NBER WORKING PAPER SERIES

SOLVING HETEROGENEOUS AGENT MODELS WITH THE MASTER EQUATION

Adrien Bilal

Working Paper 31103
http://www.nber.org/papers/w31103

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2023, Revised October 2023

The first version of this paper was posted online, circulated and presented in 2021. I thank Rodrigo Adao, Fernando Alvarez, Anmol Bhandari, Pierre Cardaliaguet, Gabriel Chodorow-Reich, David Evans, Anders Humlum, Erik Hurst, Hugo Lhuillier, Greg Kaplan, Ben Moll, Esteban Rossi-Hansberg, Panagiotis Souganidis, Ludwig Straub, Gianluca Violante, Christian Wolf and Nathan Zorzi for valuable suggestions and comments. I thank Leonardo D'Amico, Santiago Franco and Shlok Goyal for superb research assistance. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Solving Heterogeneous Agent Models with the Master Equation
Adrien Bilal
NBER Working Paper No. 31103
April 2023, Revised October 2023
JEL No. C02,C6,E10

ABSTRACT

This paper proposes an analytic representation of perturbations in heterogeneous agent economies with aggregate shocks up to any order. Treating the underlying distribution as an explicit state variable, the value function defined on an infinite-dimensional state space summarizes the entire economy and satisfies the 'Master Equation' introduced in the mathematics mean field games literature. I show that analytic perturbations of the Master Equation deliver dramatic simplifications. The First-order Approximation to the Master Equation (FAME) reduces to a standard Bellman equation for the directional derivatives of the value function with respect to the distribution and aggregate shocks. The FAME has six main advantages: (i) finite dimension; (ii) closed-form mapping to steady-state objects; (iii) applicability when many distributional moments or prices enter individuals' decision such as dynamic trade, urban or job ladder settings; (iv) block-recursivity bypassing further fixed points; (v) characterization of stability and of the stochastic steady-state; (vi) fast implementation using standard numerical methods. I develop the Second-order Approximation to the Master Equation (SAME) and show that it shares these properties, making it suitable to handle nonlinearities, aggregate risk, and asset pricing. I illustrate the FAME and the SAME with two applications. First, I show that the cost of business cycles is several orders of magnitude larger in incomplete market models. Second, I show that dynamic spatial migration models display near linearity at conventional aggregate shocks.

Adrien Bilal
Department of Economics
Harvard University
1805 Cambridge Street
Cambridge, MA 02138
and NBER
adrienbilal@fas.harvard.edu

Introduction

A vibrant frontier in macroeconomics incorporates rich cross-sectional heterogeneity in dynamic general equilibrium. Recent numerical advances have remarkably accelerated the computation of impulse response functions up to first-order perturbations in leading incomplete credit market models (Ahn et al., 2018, Auclert et al., 2021). Yet, a conceptual framework rationalizing these numerical techniques and expanding their applicability has remained elusive. What is the relationship between economic fundamentals and equilibrium outcomes in perturbations of heterogeneous agent models? How to handle settings that fall beyond the natural scope of these methods, such as frictional labor markets or dynamic discrete choice? Are higher order perturbations that include nonlinearities and aggregate risk feasible?

In this paper, I propose answers to these questions using a new conceptual framework for perturbations in heterogeneous agent economies. I propose a representation of dynamic economies with heterogeneity that is analytic, low-dimensional, handles flexible general equilibrium feedbacks between the distribution and individual decisions, and applies systematically to perturbations of any order. These results rely on two key ideas. First, I use a state-space approach and treat the distribution of underlying heterogeneity as an explicit state variable in individual decisions. A single value function equation set on the space of distributions summarizes the equilibrium: the 'Master Equation.' Second, I take analytic—instead of numerical—perturbations of the Master Equation in the distribution and aggregate shocks and characterize the directional derivatives of the value function.

Specifically, the first core idea in this paper is to represent dynamic general equilibrium economies in fully recursive form. For concreteness, consider a standard incomplete market model as in Aiyagari (1994) and Krusell and Smith (1998). Households face uninsurable idiosyncratic labor productivity risk, and may borrow and save in a risk-free asset. A representative firm rents capital and hires labor. Abstract from aggregate shocks for now: the economy simply starts out of steady-state. Households' forward-looking consumption and savings decisions are fully determined by the sequence of future interest and wage rates. These prices in turn depend on the underlying distribution of asset holdings and idiosyncratic productivity through the firms' decision and market clearing. The distribution of assets and productivity evolves over time according to the optimal savings decisions of individuals. The classic difficulty in characterizing this economy is that individual decisions are forward-looking in time, while the evolution of the infinite-dimensional distribution is backward-looking in time. Prices are the fixed point of this forward-backward system that clear the capital and labor market.

I include the underlying distribution of heterogeneity as an explicit state variable in households' decision problem. Knowledge of the distribution fully characterizes prices. Households know the law of motion of the distribution. Thus, they forecast the future path of the distribution and hence prices. Households' decision problem then depends on the distribution just as on any other state variable. The only notable difference is that the distribution is an infinite-dimensional object, rather than a finite-dimensional state vector.

The resulting representation of the economy is the Master Equation. It was recently characterized in the mathematics mean field games literature by Cardaliaguet et al. (2019). It consists of a single Bellman equation that describes the entire behavior of a system of interacting agents. In the Krusell and Smith (1998) example, the Master Equation defines a value function that depends on a given household's idiosyncratic states—assets and productivity—as well as the underlying distribution of assets and productivity of all other households of the economy. The Master Equation is a Markovian representation of the economy because it includes as a state variable all the necessary information to forecast the evolution of the economy. It merges the fixed point on decisions, prices and the distribution into a single object.

