, National Bureau of Economic Research.

Cardaliaguet, Pierre (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, Pierre, François Delarue, Jean-Michel Lasry and Pierre-Louis Lions (2019), *The master equation and the convergence problem in mean field games:(ams-201)*, Princeton University Press.

Carmona, René and Francois Delarue (2014), The master equation for large population equilibriums, *in* 'Stochastic Analysis and Applications 2014', Springer, pp. 77–128.

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

Carmona, René, François Delarue and Aimé Lachapelle (2013), 'Control of mckean–vlasov dynamics versus mean field games', *Mathematics and Financial Economics* **7**(2), 131–166.

Carmona, René, François Delarue and Daniel Lacker (2016), 'Mean field games with common noise', *The Annals of Probability* **44**(6), 3740–3803.

Fernández-Villaverde, Jesús, Samuel Hurtado and Galo Nuno (2023), 'Financial frictions and the wealth distribution', *Econometrica* **91**(3), 869–901.

Gabaix, Xavier, Jean-Michel Lasry, Pierre-Louis Lions and Benjamin Moll (2016), 'The dynamics of inequality', *Econometrica* **84**(6), 2071–2111.

Gu, Zhouzhou, Mathieu Laurière, Sebastian Merkel and Jonathan Payne (2023), Deep learning solutions to master equations for continuous time heterogeneous agent macroeconomic models, Technical report.

Gu, Zhouzhou, Mathieu Lauriere, Sebastian Merkel and Jonathan Payne (2024), 'Global solutions to master equations for continuous time heterogeneous agent macroeconomic models', *arXiv preprint arXiv:2406.13726* .

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

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

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

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

Moll, Benjamin (2025), 'The trouble with rational expectations in heterogeneous agent models: A challenge for macroeconomics', *The Economic Journal* p. ueaf104.

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

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

Schaab, Andreas (2020), 'Micro and macro uncertainty', *Available at SSRN 4099000* .

Schaab, Andreas (2021), Essays in Macroeconomics, PhD thesis, Harvard University.

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

Yang, Yucheng, Chiyuan Wang, Andreas Schaab and Benjamin Moll (2026), 'Structural reinforcement learning for heterogeneous agent macroeconomics', *arXiv preprint arXiv:2512.18892* .

Young, Eric 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.

A A primer on the Master Equation

In this section, we consider the most general Heterogeneous Agents model: each agent is part of a continuum and interacts with each other. The complete reference for the derivation of the Master Equation can be found in Lions' lectures in College de France, Cardaliaguet et al. (2019) Cardaliaguet (2013/2018) and Carmona and Delarue (2018).

We consider the general case when drift, variance, jump process, and instantaneous return depend on its individual states $x_{i,t} := x \in \mathbb{X} \subset \mathbb{R}^d$, as well as the states of all the other agents, through their distribution $m \in \mathcal{P}(\mathbb{X})$ (also called measure), as well as through an aggregate state $\mathcal{X}_t \in \mathbb{R}^\ell$. For ease of notations, we denote $x_{i,t} := x$ the individual state. The agent solves the following optimal control problem, choosing control $c \in \mathbb{C} \subset \mathbb{R}^m$

$$V(x_{t_0}, m_{t_0}, \mathcal{X}_{t_0}) = \max_{c_t} \mathbb{E}_{t_0} \int_{t_0}^{\infty} e^{-\rho t} \mathcal{L}(x_t, m, \mathcal{X}_t, c_t)\ dt$$

and subject of the dynamics of the individual states of all agents as well as the aggregate states.

$$dx_t = b(x_t, m_t, \mathcal{X}_t, c_t)dt + \sigma(x_t, m_t, \mathcal{X}_t, c_t)dB_t + \overline{\sigma}(x_t, m_t, \mathcal{X}_t, c_t)dB_t^0 + \gamma(x_t, m_t, \mathcal{X}_t, c_t)dJ_t$$

With dB_t the idiosyncratic Brownian noise, dJ_t^i the idiosyncratic jump process with intensity $\lambda(x_t, m_t, \mathcal{X}_t, c_t)$ over n_J states and dB_t^0 the common noise.

Dynamics of the aggregate states:

$$d\mathcal{X}_t = \mu(m_t, \mathcal{X}_t)dt + \widehat{\sigma}(m_t, \mathcal{X}_t)dB_t^0 + \widehat{\gamma}(m_t, \mathcal{X}_t)dJ_t^0$$

with dB_t^0 the common noise, and dJ_t^0 the common jump-process with intensity $\widehat{\lambda}(m_t, \mathcal{X}_t)$ over n_J^0 states.

Before deriving the master equation, let us consider the optimal control problem of the agents. We use Itô formula and the dynamic programming principle, and define the Hamiltonian of the control problem:

$$\begin{aligned}\mathcal{H}(x, m, \mathcal{X}, u, p, q, c) = \mathcal{L}(x, m, \mathcal{X}, c) + b(x, m, \mathcal{X}, c) \cdot p + \mathrm{Tr}\big([\sigma\sigma' + \overline{\sigma\sigma}'](x, m, \mathcal{X}, c)\ \ q\big)\\ \textstyle\sum_{n=1}^{n_J^i} \lambda^n(x, m, \mathcal{X}, c)\Big(u^n(x+\gamma(x, m, \mathcal{X}, c), \cdot) - u\Big)\end{aligned}$$

with $u \in \mathbb{R}$, $p \in \mathbb{R}^d$ and $q \in \mathbb{R}^{d\times d}$ representing respectively, the value function, its gradient, and its Hessian. The optimal control $c^\star$ is naturally maximizing:

$$c^\star \in \operatorname*{argmax}_c \mathcal{H}\big(x, m, \mathcal{X}, V, D_x V, D_{xx} V, c\big)$$

This optimal control provides the agents' dynamics, the optimal drift, variance, and jumps. Since it depends on the value V and its derivatives, we will turn now on the derivation of the master equation. For that, we need to understand how the interactions between agents affect their value.

A primer on Lions' derivative on the space of distribution

In the general case, in time-varying models with aggregate shocks, the value function of the agents is a function of the distribution $V(x, \boldsymbol{x}, \mathcal{X}) := V(x, m_{\boldsymbol{x}}, \mathcal{X})$. To apply dynamic programming, we hence rely on a concept of derivative – pioneered by P.L. Lions – of functions on the space of distributions.

Empirical distribution interpretation We observe that in a $N-$agents equilibrium, the agents' states can be summarized by the N-vector $\boldsymbol{x}$ and we have that the value of each agents depends on the empirical distribution as:

$$U(\boldsymbol{x}) := U(m_{\boldsymbol{x}}^N) \qquad\qquad m_{\boldsymbol{x}}^N(x) = \frac{1}{N}\sum_{n=1}^{N}\delta_{\{x_i=x\}}$$

with $m_{\boldsymbol{x}}^N(x)$ the empirical distribution for the N player

In that case, the Lions derivative is simply $D_{x_i}U(\boldsymbol{x}) = \frac{1}{N}D_mU(m_{\boldsymbol{x}}^N, x_i)$ where we consider the change of one agents x_i only. Note that in practice, we will take the derivative with respect to all agents' states, and then we integrate over them – hence the dummy variable y which is the one we integrate over in the derivation below.

Connection with Fréchet derivative. Note that the Lions derivative is *not* the same as the Fréchet derivative. The connection is shown with this relation: $D_mU(\cdot, m_{\boldsymbol{x}}; y) = D_y\frac{\delta U}{\delta m}(m; y) \in \mathbb{R}^d$ with $\frac{\delta U}{\delta m}(m)(h)$ the Frechet derivative in the space of signed measures in m, which is an operator on $\mathcal{P}(\mathbb{R}^d)$ and in the direction h. Hence, $\frac{\delta U}{\delta m}(m)(\cdot)$ is a signed measure $\mathcal{P}(\mathbb{R}^d)$, but could also be "identified" with an element of L^2, hence a "function" of $y \in \mathbb{R}^d$, where we can take the gradient[10], w.r.t. $y \in \mathbb{R}^d$.

Probabilistic interpretation. An alternative interpretation behind the derivative D_mU in the space of distribution is to represent a function of a measure $U(m)$ as a function of the random variable following this law $\widetilde{U}[X]$ and $X \sim m$ – called "lifting" or "extension". The derivative w.r.t. to the measure is the derivative of this lifted function: $D_mU(m, y) := D_y\widetilde{U}([X], y)$. One can also use the notion using "intrinsic derivative" $\widetilde{U}[X] := \frac{\delta U}{\delta m}$, on the space of signed measure. As a result, $D_mU(m, y) = D_y\frac{\delta U}{\delta m}(m, y)$.

[10]For derivative of a measure in L^2 that is not continuous/differentiable, we can consider the derivative in the weak sense, since we integrate against m which is assumed to be at least L^2.

Informal derivation of the master equation

More details are needed here.

We obtain the complete master equation:

$$
\begin{aligned}
-\partial_t V + \rho V &= \max_c \mathcal{H}\big(x, m, \mathcal{X}, V, D_x V, D_{xx} V, c^\star\big) \\
&+ \mu(m, \mathcal{X}) \cdot D_{\mathcal{X}} V + \operatorname{Tr}\big(\widehat{\sigma}\widehat{\sigma}' D_{\mathcal{X}\mathcal{X}} V\big) + \sum_{n=1}^{n_J^0} \widehat{\lambda}^n(m, \mathcal{X}) \Big(V \circ \widehat{\gamma}^n(m, \mathcal{X}) - V\Big) \\
&+ \int_{\mathbb{X}} D_m V(x, m, \mathcal{X}; y) \cdot b(y, m, \cdot) m(dy) + \int_{\mathbb{X}} \operatorname{Tr}\big[(\sigma\sigma' + \overline{\sigma}\overline{\sigma}')(y, \cdot) D_y\big(D_m V(x, m, \mathcal{X}; y)\big)\big] m(dy) \\
&+ 2 \int_{\mathbb{X}} \operatorname{Tr}\big(\overline{\sigma}(x, \cdot)\overline{\sigma}(y, \cdot)' D_x D_m V(x, \cdot; y)\big) m(dy) + \int_{\mathbb{X}} \operatorname{Tr}\big(\overline{\sigma}(y, \cdot)\widehat{\sigma}(\cdot)' D_m D_{\mathcal{X}} V(x, \cdot; y)\big) m(dy) \\
&+ \iint_{\mathbb{X}} \operatorname{Tr}\big(\overline{\sigma}(y, \cdot)\overline{\sigma}(y', \cdot)' D^2_{mm} V\big)(x, \cdot; y, y') m(dy) m(dy')
\end{aligned}
$$

Let us also do some remarks on this second-order master equation and the effects of the aggregate uncertainty:

- This equation is of second-order, since the derivative w.r.t. the distribution m is of second order both with the terms $D_y(D_m V)$ and $D^2_{mm} V$.
- As in the case without common noise, the terms in the first two lines correspond to the control dynamics, involved in the HJB, and the anticipation of an agent of its future – individual and aggregate – states.
- The last five terms are non-local terms and describe the agents' evolution – as would appear in a Kolmogorov Forward equation. To be more specific, these terms show how the value V changes when:

(i) the agent's control and drift $b(\cdot)$ change the state x and as a result it impacts the evolution of the distribution. Since the agents observe and anticipate all the other agents' moves, the integral is drawn on all $\mathbb{X}$.

(ii) the shape of the distribution is distorted due to idiosyncratic $\sigma(\cdot)$ and aggregate $\bar{\sigma}(\cdot)$ risk, for each agent' state y

(iii) a comovement between the state of the agent x and the distribution m due to the impact of the aggregate shocks $\overline{\sigma}$ – that affect all the agents y.

(iv) a comovement between the aggregate state $\mathcal{X}$ due to risk $\widehat{\sigma}$ and the distribution m due to the impact of the aggregate risk $\overline{\sigma}$ – that also affect all the agents y.

(v) the comovement between the other agents. This diffusion term is thus integrated over the whole space twice: the agent x will account for how the agents' states y comove with the other players y' due to aggregate risk $\overline{\sigma}$.

Example 1: no direct aggregate risk on agent's states

In this example, we consider the case where the aggregate risk doesn't affect additively the individual states of the agents, $\overline{\sigma} = 0$. As a result, the agents' distribution doesn't directly comove with the aggregate risk, making many of the cross- and second-derivatives

Note that, for the same reasons, the aggregate jump-process dJ_t^0 does not feature any cross derivatives with the distribution, as it doesn't affect directly the individual states (and hence the distribution m).

In that context, the master equation rewrites with simplification:

$$-\partial_t V + \rho V = \max_c \mathcal{H}\big(x, m, \mathcal{X}, V, D_x V, D_{xx} V, c^\star\big) + \mu(\cdot) \cdot D_{\mathcal{X}} V + \operatorname{Tr}\big(\widehat{\sigma}\widehat{\sigma}' D_{\mathcal{X}\mathcal{X}} V\big) + \textstyle\sum_n \widehat{\lambda}^n \Big(V \circ \widehat{\gamma}^n - V\Big)$$
$$+ \int_{\mathbb{X}} D_m V(x, \cdot; y) \cdot b(y, m, \cdot) m(dy) + \int_{\mathbb{X}} \operatorname{Tr}\big[\sigma\sigma'(y, \cdot) D_y\big(D_m V(x, \cdot; y)\big)\big] m(dy)$$

The value is still affected indirectly by the dynamics of the aggregate states $\mathcal{X}$, whose evolution is well anticipated by agents in the terms $\mu \cdot D_{\mathcal{X}} V + \widehat{\sigma}\widehat{\sigma}' D_{\mathcal{X}\mathcal{X}} V$

This example encompasses a large class of Heterogeneous Agent models, for example Krusell and Smith (1998) or many other HANK models, where typically productivity shocks affects TFP Z_t that then enter indirectly in agents budget constraints/wealth accumulation through wages and interest rates.