The logic underpinning the Master Equation representation is more general than the Krusell and Smith (1998) example. At the same time, the analysis in Cardaliaguet et al. (2019) imposes restrictions that are at odds with most economic applications of interest. Therefore, I expand the scope of the Master Equation to encompass a large class of continuous-time dynamic general equilibrium economies that nest many applications of interest. In particular, I introduce the formalism of weak derivatives to handle mass points symmetrically to a smooth density.[1]

The second core idea is to simplify the Master Equation by focusing on local perturbations around a deterministic steady-state. Consider an impulse in the distribution, that moves the economy away from its steady-state. I explicitly perturb the Master Equation along any such distributional impulse. To do so, I make use of Fréchet derivatives, which are generalized derivatives in infinite-dimensional spaces. I preserve the full nonlinearity of individual decisions with respect to idiosyncratic states. In contrast to numerical techniques that first discretize and next linearize, I take an analytic perturbation first, before any computational discretization is applied. The use of continuous time streamlines the mapping between individual decisions and the evolution of the distribution. It also makes handling binding borrowing constraints easier: in continuous time, the first-order optimality condition continues to hold with equality, sidestepping Lagrange multipliers.

The First-order Approximation to the Master Equation (FAME) without aggregate shocks results in a Bellman equation with five key properties. First, it is low-dimensional. Its solution, the 'Impulse Value', consists of the directional derivatives of the value function with respect to the distribution. It encodes how individuals value changes in the distribution. In the FAME, the Impulse Value depends on only twice the number of idiosyncratic states, down from infinity in the fully nonlinear Master Equation. In the Krusell and Smith (1998) example, the Impulse Value has dimension four. This drastic dimension reduction is a feature of the local perturbation. To know their Impulse Value, individuals must know their own idiosyncratic states—for instance assets and productivity. Individuals must also know where the distributional impulse is happening—at another possible pair of assets and productivity. Thus, they must keep track of another set of idiosyncratic states that index which distributional impulse they

[1] Individual states follow controlled jump-diffusion processes that are flexible functionals of the underlying distribution to handle job search and dynamic discrete choice models. Idiosyncratic states face constraints, to handle occasionally binding borrowing constraints. Mass points develop in the distribution, consistently with borrowing constraints.

are contemplating.

Second, the FAME depends in closed form on known and interpretable steady-state objects through a systematic structure. The FAME depends on the steady-state law of motion of idiosyncratic states, such as assets and productivity. It also depends on the direct impact of distributional impulses on individual utility and transition probabilities between states. Finally, the FAME depends on the general equilibrium response of all individuals in the economy through its effect on the law of motion of the distribution. By virtue of the local perturbation, these objects all have explicit expressions that are linked together in equilibrium by the FAME.

Third, the FAME applies equally well to settings in which few or many prices summarize feedback between the general equilibrium and individual decisions. For instance, in dynamic spatial models, households must keep track of several prices per location to decide where to migrate. In search and matching models with job-to-job search, the entire distribution of wage offers matters directly for workers. Because the FAME is a perturbation with respect to the entire distribution, it applies directly to those examples. For instance, Bilal and Rossi-Hansberg (2023) use the FAME to evaluate the cost of climate change in a model of the U.S. economy disaggregated into over 3,000 counties.

Fourth, the FAME provides a block-recursive representation of equilibrium and impulse response functions. The FAME inherits the block-recursivity from the Master Equation, in that it merges the value function and the distributional fixed points. Once the Impulse Value is known, the evolution of the distribution over time is obtained without solving any additional fixed point.

Fifth, the FAME offers a streamlined and efficient implementation using standard finite difference methods. Building on the analytic representation of the the FAME, I show that it displays a specific separability structure because an individual's own idiosyncratic state and the distributional impulse propagate symmetrically but independently. Once discretized, the FAME takes the form of a modified Sylvester matrix equation for which standard routines exist. In the Krusell and Smith (1998) example, computation typically takes a tenth of a second and requires a couple dozen lines of code.

The FAME extends readily to the presence of aggregate shocks. A similar perturbation approach delivers two key insights. To first order, the Impulse Value splits into two distinct components. The first component is simply the Impulse Value from the deterministic FAME, the 'deterministic Impulse Value'. The second component is the 'stochastic Impulse Value'. The latter represents how individuals value an aggregate shock. It satisfies a similar FAME to the distributional Impulse Value and may be solved analogously. The economy remains block-recursive because the deterministic Impulse Value is independent from aggregate shocks. The linearized law of motion of the distribution evolution equation now features an additional component that represents the response of individual decisions to aggregate shocks.

The sixth key property of the FAME characterizes dynamics with aggregate shocks. The FAME provides, to the best of my knowledge, the first stability criteria and a description of the stochastic steady-state in heterogeneous agent economies. Leveraging the closed-form mapping between steady-state objects and the law of motion in the FAME, I show that dynamic stability and exponential convergence back to steady-state obtain when the steady-state transition probabilities satisfy either a mixing condition or a Lyapunov function condition. Crucially, if these conditions fail, checking the dominant eigenvalue in numerically discretized economies can be misleading about true convergence rates: the numerical dominant eigenvalue will converge to zero as the discretization becomes finer if the underlying law of motion exhibits a continuous spectrum that includes zero. The mixing or Lyapunov conditions rule out this possibility and ensure a spectral gap in the law of motion of the economy. Building on these stability conditions and the linearity of the law of motion in the FAME, I further characterize the conditional average of the distribution in the stochastic steady-state. It satisfies a linear equation that, after discretization, is a standard Sylvester matrix equation.