Moreover, this equation is analogous to the infinite-dimensional HJB expressed in equation (43) of Appendix A.1. of Ahn, Kaplan, Moll, Winberry and Wolf (2018)

Also Schaab's JMP and Bilal's paper

Example 2: no aggregate risk

In this example, we consider the case where there is no aggregate risk in any of the dynamics, either Brownian risk $\overline{\sigma} = \widehat{\sigma} = 0$, nor Jump-process $\widehat{\lambda}^n = 0$

In that context, the master equation rewrites with simplification:

$$-\partial_t V + \rho V = \max_c \mathcal{H}\big(x, m, \mathcal{X}, V, D_x V, D_{xx} V, c^\star\big) + \mu(\cdot) \cdot D_{\mathcal{X}} V$$
$$+ \int_{\mathbb{X}} D_m V(x, \cdot; y) \cdot b(y, m, \cdot) m(dy) + \int_{\mathbb{X}} \operatorname{Tr}\big[\sigma\sigma'(y, \cdot) D_y\big(D_m V(x, \cdot; y)\big)\big] m(dy)$$

We see that the term depending on the transport of the distribution $b(y, \cdot) D_m V(x, \cdot, y)$ at point y and the distortion due to idiosyncratic risk remains the same as in the previous case. The indirect impact also disappears this time: $\widehat{\sigma}\widehat{\sigma}' D^2_{\mathcal{X}\mathcal{X}} V = 0$

B A general derivation of the master equation

For completeness, we derive the master equation from the N player game, in the general case when the drift, variance, jumps and instantaneous returns depend on all states $x_{i,t} \in \mathbb{X} \subset \mathbb{R}^d$, the states of other agents $\boldsymbol{x} = \{x_j\}_{j=1\dots N}$, additional aggregates states $\mathcal{X}_t \in \mathbb{R}^\ell$

$$V(x_{i,t_0}, \boldsymbol{x}_{t_0}, \mathcal{X}_{t_0}) = \max_{c_{i,t}} \int_{t_0}^{\infty} e^{-\rho t} \mathcal{L}(x_{i,t}, \boldsymbol{x}_t, \mathcal{X}_t, c_{i,t})\ dt$$

Dynamics of individual states:

$$dx_{i,t} = b(x_{i,t}, \boldsymbol{x}_t, \mathcal{X}_t, c_{i,t})dt + \sigma(x_{i,t}, \boldsymbol{x}_t, \mathcal{X}_t, c_{i,t})dB_t^i + \overline{\sigma}(x_{i,t}, \boldsymbol{x}, \mathcal{X}_t, c_{i,t})dB_t^0 + \gamma(x_{i,t}, \boldsymbol{x}_t, \mathcal{X}_t, c_{i,t})dJ_t^i$$

With dB_t^i idiosyncratic Brownian noise, dJ_t^i the idiosyncratic jump process with intensity $\lambda(x_{i,t}, \boldsymbol{x}_t, \mathcal{X}_t, c_{i,t})$ over n_J states and dB_t^0 the common noise.

Dynamics of the aggregate states:

$$d\mathcal{X}_t = \mu(\boldsymbol{x}_t, \mathcal{X}_t)dt + \widehat{\sigma}(\boldsymbol{x}_t, \mathcal{X}_t)dB_t^0 + \widehat{\gamma}(\boldsymbol{x}_t, \mathcal{X}_t)dJ_t^0$$

with dB_t^0 the common noise, and dJ_t^0 the common jump-process with intensity $\widehat{\lambda}(\boldsymbol{x}_t, \mathcal{X}_t)$ over n_J^0 states.

Define the Hamiltonian of the control problem:

$$\begin{aligned}\mathcal{H}(x, \boldsymbol{x}, \mathcal{X}, u, p, q) = \max_c\ & \mathcal{L}(x, \boldsymbol{x}, \mathcal{X}, c) + b(x, \boldsymbol{x}, \mathcal{X}, c) \cdot p + \mathrm{Tr}\big([\sigma\sigma' + \overline{\sigma\sigma}'](x, \boldsymbol{x}, \mathcal{X}, c)\ \ q\big) \\ & \textstyle\sum_{n=1}^{n_J^i} \lambda^n(x, \boldsymbol{x}, \mathcal{X}, c)\Big(u^n(x+\gamma(x, \boldsymbol{x}, \mathcal{X}, c), x, \boldsymbol{x}, \mathcal{X}) - u\Big)\end{aligned}$$

with $p \in \mathbb{R}^d$ and $q \in \mathbb{R}^{d\times d}$. We define the optimal control and optimal d-dimensional drift:

$$\begin{aligned} & c^\star \in \underset{c}{\mathrm{argmax}} \mathcal{L}(\cdot, c) + b(\cdot, c) \cdot p + \mathrm{Tr}\big([\sigma\sigma' + \overline{\sigma\sigma}'](\cdot, c)q\big) + \textstyle\sum \lambda^n(\cdot, c)(u^n \circ \gamma(\cdot, c) - u) \\ \Rightarrow \qquad & b(x, \boldsymbol{x}, \mathcal{X}, c^\star) = D_p\mathcal{H}(x, \boldsymbol{x}, \mathcal{X}, D_x V, D_{xx}^2 V)\end{aligned}$$

Dynamic Programming principle:

$$V(x_{i,t_0}, \boldsymbol{x}_{t_0}, \mathcal{X}_{t_0}) = \max_{c_{i,t}} \int_{t_0}^{t_1} e^{-\rho t} \mathcal{L}(x_{i,t}, \boldsymbol{x}_t, \mathcal{X}_t, c_{i,t})\ dt + e^{-\rho(t_1 - t_0)} V(x_{i,t_1}, \boldsymbol{x}_{t_1}, \mathcal{X}_{t_1})$$

Ito Formula:
The first terms are standard, but we also need to account for all the dynamics of all agents!

Value function $V^i := V(x_{i,t}, \boldsymbol{x}_t, \mathcal{X}_t)$, the derivatives with respect to its own states, $D_{x_j} V^i \equiv D_{x_j} V(x_{i,t}, \boldsymbol{x}_t, \mathcal{X}_t) \in \mathbb{R}^d$, derivatives w.r.t to other agents j: $D_{x_i x_j} V$ and $D_{x_j x_j} V^i \equiv D_{x_j x_j} V(x_{i,t}, \boldsymbol{x}_t, \mathcal{X}_t) \in \mathbb{R}^{d\times d}$ and the derivatives with respect to the aggregate state $D_{\mathcal{X}} V \equiv D_{\mathcal{X}} V(x_{i,t}, \boldsymbol{x}_t, \mathcal{X}_t) \in \mathbb{R}^\ell$ and $D_{\mathcal{X}\mathcal{X}} V = D_{\mathcal{X}\mathcal{X}} V(x_{i,t}, \boldsymbol{x}_t, \mathcal{X}_t)$.

The optimal drift $D_p\mathcal{H}^i \equiv D_p\mathcal{H}(x_{i,t}, \boldsymbol{x}_t, \mathcal{X}_t, D_{x_i} V^i, D_{x_i x_i} V^i)$. We can derive the master equation:

$$
\begin{aligned}
dV^i(x_{i,t},\boldsymbol{x}_t,\mathcal{X}_t) &= \partial_t V^i - \rho V^i + \mathcal{H}^i\big(x_{i,t},\boldsymbol{x}_t,\mathcal{X}_t,V^i,D_{x_i}V^i,D_{x_ix_i}V^i\big) \\
&\quad + \mu(\boldsymbol{x}_t,\mathcal{X}_t)\cdot D_{\mathcal{X}}V^i + \mathrm{Tr}\big(\widehat{\sigma}\widehat{\sigma}' D_{\mathcal{X}\mathcal{X}}V^i\big) + \sum_{n=1}^{n_J^0}\widehat{\lambda}(\boldsymbol{x}_t,\mathcal{X}_t)\Big(V^i\circ\widehat{\gamma}^i(\boldsymbol{x}_t,\mathcal{X}_t) - V^i\Big) \\
&\quad + \sum_{j\neq i} D_{x_j}V^i\cdot D_p\mathcal{H}^j \ + \sum_{j\neq i}\ \textstyle\sum_{n=1}^{n_J^i}\lambda^n(x_j,\boldsymbol{x}_t,\cdot)\Big(V^i(\cdot,\boldsymbol{x}_t\circ\gamma(x_j,\boldsymbol{x}_t,\cdot),\boldsymbol{x},\mathcal{X}) - V^i\Big) \\
&\quad + \sum_{j\neq i}\mathrm{Tr}\big((\sigma^j\sigma^{j\prime}+\overline{\sigma}^j\overline{\sigma}^{j\prime})D_{x_jx_j}V^i\big) + \sum_{\substack{j,k\neq i\\ j\neq k}}\mathrm{Tr}\big(\overline{\sigma}^j\overline{\sigma}^{k\prime}D_{x_jx_k}V^i\big) + \sum_j \mathrm{Tr}\big(\overline{\sigma}^j\widehat{\sigma}' D_{x_j\mathcal{X}}V^i\big)
\end{aligned}
$$

Now, let reformulate the value function as a function of the measure $V(x,\boldsymbol{x},\mathcal{X}) \equiv V(x,m_{\boldsymbol{x}},\mathcal{X})$. We hence rely on the derivatives in the space of measure. Starting from the observation that $V(\cdot,\boldsymbol{x},\cdot) \equiv V(\cdot,m_{\boldsymbol{x}},\cdot)$.

We use the Lions derivative $D_mV(\cdot,m_{\boldsymbol{x}};y) = D_y\frac{\delta V}{\delta m}(m;y) \in \mathbb{R}^d$ with $\frac{\delta V}{\delta m}(m)(h)$ the Frechet derivative in the space of signed measures in m, which is an operator on $\mathcal{P}(\mathbb{R}^d)$ and in the direction h. Hence, $\frac{\delta V}{\delta m}(m)(\cdot)$ is a signed measure $\mathcal{P}(\mathbb{R}^d)$, but could also be "identified" with an element of L^2, hence a "function" of $y \in \mathbb{R}^d$, where we can take the gradient, w.r.t. $y \in \mathbb{R}^d$.

Derivative in the space of measure. The idea behind the derivative D_mV in the space of functions in the Wasserstein space 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 : $D_mg(m,y) \equiv D_y\widetilde{g}([X],y)$. One can also use the notion using "intrinsic derivative" $\frac{\delta g}{\delta m}$ which is nothing else than the derivative in the Frechet sense, when extending $\mathcal{P}(\mathbb{R})$ to the space of signed measure. As a result, $D_mg(m,y) \equiv D_y\frac{\delta g}{\delta m}(m,y)$. The complete references are found in Cardaliaguet (2013/2018) and Carmona and Delarue (2018).

First order derivative:

$$
D_{x_j}V(\cdot,\boldsymbol{x},\cdot) \simeq \frac{1}{N-1}D_mV(\cdot,m_{\boldsymbol{x}};y) \ = D_mV(\cdot,m_{\boldsymbol{x}};y)m_{\boldsymbol{x}}(dy)
$$

So, the first order terms become:

$$
\sum_{j\neq i} D_{x_j}V^i\cdot D_p\mathcal{H}^j \ \simeq \int_{\mathbb{X}} D_mV^i(x,m_{\boldsymbol{x}};y)\cdot D_p\mathcal{H}(y,m_{\boldsymbol{x}},\cdot)m_{\boldsymbol{x}}(dy)
$$

Second order derivative w.r.t other agents:

$$
\mathrm{Tr}\big(D_{x_jx_j}V(\cdot,\boldsymbol{x},\cdot)\big) \simeq \frac{1}{N-1}\operatorname*{div}_y\big[D_mV\big](x,m_{\boldsymbol{x}};y) + \frac{1}{(N-1)^2}\mathrm{Tr}\big[D^2_{mm}V\big](x,m_{\boldsymbol{x}};y)
$$

So, at the limit $N \to \infty$, those second-order terms in idiosyncratic noise become:

$$\sum_{j\neq i} \mathrm{Tr}\big((\sigma^j\sigma^{j'}+\overline{\sigma}^j\overline{\sigma}^{j'}) + D_{x_jx_j}V^i\big) \ \simeq \ \int_{\mathbb{X}} \mathrm{Tr}\big[(\sigma^j\sigma^{j'}+\overline{\sigma}^j\overline{\sigma}^{j'})D_y\big(D_mV^i(x,m_{\boldsymbol{x}},\mathcal{X};y)\big)\big](y,m_{\boldsymbol{x}},\mathcal{X})m_{\boldsymbol{x}}(dy)$$

where $D_y\big(D_mV^i(x,m_{\boldsymbol{x}};y)\big)$ is a matrix $\in \mathbb{R}^{d\times d}$.

Second order cross derivative w.r.t other agents:

$$\begin{aligned}
\sum_{(j,k)\neq(i,i)} \mathrm{Tr}\big(D_{x_jx_k}V(\cdot,\boldsymbol{x},\cdot)\big) &= 2\sum_{k\neq i}\mathrm{Tr}\big(D_{x_ix_k}V(\cdot,\boldsymbol{x},\cdot)\big) + \sum_{j,k\neq i}\mathrm{Tr}\big(D_{x_jx_k}V(\cdot,\boldsymbol{x},\cdot)\big)\\
2\sum_{k\neq i}\mathrm{Tr}\big(D_{x_ix_k}V(\cdot,\boldsymbol{x},\cdot)\big) &\simeq \frac{2}{N-1}\sum_{k\neq i}\mathrm{Tr}\big(D_xD_mV(x,\boldsymbol{x},\cdot;y)\big)\\
&\simeq 2\int_{\mathbb{X}}\mathrm{Tr}\big(D_xD_mV(x,\boldsymbol{x},\cdot;y)\big)m_{\boldsymbol{x}}(dy)\\
\sum_{j,k\neq i}\mathrm{Tr}\big(D_{x_jx_k}V(\cdot,\boldsymbol{x},\cdot)\big) &\simeq \frac{1}{(N-1)^2}\sum_{j,k\neq i}\mathrm{Tr}\big(D^2_{mm}V\big)(x,\boldsymbol{x},\mathcal{X};x_j,x_k)\\
&\simeq \iint_{\mathbb{X}}\mathrm{Tr}\big(D^2_{mm}V\big)(x,\boldsymbol{x},\mathcal{X};y,y')m_{\boldsymbol{x}}(dy)m_{\boldsymbol{x}}(dy')
\end{aligned}$$