The Master Equation provides a systematic approach to perturbations of increasing order. Conceptually and practically, the Second-order Approximation to the Master Equation (SAME) is the same as the FAME. I show that the SAME defines a value function that depends again only on steady-state objects in closed form. Its solution now depends on three times the number of idiosyncratic states because pairwise impulses in the distribution matter to second order. I show how to compute the solution to the SAME using tensor Sylvester equations. I further characterize second moments of the stochastic steady-state distribution and provide a formula for welfare in the SAME. In sum, the SAME is well-suited for applications that focus on non-linearities, aggregate risk or asset pricing which requires second-order perturbations to depart from certainty equivalence. The Master Equation can handle third and higher-order perturbations as well, although I do not explicitly derive them in this paper.

I illustrate the Master Equation approach with two distinct applications. The first application shows how to use the FAME and the SAME in a workhorse incomplete market economy à la Aiyagari (1994) and Krusell and Smith (1998). I evaluate the cost of business cycles in a setup similar to the one in Krusell et al. (2009) that adds countercyclical income risk. By contrast however, I use a low-liquidity, high-Marginal Propensity to Consume (MPC) calibration that targets an average MPC of 0.2 and a volatility of aggregate consumption of 0.032 as in Lucas (1987). I find that the combination of countercyclical income risk and incomplete markets leads to an aggregate cost of business cycles of 2.3% of steady-state consumption, 23 times larger than Lucas (1987)'s seminal calculation. Crucially, losses are concentrated on the low-wealth and unemployed individuals, who can gain over 10% from the elimination of business cycles. Calculating the solution to the FAME and the SAME to reach these conclusions takes a couple seconds on a laptop.

The second application shows how the use the FAME and the SAME in a dynamic spatial model. I consider the United States (US) economy disaggregated into 381 Metropolitan Statistical Areas (MSAs). Individuals can migrate between MSAs and face bilateral migration costs and idiosyncratic preference shocks. Individuals spend their wage on a final good and housing which is locally supplied. Locations are differentially exposed to an aggregate productivity shock. This setting defines a heterogeneous agent economy in which a household's current location is an individual state variable, and the population distribution is an aggregate state variable. The FAME and the SAME take a couple seconds to solve on a laptop. After estimating the framework on US data, I show that the response of population and welfare across all locations is virtually identical in the FAME and the SAME for aggregate shocks up to 30%. This result indicates that first-order perturbations are sufficient for many applications of interest.

This paper relates to four strands of literature. First, I build on the mathematics mean field games literature and its Master Equation formulation in Cardaliaguet et al. (2019).[2] I complement this literature by proposing a flexible formulation of the Master Equation that is amenable to a wide class of economic applications, and by characterizing explicitly its first- and second-order perturbations.

Second, this paper relates to the set of papers that characterize impulse response functions analytically in specific heterogeneous agent models by studying their spectral properties (Gabaix et al., 2016, Alvarez and Lippi, 2021, Liu and Tsyvinski, 2020). I complement this literature by providing sufficient conditions for stability and the emergence of a spectral gap that can be checked in a wide class of economies.

Third, this paper connects to literature proposing computational methods for first-order perturbations of impulse response functions in heterogeneous agent economies in state space form (Reiter, 2009, Ahn et al., 2018). These methods first discretize, then linearize to first order, an economy with heterogeneity. They treat the resulting finite but high-dimensional system as a standard rational expectations system. By reversing the order—linearizing first, discretizing next—the FAME is the internally consistent foundation for this computational approach.[3] The FAME provides an economic interpretation of numerical output that may be otherwise difficult to see through and simplifies implementation. Crucially, the Master Equation approach also delivers a systematic approach to higher order perturbations such as the SAME.[4] Since this paper was first circulated in 2021, Bhandari et al. (2023) have developed complementary techniques to compute perturbations in discrete time heterogeneous agent economies.

Fourth, the Master Equation approach relates to numerical linearization techniques that leverage the sequence-space representation of the economy (Boppart et al., 2018, Auclert et al., 2021). These sequence-space approaches are designed for first-order perturbations, in contrast to the Master Equation approach that scales to any order. In a companion paper (Bilal, 2023) I derive and charaterize analytic sequence-space Jacobians in continuous time and describe their connection to the FAME.

Fifth, the Master Equation approach connects to numerical methods to solve heterogeneous agent

[2]Mean-field games were first introduced in sequential form by Lasry and Lions (2006) and Huang et al. (2006). The Master Equation was initially discussed in Lions (2011). See also Carmona and Delarue (2018a) and Carmona and Delarue (2018b) for the dual, probabilistic approach to mean-field games and the Master Equation.

[3]I discuss the connection in more detail in Section 1.7. Ahn et al. (2018) mention the Master Equation in their Appendix as a possible justification for numerical state-space approaches but do not establish the connection formally.

[4]The analytic nature of perturbations in the FAME and the SAME relate to the analytic perturbation approach in Bhandari et al. (2021). They also propose first and second-order perturbation of heterogeneous agent economies. However, they require that all shocks, both aggregate and idiosyncratic, are small enough. The FAME and the SAME instead preserve full nonlinearity with respect to idiosyncratic uncertainty by only requiring that aggregate shocks are small. Childers (2018) proposes a hybrid approach where some differentiation is analytic but does not recover smaller-dimensional Bellman equations as in the FAME.

models globally. Schaab (2021) also uses the Master Equation to propose a global, adaptative sparse grid strategy that builds on Brumm and Scheidegger (2017). Kahou et al. (2021), and Azinovic et al. (2022) use neural networks to obtain global solutions. Global methods do not rely on local perturbations, but can be delicate to implement and more time-consuming. This paper instead proposes an analytic perspective on local perturbations of the Master Equation.

The remainder of this paper is organized as follows. Section 1 presents the intuition behind the Master Equation approach in the context of the Krusell and Smith (1998) economy and a dynamic migration model. Section 2 defines the Master Equation for a general economy. Section 3 derives the FAME. Section 4 describes the SAME. Section 5 presents the two applications. The last section concludes. Proofs and additional details may be found in the Appendix.

1 Motivating examples

This section illustrates the Master Equation approach in two specific economies. I start with the Krusell and Smith (1998) economy as an expositional device. Therein, I show how to use the Master Equation approach to solve for impulse response functions, and derive connections with existing numerical methods. In the last part of this section, I provide another example in a dynamic discrete choice setting.

1.1 Setup

The setup follows closely the continuous-time version of the Krusell and Smith (1998) economy in Achdou et al. (2021). Time $t \geq 0$ is continuous and runs forever. There are no aggregate shocks for now, but the economy may start out of steady-state. Individuals are endowed with idiosyncratic time-varying productivity $y_t$which follows a stationary stochastic process. This process is independent across individuals and is defined by its generator $M(y)$—a functional operator that encodes conditional expectations under the income process. For instance, if productivity follows a diffusion, $dy_t = \mu(y_t)dt + \sigma(y_t)dW_t$, then the generator is the functional operator $M(y)[V] = \mu(y)V'(y) + \frac{\sigma(y)^2}{2}V''(y)$. Households are endowed with initial asset holdings a_0.

Households solve a standard income fluctuation problem by deciding how much to consume and save every period in a single risk-free asset a. For brevity, I denote by $x = (a, y)$ the pair of idiosyncratic states of households. The value function of households $V_t(x)$ satisfies the Hamilton-Jacobi-Bellman equation:[5]

$$\rho V_t(x) = \max_{c \geq 0} u(c) + L_t(x, c)[V_t] + \frac{\partial V_t}{\partial t}(x) \quad , \quad L_t(x, c)[V] \equiv (r_t a + w_t y - c)\frac{\partial V}{\partial a}(x) + M(y)[V]. \tag{1}$$

$L_t(x, c)[V]$ is the continuation value from changes in assets a_t that evolve according to the budget constraint $da_t = (r_t a_t + w_t y_t - c_t)dt$, and changes in productivity that follows its exogenous stochastic

[5]The value function is restricted to have at most linear growth as assets and income approach infinity. This restriction is the recursive analogue to the No-Ponzi condition in the sequential formulation of problem (1).

process. The functional operator $L_t(x,c)$ is the generator of the stochastic process for the pair of idiosyncratic states of households $x = (a, y)$. Denote by $\hat{c}_t(x)$ the optimal consumption decision of households.

A representative firm operates a production technology $Y = \bar{Z}K^{\alpha}N^{1-\alpha}$ and rents assets from households at the interest rate r_t. The firm transforms assets into productive capital K. For expositional simplicity, capital does not depreciate. The optimality conditions of the firm together with market clearing imply that the real interest rate r_t and the wage rate w_t clear the capital and labor market:

$$r_t = \alpha \left(\frac{\iint y g_t(x) dx}{\iint a g_t(x) dx} \right)^{1-\alpha} \equiv \mathcal{R}(g_t) \quad ; \quad w_t = (1-\alpha) \left(\frac{\iint a g_t(x) dx}{\iint y g_t(x) dx} \right)^{\alpha} \equiv \mathcal{W}(g_t), \tag{2}$$

where $g_t(x)$ is the probability distribution function of households over assets and income at calendar time t. The functionals $\mathcal{R}, \mathcal{W}$ capture how the interest and wage rates depend on the current distribution of households.

The distribution $g_t(x)$ evolves over time according to its law of motion, the Kolmogorov Forward equation:

$$\frac{\partial g_t}{\partial t}(x) \quad = \quad L_t^*(x, \hat{c}_t(x))[g] \quad , \quad L_t^*(x, \hat{c}_t(x))[g] \equiv -\frac{\partial}{\partial a}\Big(s_t(x) g_t(x) \Big) + M^*(y)[g_t], \tag{3}$$

and $s_t(x) = r_t a + w_t y - \hat{c}_t(x)$ denotes the equilibrium savings rate. M^* is the adjoint of the operator M, which is the functional equivalent of the matrix transpose. Similarly, the operator L^* is the adjoint of the operator L.

1.2 The Master Equation

The Master Equation approach considers the individual decision problem (1) in state-space, or recursive, form. This approach is distinct from the sequential view of (1), that takes the path of interest and wage rates $(r_t, w_t), t \geq 0$ as an input.

Viewing the individual decision problem (1) in the state space requires defining the value function on the relevant state space: the space of distributions g. While the distribution is infinite-dimensional, it may be viewed as a larger analog of any other finite-dimensional state variable. Just as with any other state variable, including the distribution g as an explicit state variable requires knowing its evolution over time. Crucially, the Kolmogorov Forward equation (3) encodes precisely that law of motion.

I build towards the Master Equation in three steps. The first step is to recognize how the value of a household depends on equilibrium objects. In this example, the value depends only on the interest and wage rates r_t, w_t.

The second step is to express prices and any other equilibrium objects that may enter the value function, as functionals of the underlying primitive distributions that define the relevant aggregate state. In this example, interest and wage rates $r_t = \mathcal{R}(g_t), w_t = \mathcal{W}(g_t)$ depend on the ratio of marginals of the joint distribution of assets and wealth from the capital and labor market clearing conditions (2). Thus, the capital and labor market clearing conditions (2) provides the required map directly through the functionals $\mathcal{R}, \mathcal{W}$. Substituting into the value of households, I obtain

$$\rho V_t(a,y) \quad = \quad \max_c u(c) + L(x,c,g_t)[V_t] + \frac{\partial V}{\partial t}(a,y), \tag{4}$$

where $L(x,c,g_t)[V] = (\mathcal{R}(g_t)a + \mathcal{W}(g_t)y - c)\frac{\partial V}{\partial a}(a,y) + M(y)[V]$. So far, the transformation of the Bellman equation (1) into (4) is mostly notational: I have substituted the time-dependent price sequence with the distribution-dependent price functionals.