Hence, the second order terms showing interaction with other agents:

$$\begin{aligned}
\sum_{(j,k)\neq(i,i)}\mathrm{Tr}\big(\overline{\sigma}^j\overline{\sigma}^{k'}D_{x_jx_k}V^i\big) &= 2\int_{\mathbb{X}}\mathrm{Tr}\big(\overline{\sigma}(x,\cdot)\overline{\sigma}(y,\cdot)'D_xD_mV(x,\boldsymbol{x},\cdot;y)\big)m_{\boldsymbol{x}}(dy)\\
&+\iint_{\mathbb{X}}\mathrm{Tr}\big(\overline{\sigma}(y,\cdot)\overline{\sigma}(y',\cdot)'D^2_{mm}V\big)(x,\boldsymbol{x},\mathcal{X};y,y')m_{\boldsymbol{x}}(dy)m_{\boldsymbol{x}}(dy')
\end{aligned}$$

Second-order cross derivative between other agents and aggregate noise:

$$\sum_j\mathrm{Tr}\big(\overline{\sigma}^j\widehat{\sigma}(\mathcal{X}_t)'D_{x_j\mathcal{X}}V^i\big) = \int_{\mathbb{X}}\mathrm{Tr}\big(\overline{\sigma}(y,\cdot)\widehat{\sigma}(\mathcal{X}_t)'D_mD_{\mathcal{X}}V\big)m(dy)$$

We obtain the complete master equation:

$$\begin{aligned}
-\partial_tV+\rho V &= \mathcal{H}\big(x,m,\mathcal{X},V,D_xV,D_{xx}V,c^\star\big)\\
&+\mu(m,\mathcal{X})\cdot D_{\mathcal{X}}V+\mathrm{Tr}\big(\widehat{\sigma}\widehat{\sigma}'D_{\mathcal{X}\mathcal{X}}V\big)+\sum_{n=1}^{n_J^0}\widehat{\lambda}^n(m,\mathcal{X})\Big(V\circ\widehat{\gamma}^n(m,\mathcal{X})-V\Big)\\
&+\int_{\mathbb{X}}D_mV(x,\cdot;y)\cdot D_p\mathcal{H}(y,\cdot)m(dy)+\int_{\mathbb{X}}\sum_{n=1}^{n_J^0}\lambda^n(y,\cdot)\Delta_mV(x,\cdot;y)\circ\gamma(y,\cdot)m(dy)\\
&+\int_{\mathbb{X}}\mathrm{Tr}\big[(\sigma\sigma'+\overline{\sigma}\overline{\sigma}')(y,\cdot)D_y\big(D_mV(x,m,\mathcal{X};y)\big)\big](y,m,\mathcal{X})m(dy)\\
&+2\int_{\mathbb{X}}\mathrm{Tr}\big(\overline{\sigma}(x,\cdot)\overline{\sigma}(y,\cdot)'D_xD_mV(x,\cdot;y)\big)m(dy)+\int_{\mathbb{X}}\mathrm{Tr}\big(\overline{\sigma}(y,\cdot)\widehat{\sigma}(\mathcal{X}_t)'D_mD_{\mathcal{X}}V(x,m,\mathcal{X};y)\big)m(dy)\\
&+\iint_{\mathbb{X}}\mathrm{Tr}\big(\overline{\sigma}(y,\cdot)\overline{\sigma}(y',\cdot)'D^2_{mm}V\big)(x,\cdot;y,y')m(dy)m(dy')
\end{aligned}$$

$$\mathcal{H}(x,m,\mathcal{X},V,D_xV,D_{xx}V) = \max_c \ \mathcal{L}(x,m,\mathcal{X},c) + b(x,m,\mathcal{X},c)\cdot D_xV + \mathrm{Tr}\big([\sigma\sigma' + \overline{\sigma\sigma}'](x,m,\mathcal{X},c)\ D_{xx}V\big)$$

$$\textstyle\sum_{n=1}^{n_J^i} \lambda^n(x,m,\mathcal{X},c)\Big(V^n(x{+}\gamma(x,m,\mathcal{X},c),x,m,\mathcal{X}) - V\Big)$$

$$\begin{aligned}
-\partial_t V + \rho V &= \mathcal{H}\big(x,m,\mathcal{X},V,D_xV,D_{xx}V,c^\star\big)\\
&+ \mu(m,\mathcal{X})\cdot D_{\mathcal{X}}V + \mathrm{Tr}\big(\widehat{\sigma}\widehat{\sigma}' D_{\mathcal{X}\mathcal{X}}V\big) + \sum_{n=1}^{n_J^0} \widehat{\lambda}^n(m,\mathcal{X})\Big(V\circ\widehat{\gamma}^n(m,\mathcal{X}) - V\Big)\\
&+ \int_{\mathbb{X}} D_mV(x,\cdot;y)\cdot D_p\mathcal{H}(y,\cdot)m(dy) + \int_{\mathbb{X}} \sum_{n=1}^{n_J^0} \lambda^n(y,\cdot)\Delta_m V(x,\cdot;y)\circ\gamma(y,\cdot)m(dy)\\
&+ \int_{\mathbb{X}} \mathrm{Tr}\big[(\sigma\sigma'{+}\overline{\sigma\sigma}')(y,\cdot)D_y\big(D_mV(x,m,\mathcal{X};y)\big)\big](y,m,\mathcal{X})m(dy)\\
&+ 2\int_{\mathbb{X}} \mathrm{Tr}\big(\overline{\sigma}(x,\cdot)\overline{\sigma}(y,\cdot)'D_xD_mV(x,\cdot;y)\big)m(dy) + \int_{\mathbb{X}} \mathrm{Tr}\big(\overline{\sigma}(y,\cdot)\widehat{\sigma}(\mathcal{X}_t)'D_mD_{\mathcal{X}}V(x,m,\mathcal{X};y)\big)m(dy)\\
&+ \iint_{\mathbb{X}} \mathrm{Tr}\big(\overline{\sigma}(y,\cdot)\overline{\sigma}(y',\cdot)'D^2_{mm}V\big)(x,\cdot;y,y')m(dy)m(dy')
\end{aligned}$$

C Derivative in the space of probability measure

For this part of the document, I use the lecture notes of P. Cardaliaguet (2013/2018).

C.1 Derivative in L^2 sense

We consider a function $F:\mathcal{P}_2(\mathbb{R}^d)\to\mathbb{R}$ with p the dimension of heterogeneity and defined on the space of probability measures m with a finite second order moment: $\int_{\mathbb{R}^d}|x|^2m(dx)<\infty$ and we could us the Wasserstein distance (used a lot in MFG and optimal transport theory):

$$\boldsymbol{d}_2(g_1,g_2) := \Big(\inf_\pi \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2\pi(dx,dy)\Big)^{1/2} \qquad s.t.\quad \pi\in\mathcal{P}(\mathbb{R}^d\ \times\mathbb{R}^d)$$

$$\int_{\mathbb{R}^d\times\mathbb{R}^d}\phi(x)\pi(dx,dy) = \int_{\mathbb{R}^d}\phi(x)m_1(dx) \qquad \int_{\mathbb{R}^d\times\mathbb{R}^d}\phi(y)\pi(dx,dy) = \int_{\mathbb{R}^d}\phi(x)m_1(dx) \qquad \forall\phi\in\mathcal{C}^0_b(\mathbb{R}^d)$$

Definition C.1. *We say that the function $F:\mathcal{P}_2(\mathbb{R}^d)\to\mathbb{R}$ is continuously differentiable, i.e. $\mathcal{C}^1$ in the L^2-sense if there exists a bounded continuous map $\frac{\delta F}{\delta m}:\mathcal{P}^2\times\mathbb{R}^d\to\mathbb{R}$ such that for any $m,m'\in\mathcal{P}_2$*

$$F(m') - F(m) = \int_0^1\int_{\mathbb{R}^d}\frac{\delta F}{\delta m}((1-s)m+sm',y)d(m'-m)(y)$$

We say that $\frac{\delta F}{\delta m}$ is the L^2-derivative of F.

Notes :

- This L^2-derivative of F is also its Fréchet derivative in the space L^2 in the "direction" $m'-m$.

- Indeed, writing the relation differently:

$$\lim_{t\to 0^+} \frac{F(m+t(m'-m))-F(m)}{t} = \int_{\mathbb{R}^d} \frac{\delta F}{\delta m}(m,y)d(m'-m)(y) = \langle \frac{\delta F}{\delta m}(m,\cdot), m'-m\rangle$$

where $\langle \cdot, \cdot \rangle$ is a duality bracket in L^2

To be continued ...

C.2 Intrinsic derivative

Definition C.2. *We say that the function $F : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is continuously differentiable, i.e. $\mathcal{C}^1$ in the intrinsic sense if F is $\mathcal{C}^1$ in the L^2-sense and if V is differentiable with respect to the space variable with $D_y \frac{\delta F}{\delta m}$ continuous and bounded on $\mathbb{R}^d \times \mathcal{P}_2$. We define the intrinsic derivative $D_m F : \mathcal{P}_1 \times \mathbb{R}^d \to \mathbb{R}^d$ as*

$$D_m V(m,y) := D_y \frac{\delta F}{\delta m}(m,y)$$

there exists a bounded continuous map $\frac{\delta F}{\delta m} : \mathcal{P}^2 \times \mathbb{R}^d \to \mathbb{R}$ such that for any $m, m' \in \mathcal{P}_2$

C.3 Probabilistic interpretation and lifting

Lifting + approach of Carmona and Delarue:

Need to rewrite this part:

When one consider the measure $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, the Hilbertian structure will be used to study $D_\mu H(\mu)$: one need to analyze a *lifting* (or 'extension') $\widetilde{H}(\tilde{X})$ where $\widetilde{H}$ depends on the random variable $\tilde{X}$ (and thus defined on L^2) such that $\widetilde{H}(\tilde{X}) = H(\mathbb{P}_{\tilde{X}})$. The function H is said to differentiable in μ_0 if there exists a r.v. $X_0 \sim \mu_0$ such that $\widetilde{H}$ is (Fréchet) differentiable at $\tilde{X}_0$, and $D\widetilde{H}(\tilde{X}_0)$ is the 'representation' of the derivative of H at μ_0. We thus call $x \mapsto D_\mu H(\mu_0)(\cdot)$ the derivative of H w.r.t. μ_0 as a deterministic function and :

$$\begin{aligned} H(\mu) &= H(\mu_0) + D\widetilde{H}(\tilde{X}_0)\cdot(\tilde{X}-\tilde{X}_0) + o(||\tilde{X}-\tilde{X}_0||_2) \\ &= H(\mu_0) + D_\mu H(\mu_0)(\tilde{X}_0)\cdot(\tilde{X}-\tilde{X}_0) + o(||\tilde{X}-\tilde{X}_0||_2) \end{aligned}$$

A concrete example to show that this derivative is *not* a standard (Frechet) differential, is when we define: $H(\mu) = \int_{\mathbb{R}^d} h(x)\mu(dx) = \langle h, \mu\rangle$ (which is linear in L^2). Its lifting will be $\widetilde{H}(\tilde{X}) = \widetilde{\mathbb{E}}[h(\tilde{X})]$ and its derivative $D\widetilde{H}(\tilde{X})\cdot Y = \widetilde{\mathbb{E}}[Dh(\tilde{X})\cdot Y]$ and consequently: $D_\mu H(\mu_0)(\cdot) \equiv Dh(\cdot)$ (which is not equal to the Frechet-differential h).

D Huggett Model with Aggregate Shocks

We restate the Huggett consumption-saving problem in our setting. We follow a framework as in the Krusell-Smith model: we consider idiosyncratic state z as an n_z-states Markov jump process, and the aggregate risk on TFP Z_t follows a mean-reverting diffusion process – a Ornstein-Uhlenbeck diffusion process – which is analogous to $AR(1)$ process in discrete time.

The firm is using labor with a linear technology $Y_t = Z_t \bar{L}$ where $\bar{L} = \langle z, g_t \rangle = \int_{a,z} z g(da, z) = \mathbb{E}[z]$ is the exogenous labor supply, normalized to a constant $\bar{L} = 1$ to alleviate the notations. Since the firm is not using capital, the interest rate for agents bonds will be determined *implicitly* as we will see later.

As a result, Households, subject to these processes, save in assets a, and earn labor income zw where the wage w_t is the Marginal Product of Labor $MPL = Z_t$.

$$V_0 = \max_{\{c_t\}_t} \mathbb{E}_0 \int_0^\infty e^{-\rho t} u(c_t) dt \tag{24}$$

$$\text{s.t.} \quad da_t = s_t(a, z, Z, r, c)dt = \Big(z_t \underbrace{Z_t}_{\equiv w_t} + r_t\, a_t \; - c_t\Big)dt \qquad \& \qquad a_t \geq \underline{a} \tag{wealth}$$

$$dz_t = dJ_t \qquad \text{intensity} \quad \lambda_{ij} = \lambda(z_{it}, dz_{jt}) \tag{labor productivity}$$

$$dZ_t = \mu(Z_t)dt + \widehat{\sigma} dB_t^0 \tag{aggreg. productivity}$$

We denote g_t the distribution of states $(a_t, z_t) \in \mathbb{X} = [\underline{a}, \infty) \times \{z_1, \dots z_{n_z}\}$. The difference with the Krusell-Smith model is that the market clears in zero net supply, such that the aggregate saving is null.

$$\iint_{\mathbb{X}} a g_t(da, z) = \bar{B} = 0 \tag{25}$$

Therefore, aggregate saving $\bar{B}$ can not be used as an aggregate state variable.

As a result, the interest rate r is defined *implicitly* as a function of the distribution g. Additionally, in this class of model, the market clearing for bonds or for goods are equivalent by Walras law. As a result, the good resource constraint implies that output equals the aggregation of individual consumptions $\{c_t^\star(\cdot)\}$, which are policy decisions as a function of states that still remain to be determined (see below). When c is chosen optimally as a result of the stochastic control by Households, we can aggregate:

$$Y_t = C_t = \int_{a,z} c_t^\star(\cdot)\, g_t(da, z) \qquad \int_{a,z} s_t(a, z, Z, r, c^\star)\, g_t(da, z) = \int_{a,z} (zZ + ra - c^\star)\, g_t(da, z) = 0 \tag{26}$$

As a result, we can start with a guess for the dynamics of the interest rate:

$$dr_t = \mu_r dt + \sigma_r dB_t^0$$

where $\{\mu_r, \sigma_r\}$ are parameters determined in equilibrium. The goal of the projection will be to make those parameters explicit as a function of the rest of the model's dynamics.