The cornerstone of the Master Equation approach lies in the third step: replacing the time dependence of the value function itself by an explicit dependence on the distribution g. This substitution amounts to a change of variables: $V_t(x) \equiv V(x, g_t)$. Using this identity, the decision problem becomes fully recursive by re-expressing the time derivative in the Bellman equation (4). The first step is to recognize that, by the chain rule:

$$\frac{\partial V_t}{\partial t}(a,y) = \int \frac{\partial V}{\partial g}(x,x',g_t)\frac{\partial g_t}{\partial t}(x')dx' \equiv \left\langle \frac{\partial V}{\partial g}(x,g_t), \frac{\partial g_t}{\partial t} \right\rangle. \tag{5}$$

The chain rule in (5) is one that applies in infinite-dimensional spaces. It involves slightly more notation than the usual chain rule in finite dimension, but follows the exact same logic.

The brackets $\langle \cdot, \cdot \rangle$ denote an inner product in the appropriate functional space. In this application, the inner product turns out to be $\langle \varphi, \psi \rangle = \int \varphi(a,y)\psi(a,y)dady$ on the Hilbert space of square integrable functions. This inner product is the natural generalization of the Euclidean inner product $\langle \varphi, \psi \rangle = \sum_{n=1}^{N} \varphi_n \psi_n$ when dealing with functions rather than vectors.

The derivative of the value with respect to the distribution, $\frac{\partial V}{\partial g}$, must be understood in an appropriate space for the distribution g. The relevant notion in most economic applications turns out to be that of Fréchet derivative, the natural generalization of derivatives in finite-dimensional spaces to infinite-dimensional Hilbert spaces.

To gain intuition, suppose temporarily that the possible set of assets and wages was discrete and finite, indexed by n. The value function would become a vector $(V_n)_{n=1}^N$, and the derivative $\frac{\partial V}{\partial g}$ would simply represent the gradient of the value vector with respect to the mass at each ones of these points. Namely, one could write $g \equiv (g_n)_{n=1}^N$, and thus $\frac{\partial V}{\partial g} = \left(\frac{\partial V_n}{\partial g_1}, ..., \frac{\partial V_n}{\partial g_N}\right)_{n=1}^N$.

The Fréchet derivative extends the notion of gradient to the case when the underlying idiosyncratic state space is continuous rather than discrete. In particular, the Fréchet derivative $\frac{\partial V}{\partial g}(x,x',g)$ is itself a function of the direction in which the derivative is taken, x'—just as with a finite dimensional gradient.[6]

The second step to remove the time derivative is to recognize that the change in the distribution, $\frac{\partial g_t}{\partial t}$ is precisely given by the law of motion (3). Because prices are functionals of the distribution g,

[6] I use this notation to empahsize this functional dependence, rather than the notation $\frac{\partial V}{\partial g(x')}(x,g)$ that would be more directly analogous to the discrete case. The notation $\frac{\partial V}{\partial g}(x,g)$ in the inner product implicitly omits the dependence on x', but I also write explicitly $\frac{\partial V}{\partial g}(x,x',g)$ when needed.

the savings rate writes $s_t(x) \equiv s(x, g_t) = \mathcal{R}(g_t)a + w - \hat{c}(x, g_t)$. Hence, the evolution equation (3) also only depends on the distribution g. Thus, I change variables in the functional operator that encodes the evolution of the distribution $L_t^*(x, \hat{c}_t(x)) \equiv L^*(x, g_t)$ similarly to the value function. The dependence on the distribution g is both explicit through the interest rate, and implicit through the optimal consumption decision:[7]

$$\frac{\partial g_t}{\partial t}(x) = L^*(x, g_t)[g_t]. \tag{6}$$

Combining the chain rule (5) with the law of motion of the distribution (6), I finally obtain

$$\frac{\partial V_t}{\partial t}(x) = \int \frac{\partial V}{\partial g}(x, x', g_t) L^*(x', g_t)[g_t] dx'.$$

Combining the previous observations, I rewrite the Bellman equation (4) as

$$\rho V(x, g) \quad = \quad \max_c u(c) + L(x, c, g_t)[V] + \int \frac{\partial V}{\partial g}(x, x', g_t) L^*(x', g_t)[g_t] dx'. \tag{7}$$

Equation (7) is the Master Equation. Arriving at the representation (7) has required many definitions, but the payoff is substantial: the Master Equation (7) is a state-space—or recursive—representation of the household problem.

Inspection of the Master Equation (7) reveals that there is no need to keep track of a separate law of motion for the distribution. This law of motion has precisely been incorporated into the value function through its last term $\int \frac{\partial V}{\partial g}(x, x', g_t) L^*(x', g_t)[g_t] dx'$. As a result, the Master Equation (7) is the only equation that needs be solved to characterize the equilibrium. This property has lead it to be called the 'Master Equation' in the mathematics mean field games literature. In practice, the representation of the equilibrium is now block-recursive: the evolution of the distribution along any particular equilibrium realization follows ex-post, once the solution to the Master Equation is known.

The recursive nature of the Master Equation allows to leverage standard recursive methods to characterize and compute the solution to (7). However, the fully nonlinear Master Equation (7) is defined on a infinite-dimensional state space that includes the distribution g. Therefore, it remains difficult to handle nonlinearly in practice.

To overcome this practical difficulty, I combine the Master Equation (7) with local perturbation methods. It turns out that this combination provides a powerful closed form characterization of the linearized Master Equation, and drastically reduces the dimensionality of the problem.