D.1 Master equation in the Huggett model

The Master equation would write the value as a function of the individual states, assets and labor productivity $(a,z) \in \mathbb{X}$, the aggregate state Z, and the distribution $g \in \mathcal{P}(\mathbb{X})$ over (a,z), so $V = V(a,z,g,Z)$.

$$\rho V = \overbrace{\max_c u(c) + \mathcal{L}[\,V\,|c](t,a,z)}^{\text{standard HJB continuation value}} \quad \overbrace{-\theta(Z-\bar{Z})V_Z + \frac{\widehat{\sigma}^2}{2}V_{ZZ}}^{\text{direct effect of risk of } Z \text{ on } V}$$
$$+ \underbrace{\iint_{z,a} \frac{\delta V(a,z,g,Z,r)}{\delta g}[\tilde{a},\tilde{z}]\ \mathcal{L}^*[\,g\,|\,c^\star](t,d\tilde{a},\tilde{z})}_{\text{change due to distribution dynamics}} \tag{27}$$

with the generator $\mathcal{L}$ for the individual states (a,z):

$$\mathcal{L}[\,V\,|c^\star](a,z_j,g,Z,r) = V_a(a,z_j,\cdot)\underbrace{[z_j Z + r\,a - c^\star]}_{\text{change in saving}} + \underbrace{\lambda_j(V(t,a,z_{-j},\cdot) - V(t,a,z_j,\cdot))}_{\text{change in labor income}} \quad . \tag{28}$$

In equilibrium, the interest rate is determined implicitly by the market-clearing condition for bonds or goods, which are equivalent, by Walras law.

$$\iint_{\mathbb{X}} a g_t(da,z) = 0 \qquad\qquad Z + \int_{a,z} c^\star(\cdot)\, g_t(da,z) = 0$$

In this context, we can find the optimal policy in the presence of credit constraints and the value $V = V(a,z_j,g,Z)$, which makes the consumption depend on the states:

$$c^\star(a,z_j,g,Z) = \begin{cases} u'^{-1}(V_a(a,z_j,g,Z)) & a > \underline{a} \\ z_j Z + r\,a & \text{if } a \leq \underline{a} \end{cases}$$

D.2 Projection in the Huggett model

We now directly project the dynamics of the distribution onto the dynamics of the interest rate. If we consider the value projected on the interest rate $\overline{V} = \overline{V}(a,z,r,Z) = V(a,z,g,Z)$, where the agents forecast accurately and rationally the dynamics of the interest rate appendix D. In that case, the master equation / HJB of the agents writes (more simply) as:

$$\rho\overline{V} = \overbrace{\max_c u(c) + \mathcal{L}[\,\overline{V}\,|c](a,z)}^{\text{standard HJB continuation value}} \quad \overbrace{-\theta(Z-\bar{Z})\overline{V}_Z + \frac{\widehat{\sigma}^2}{2}\overline{V}_{ZZ}}^{\text{direct effect of risk of } Z \text{ on } \overline{V}}$$
$$+ \quad \overbrace{\mu_r\overline{V}_r + \frac{\sigma_r{}^2}{2}\overline{V}_{rr}}^{\text{direct effect of asset prices } r} \quad + \quad \overbrace{\widehat{\sigma}\sigma_r\,\overline{V}_{Z\,r}}^{\substack{\text{covariance btw income} \\ \text{and interest rate risk}}} \tag{29}$$

where the parameters μ_r and σ_r are the dynamics to be learned from the distributional dynamics. To make the mapping between the distributional dynamics and the interest rate dynamics, we will use the implicit function theorem in the equilibrium condition eq. (26). Similarly, the optimal consumption is expressed as a function of the states (a, z_j, r, Z):

$$c^{\star}(a, z_j, r, Z) = \begin{cases} u'^{-1}(\overline{V}_{a(a,z_j,g,Z)}) & a > \underline{a} \\ z_j Z + r\,a & \text{if } a \leq \underline{a} \end{cases}$$

For the following, we will need the elasticity of consumption to interest rates, which will be key to determining the movement of asset prices in equilibrium.

$$\frac{dc^{\star}(a, z_j, r, Z)}{dr} = \begin{cases} [u''(c(V_a))]^{-1} V_{a\,r}(\cdot) = -\gamma\; c^{\star}(a,z_j,r,Z)\; \frac{\overline{V}_{a\,r}}{\overline{V}_a} & a > \underline{a} \\ a & \text{if } a \leq \underline{a} \end{cases}$$

with $\overline{V} = \overline{V}(a, z, r, Z)$, where γ is the Relative Risk Aversion of the utility function, which is constant in CRRA but could otherwise depend on c and hence $\overline{V}_a$ when the constraint does not bind. In such case, the consumption sensitivity is inversely related to the cross derivative $V_{a\,r}$: when the marginal value of wealth increases in the interest rate, i.e. $V_{ar} > 0$, then $dc/dr < 0$, implying that the substitution effect dominates: agent save more when interests rise. We also see that when the constraint binds $a \leq \underline{a}$, the consumption sensitivity depends on saving: when $a \leq \underline{a} \leq 0$, indebted households are forced to reduce consumption when interest rates rise.

Similarly, the MPC with respect to aggregate income Z – which is wage here – can be written:

$$\frac{\partial c^{\star}(a, z_j, r, Z)}{\partial Z} = \begin{cases} -\gamma\; c^{\star}(a,z_j,r,Z)\; \frac{\overline{V}_{a\,Z}}{\overline{V}_a} & a > \underline{a} \\ z_j & \text{if } a \leq \underline{a} \end{cases}$$

where the sensitivity depends on the cross derivative $V_{a\,Z}$. When the marginal value of wealth V_a declines with income Z, the MPC dc/dZ is positive as expected. However, it can be much higher – and closer to 1 – when the constraint binds $a = \underline{a}$.

D.3 Asset prices as implicit function of states and the distribution of agents

The market-clearing/equilibrium condition in eq. (26) allow to define the asset prices as an implicit function of the distribution:

$$F(g, Z, r) = 0 \qquad \int_{a,z} (zZ + ra - c^{\star}(\cdot, r))\, g_t(da, z) = Z - \int_{a,z} c^{\star}(a, z, r, Z) g(da, z) = 0$$

Using the implicit function theorem, $F(g, Z, r) = 0$ implicitely defines a function $r = r(g, Z)$, for $g \in \mathcal{P}(\mathbb{X})$ and $Z \in \mathbb{R}_+$, and we can find their derivatives – in infinite dimension in the case of the distribution g.

$$
\begin{aligned}
r = r(g,Z) \qquad &\Rightarrow \qquad \frac{\delta r(g,Z)}{\delta g}[\tilde{a},\tilde{z}] = -[F_r(\cdot)]^{-1}\frac{\delta F(\cdot)}{\delta g}[\tilde{a},\tilde{z}] \\
&\Rightarrow \qquad \frac{dr(g,Z)}{dZ} = -[F_r(\cdot)]^{-1}F_Z(\cdot)
\end{aligned}
$$

where $\delta r/\delta g$ are again the intrinsic derivatives with respect to the distribution. In the case of the market-clearing and the constant aggregate saving, we obtain:

$$
\begin{aligned}
F_Z(g,Z,r) &= 1 - \int_{a,z}\frac{\partial c^\star(a,z,r,Z)}{\partial Z}g(da,z) =: 1 - \frac{\partial \mathcal{C}(r,Z|g)}{\partial Z} \\
F_r(g,Z,r) &= - \int_{a,z}\frac{\partial c^\star(a,z,r,Z)}{\partial r}g(da,z) =: -\frac{\partial \mathcal{C}(r,Z|g)}{\partial r} \\
\frac{\delta F(g,Z,r)}{\delta g}[\tilde{a},\tilde{z}] &= c^\star(\tilde{a},\tilde{z},r,Z)
\end{aligned}
$$

where the elasticity of consumption to the interest rate and the MPC are derived in the previous section. We notice two features of the equilibrium: First, the derivative of the market clearing w.r.t to the distribution depends on the individual decision $c^\star(\cdot)$. Second, the derivative with respect to the interest rate or aggregate productivity depends on the aggregation of agents' sensitivity, which is necessary to obtain the general equilibrium effects. This implies the equilibrium interest rate:

$$
\begin{aligned}
\frac{\delta r(g,Z)}{\delta g}[\tilde{a},\tilde{z}] &= \frac{1}{\frac{\partial \mathcal{C}(r,Z|g)}{\partial r}}c^\star(\tilde{a},\tilde{z},r,Z) \\
\frac{\partial r(g,Z)}{\partial Z} &= \frac{1-\frac{\partial \mathcal{C}(r,Z|g)}{\partial Z}}{\frac{\partial \mathcal{C}(r,Z|g)}{\partial r}}
\end{aligned}
\tag{30}
$$

D.4 Asset price dynamics and Master equation

If we know that the asset price r is a function of the aggregate productivity Z as well as the distribution g, therefore we can deduce the dynamics using Ito formula:

$$
r = r(g,Z) \qquad \Rightarrow \qquad dr = \frac{\partial r(g,Z)}{\partial Z}dZ + \int_{\tilde{a},\tilde{z}}\frac{\delta r(g,Z)}{\delta g}[\tilde{a},\tilde{z}]\ \mathcal{L}^\star(g\,|\,c^\star)[\tilde{a},\tilde{z}]g(d\tilde{a},z)dt
$$

Putting all the elements together, this implies the following HJB equation for states (a,z,r,Z):

$$
\begin{aligned}
\rho\overline{V} = \ &\overbrace{\max_c u(c) + \mathcal{L}[\,\overline{V}\,|c](t,a,z)}^{\text{standard HJB continuation value}} \quad \overbrace{-\theta(Z-\bar{Z})\Big(\overline{V}_Z + \tfrac{\partial r(g,Z)}{\partial Z}\overline{V}_r\Big)}^{\text{direct + indirect effect of drift } Z} + \overbrace{\frac{\widehat{\sigma}^2}{2}\Big(\overline{V}_{ZZ} + \big(\tfrac{\partial r(g,Z)}{\partial Z}\big)^2\overline{V}_{rr}\Big)}^{\text{direct + indirect effect of risk of } Z} \\
&\overbrace{+\ \widehat{\sigma}^2\ \tfrac{\partial r(g,Z)}{\partial Z}\,\overline{V}_{Z\,r}}^{\substack{\text{covariance btw agg income}\\ \text{and interest rate risk}}} \quad + \overbrace{\big(\tfrac{\partial \mathcal{C}(r,Z|g)}{\partial r}\big)^{-1}\int_{\tilde{a},\tilde{z}} c^\star(\tilde{a},\tilde{z},r,Z)\ \mathcal{L}^\star(g\,|\,c^\star)[\tilde{a},\tilde{z}]g(d\tilde{a},z)}^{\text{effect of the distribution on } r}
\end{aligned}
\tag{31}
$$

Note that in that case, subject to the approximation that agents value only depends on interest rate $V = \overline{V} = \overline{V}(a, z, r, Z)$, there is no additional approximation that needs to be made since the interest rate dynamics – at first and second order – is a sufficient statistics for the agents consumption and saving behavior. In particular, we can calculate perfectly the distributional dynamics:

$$\int_{\tilde{a},\tilde{z}} c^{\star}{\scriptstyle(\tilde{a},\tilde{z},r,Z)}\ \mathcal{L}^{\star}(g \,|\, c^{\star})[\tilde{a},\tilde{z}]g(d\tilde{a},\tilde{z}) = \int_{\tilde{a},\tilde{z}} \frac{c^{\star}{\scriptstyle(\tilde{a},\tilde{z},r,Z)}}{da} s(\tilde{a},\tilde{z},r,Z)\ g(d\tilde{a},\tilde{z})\ +\ \int_{\tilde{a},\tilde{z}} \lambda_j (c^{\star}{\scriptstyle(\tilde{a},\tilde{z}_{-j},\cdot)} - c^{\star}{\scriptstyle(\tilde{a},\tilde{z}_j,\cdot)})\ g(d\tilde{a},\tilde{z})$$

$$= \mathbb{C}\text{ov}\big(MPC(a,\cdot), s(a,\cdot)\big) + \mathbb{E}[\ \lambda_j \Delta_z c^{\star}(\cdot)]$$

This is the sum of two terms: a covariance between saving and the marginal propensity to consume out of wealth (or cash in hand), and second, a term representing the consumption jump caused by a discrete change in income.

D.5 Numerical methods

The previous master equation can be approximated using the numerical approach developed in the main text.

1. Master Equation with "projection": $V = \overline{V}(a, z, r, Z)$

    - Start from a guess for $\mathcal{C}^{(n)}(r, Z|g)$ and $\partial\mathcal{C}^{(n)}(r, Z|g)/\partial r$ and $\partial\mathcal{C}^{(n)}(r, Z|g)/\partial Z$, as well as the distributional term $\mathbb{C}\text{ov}\big(MPC(a,\cdot), s(a,\cdot)\big)$
    - Solve the Master HJB eq. (31), for $\overline{V}{\scriptstyle(a,z,r,Z)}$ using standard finite difference methods
    - Obtain the individual decisions $c^{\star}{\scriptstyle(a,z,r,Z)}$ and operator $\mathcal{A}[\overline{V}|c^{\star}](x)$ for $x = (a, z, r, Z)$

2. Kolmogorov forward for the global dynamical system $x = (a, z, r, Z)$

    - Solve for distribution $\widetilde{g}$ over all states (a, z, r, Z) for "free" with $\mathcal{A}^{*}[\widetilde{g}|c^{\star}]$
    - Using Radon-Nykodym / Bayes rules, obtain $g(a, z\big|r, Z)$
    - Update $\mathcal{C}^{(n+1)}(r, Z|g)$ and the other distributional terms, using $g = g(a, z\big|r, Z)$

3. Repeat until convergence of the dynamics $||\mathcal{C}^{(n+1)} - \mathcal{C}^{(n)}|| < \varepsilon$.