1.3 The FAME

I start from a locally isolated steady-state of the economy. It is given by a steady-state value function $V^{SS}(x)$ and a steady-state distribution $g^{SS}(x)$. I then consider the First-order Approximation to the Master Equation (FAME) around the steady-state distribution g^{SS}. Specifically, I only require that deviations in the distribution—that I denote by $h = g - g^{SS}$ and call distributional impulses—are small

[7] The dependence of the coefficients of the operator L^* on the distribution g_t runs through the parenthesis: $g_t \mapsto L^*(x, g_t)[\cdot]$. The action of the operator holding the coefficients fixed runs through the square brackets: $g_t \mapsto L^*(x, \cdot)[g_t]$.

in the mean squared error norm. I do not require that idiosyncratic uncertainty is small. Instead, the present approximation preserves the full nonlinearity of decisions with respect to individual states.[8] The main innovation in this paper is to take the perturbation analytically instead of numerically.

The critical observation in the FAME is that the value function $V(x, g)$ then becomes, to a first order, a linear functional of the distributional impulse h. Namely, to first order:

$$V(x, g^{SS} + h) = V^{SS}(x) + \int v(x, x')h(x')dx'. \tag{8}$$

The function $v(x, x')$ encodes how the value function evaluated at the point $x = (a, y)$ responds to small impulses in the distribution around the steady-state. To first order, only the effect of the impulse direction by direction need be considered, and the expansion is additive in the impulse h. The pairwise effects of the impulses are second order and thus drop out to first order.

I call the function $v(x, x')$ the 'determinstic Impulse Value Function' or simply the Impulse Value. This terminology is motivated by the observation that $v(x, x')$ exactly encodes how the value function responds to a small impulse h in the underlying distribution relative to steady-state. It is the general equilibrium effect of adding one household at x' on the value of a household at x.

By construction, the Impulse Value coincides with the steady-state directional derivative: $v(x, x') = \frac{\partial V}{\partial g}(x, x', g^{SS})$. To build intuition, the analogue of equation (8) with a finite state space would be $V_n(g) = V_n^{SS} + \sum_{k=1}^{N} v_{nk}h_k$ with $v_{nk} = \frac{\partial V_n}{\partial g_k}(g^{SS})$. The integral in equation (8) generalizes this notation to settings with a continuous state space.

The goal of the FAME is to derive restrictions that determine the Impulse Value. To that end, I follow a similar strategy to perturbation methods in representative agent economies such as the Real Business Cycle (RBC) model. I substitute the definition of the Impulse Value (8) into the nonlinear Master Equation (7). I then take a first-order approximation in the distributional impulse h. Since the Master Equation must hold for all h, the final step uses the method of undetermined coefficients. When linearizing the RBC model, there is a finite number of coefficients to identify—for instance one coefficient for how the value function depends on impulses in the aggregate capital stock. With heterogeneity, the only difference is that the 'coefficients' are themselves functions, such as the Impulse Value $v(x, x')$ itself.

The calculation described above leads to the FAME:

$$\rho v(x, x') = \underbrace{u'(c^{SS}(x))D(x, x')}_{\text{Direct price impact}} + \underbrace{\mathcal{L}(x)[v(\cdot, x')]}_{\substack{\text{Partial equilibrium:}\\ \text{continuation value}\\ \text{from shocks to } x}} + \underbrace{\mathcal{L}(x')[v(x, \cdot)]}_{\substack{\text{General equilibrium:}\\ \text{continuation value}\\ \text{from propagation}\\ \text{of impulse at } x'}} + \underbrace{\int v(x, x'')G(x'', x', v)dx''}_{\substack{\text{General equilibrium:}\\ \text{weighted average of changes in}\\ \text{savings rates of other households } x''\\ \text{in response to impulse at } x'}}, \tag{9}$$

[8] My approach is similar in spirit to Reiter (2009), Ahn et al. (2018) or Auclert et al. (2021). See Preston and Roca (2007), Mertens and Judd (2017), and Bhandari et al. (2021) for perturbation methods that additionally require that idiosyncratic shocks are small.

where

$$
\begin{aligned}
D(x,x') &= \big(\mathcal{R}_0 a' + \mathcal{R}_1 y'\big)a + \big(\mathcal{W}_0 a' + \mathcal{W}_1 y'\big)y \quad , \quad \mathcal{L}(x) = L(x, c^{SS}(x), g^{SS}) \\
G(x'',x',v) &= -\frac{\partial}{\partial a''}\left(g^{SS}(x'')\Big(D(x'',x') - \frac{1}{u''(c^{SS}(x''))}\frac{\partial v}{\partial a''}(x'',x')\Big)\right)
\end{aligned}
$$

and $\mathcal{R}_0, \mathcal{R}_1, \mathcal{W}_0, \mathcal{W}_1$ are constants that depend only the steady-state distribution and are given in Appendix A.1.

Equation (9) is the FAME. Its right-hand-side has four components. Each one encodes a particular force that affects how the value of household $x = (a, y)$ changes in equilibrium when an additional household enters the economy at $x' = (a', y')$.

The first component in the FAME is the direct price impact. When the distribution changes, prices also change. The movement in prices affects households' disposable income as encoded in the price impact function $D(x, x')$. The price impact function D depends linearly on the household's state: a household with more assets benefits more from a rise in the interest rate. The price impact function D also depends linearly in the point x' in the state space where the distributional impulse is occurring—where household x is contemplating an excess mass of other households x'. A given distributional impulse at x' affects the interest rate more if a' is high or if y' is high. Therefore, the function D is larger at larger a' and y'. A similar logic underlies the impact of the distributional impulse through the wage rate. The impact of the distributional impulse on households' consumption drops out due to the envelope condition—the first-order optimality condition always holds in continuous time. By virtue of the local perturbation, households then convert changes in disposable income into utils using their steady-state marginal utility of consumption.

The second component in the FAME encodes a partial equilibrium force, similar to households' continuation value in (1). Even out of steady-state, households form expectations about their own assets and labor market productivity. Crucially, by virtue of the first-order perturbation, household need only evaluate those expectations using steady-state prices and consumption policy functions. Thus, they use the steady-state continuation value operator $\mathcal{L}$ that involves only steady-state transition probabilities. This operator acts on the first argument x of the Impulse Value, that represents the dependence of their value on their own own state variable.

The third component in the FAME represents a first general equilibrium force. When contemplating the effect of an additional household at x' on the economy, households at x expect this additional household at x' to behave just as any other household. Household x' consumes, saves and receives labor market shocks. Thus, the additional household at x' travels through the state space. Keeping track of where they go matters to project the economy forward in time and evaluate what tomorrow's distribution will be. The FAME shows that this expectation is summarized by the steady-state expectation operator $\mathcal{L}$. Once more, because of the local perturbation, only steady-state transition probabilities matter to first order. Crucially, the steady-state operator $\mathcal{L}$ acts on the second argument x' of the Impulse Value value, that represents the effect of an additional household at x' on the value of household x.

The fourth component in the FAME encodes a second general equilibrium force. It represents how household x values changes in the law of motion of the distribution that arise because of an additional household at x'. Why would the law of motion change? An additional household at x' affects prices. Because prices change, all households in the economy change their savings behavior—represented by the integral over x''. This change in savings behavior affects the law of motion of the distribution to first order, and thus affects any given household x after weighting by the steady-state distrbution g^{SS} and converting to utils using the Impulse Value $v(x, x'')$. The change in savings rates of any other household x'' is then given by the innermost bracket. It involves the price impact function D net of the first-order change in consumption $\frac{\partial_{a''} v(x'',x')}{u''(c^{SS}(x''))}$. This expression for the consumption response follows from linearizing the first-order condition for consumption. It is a 'distributional Marginal Propensity to Consume' (dMPC): it represents how consumption changes in response to a distributional impulse.

Despite its notational complexity, the FAME is in fact remarkably simple. It has four key features, which Section 3 shows hold much more broadly than in the present example.

1.4 Properties of the FAME

The first property of the FAME is that it is a standard Bellman equation in finite dimension. The dimensionality of the Impulse Value is simply twice that of the original problem, instead of being infinite-dimensional like the nonlinear Master Equation (7). This drastic simplification stems from the local perturbation. Households located at x in the state space need only consider isolated impulses at any other possible x' in the distribution, since any pairwise impulses would lead to a second-order deviation in the value function.

The second property of the FAME is that it depends on the steady-state in closed form. By virtue of the analytic nature of the perturbation, all the objects entering in the FAME are explicitly linked to the steady-state. Hence, once the nonlinear steady-state of the model is known, no additional calculation is needed to write down and solve the FAME.

The third property of the FAME is block-recursivity, that it inherits from the Master Equation. The FAME is the only fixed point that must be solved to know individual behavior along any impulse response. There is no additional price or distributional fixed point to solve because such fixed points have already been embedded into the Master Equation. The FAME uncovers that this joint fixed point has a simple structure that may be solved efficiently.

In fact, once the Impulse Value is known, it is straightforward to apply a similar perturbation argument to the law of motion of the distribution (3).[9] To first order,

$$\frac{\partial h_t}{\partial t}(x) = \mathcal{L}^*(x)[h_t] + \mathcal{G}(x)[h_t] \quad , \quad \mathcal{G}(x)[h] \equiv \int G(x, x', v) h(x') dx'. \tag{10}$$

Equation (10) encodes the time evolution of any impulse h in the distribution over time. Its two

[9] In fact, equation (10) is already known. It follows immediately from the linearization of the last component in (7).

components represent distinct forces that mirror those that the FAME (9) uncovered.

The first term $\mathcal{L}^*[h]$ represents the partial equilibrium evolution of the impulse h. This partial equilibrium evolution is computed holding transition probabilities fixed at their steady-state values.

The second term $\mathcal{G}[h]$ represents the general equilibrium response of the economy. It encodes changes in savings behavior of all households in response to the distributional impulse. These changes in behavior are embedded in the price impact and dMPCs that enter the definition of the kernel G. The kernel G coincides with the expression in the general equilibrium component of the FAME (9) because both represent changes in savings rates.

The fourth property of the FAME is that it lends itself to an efficient numerical solution method. Due to its standard Bellman equation structure, the FAME can be solved efficiently using standard finite dimension schemes. Inspection of equation (9) reveals that a discretized version of the Impulse Value, $\mathrm{v}_{ij} = v(x_i, x_j)$ on a grid $(x_i)_i$ satisfies the nonlinear Sylvester matrix equation

$$\rho \mathrm{v} = \mathrm{D} + \mathrm{L}\mathrm{v} + \mathrm{v}\mathrm{L}^T + \mathrm{v}\mathrm{G}\mathrm{v},$$

where D is a matrix that discretizes $u'(c^{SS})D$, L is a matrix that discretizes the steady-state transition operator $\mathcal{L}$, L^T denotes the transpose of L, and G is a matrix that discretizes the operator $\mathcal{G}$. This Sylvester equation can be readily solved using standard routines and an iterative scheme. I describe this solution method in more detail in Section 3.7.

1.5 Aggregate shocks with the FAME

So far, there were no aggregate shocks in order to focus on the role of the distribution. I now introduce aggregate shocks in the economy. Aggregate productivity Z is no longer constant, but fluctuates over time: $Z_t = \bar{Z}e^{\varepsilon z_t}$. Log productivity follows a continuous-time AR(1) process: $dz_t = -\mu z_t dt + dW_t$ with associated generator $\mathcal{A}(z) = -\theta z \partial_z + \frac{1}{2}\partial_{zz}$. The parameter $\varepsilon > 0$ represents the overall scale of aggregate shocks.

The Master Equation (7) enriched with aggregate shocks becomes

$$\rho V(x,z,g) = \max_c u(c) + L(x,c,\varepsilon z,g)[V] + \int \frac{\partial V}{\partial g}(x,x',\varepsilon z,g) L^*(x',\varepsilon z,g)[g]dx' + \mathcal{A}(z)[V]. \tag{11}$$

Relative to the Master Equation without aggregate shocks (7), the Master Equation with aggregate shocks (11) now depends on aggregate productivity εz_t. This dependence is explicit in the operator L through the price functionals that now depend on aggregate productivity. In addition, households expect aggregate productivity to fluctuate over time. These expectations result in the additional continuation value $\mathcal{A}(z)[V]$. It is also possible to consider deterministic transitional dynamics in response to a one-time shock. In that case, one needs simply to keep track of time t instead of the aggregate productivity state z.[10]

[10] In that case, the Master Equation approach simply removes the contribution to the time derivative in the individual decision problem coming from the distribution g_t, but keeps the time derivative coming from the aggregate shock.

The law of motion of the distribution (3) also extends with aggregate shocks:

$$dg_t(x) = L^*(x, \varepsilon z_t, g_t)[g_t]\ dt. \tag{12}$$

The law of motion in (12) is now a stochastic version of the Kolmogorov forward equation (3).[11]

With the Master Equation (11) at hand, I extend the perturbation argument in the deterministic FAME to aggregate shocks. There are now two objects that are small. The first one is, as before, the deviation in the distribution $g - g^{SS}$. The second object is the scale of aggregate shocks ε. The natural benchmark is that both objects are of the same scale, when $\varepsilon h \equiv g - g^{SS}$ has scale ε.[12]

When ε is small, the first-order approximation to the value function is:

$$V(x, z, g^{SS} + \varepsilon h) = V^{SS}(x) + \varepsilon \left\{ \int v(x, x')h(x')dx' + \omega(x, z) \right\}. \tag{13}$$

ω is the 'stochastic Impulse Value': the direct effect of aggregate shocks on the value function.[13]

Just as in any Taylor expansion, the first-order approximation is additive in the response to a distributional impulse h and the aggregate shock z. Any pairwise perturbation involving an impulse in the aggregate shock together with an impulse in the distribution h is of scale ε^2, and thus second-order.

The separability in equation (13) is critical and represents certainty equivalence. It implies that the deterministic Impulse Value $v(x, x')$ with respect to the distributional impulse h satisfies the deterministic FAME (9) evaluated at $Z = \bar{Z}$. In particular, the deterministic Impulse Value x can be solved for independently from aggregate shocks. This observation mirrors the linearization of representative agent models such as the RBC model. There too, the deterministic component in the value function is independent from the stochastic component.

I identify coefficients on the aggregate shock after substituting the first-order expansion (13) into the nonlinear Master Equation (7), similarly to the derivation of deterministic FAME (9). The stochastic Impulse Value $\omega(x, z)$ then satisfies the stochastic FAME:

$$\rho\omega(x, z) = \underbrace{z\Omega(x)u'(c^{SS}(x))}_{\text{Direct price impact}} + \underbrace{\mathcal{L}(x)[\omega(\cdot, z)]}_{\substack{\text{Partial equilibrium:}\\ \text{continuation value}\\ \text{from shocks to } x}} + \underbrace{\mathcal{A}(z)[\omega(x, \cdot)]}_{\substack{\text{Continuation value from}\\ \text{aggregate shocks } z}} + \underbrace{\int v(x, x'')S(x'', z, \omega)dx''}_{\substack{\text{General equilibrium:}\\ \text{weighted average of changes in}\\ \text{savings rates of other households } x''\\ \text{in response to aggregate shock } z}}, \tag{14}$$

[11]It is a stochastic partial differential equation. The coefficients of the operator L^* change stochastically following shocks to aggregate productivity z_t, and thus define a stochastic partial differential equation. The logic underlying the law of motion (12) is entirely similar to that of finite-dimensional evolution equations. For instance, the law of motion for assets for employed individuals is a unidimensional stochastic differential equation $da_t = s_t(x_t)dt$. The law of motion (12) is an analogue, but in infinite dimension. The only addition is that interactions between the entries of the infinite-dimensional vector $g_t(x)$ that are relevant for its time dynamics are picked up by cross-sectional derivatives. The Master Equation (11) is non-stochastic, while the SPDE (12) is, because the Master Equation (11) conditions on the current value of the aggregate shock z, while the SPDE (12) takes the sequence of realized aggregate shocks z_t as given.

[12]In the stochastic steady-state, provided it is stable, g_t will remain in a neighborhood of g^{SS} of typical size ε because the typical size of aggregate shocks is ε.

[13]Depending on the stochastic process for aggregate productivity z, the stochastic Impulse Value ω need not be linear in z. In this section's example, it turns out that it can be easily proven because the productivity process is an unrestricted diffusion process. However, when there are reflecting boundaries or non-symmetric jump terms in the productivity process, ω is no longer linear in general. By contrast, the process for the distribution is always an unrestricted diffusion process, and thus the distribution always enters as a linear functional.

where $\Omega(x) \equiv \mathcal{R}_2 a + \mathcal{W}_2 y$, $\mathcal{R}_2, \mathcal{W}_2$ are constants that depend only on g^{SS} and are given in Appendix A.1, and

$$S(x,z) = -\frac{\partial}{\partial a}\left(g^{SS}(x)\left(\Omega(x)z - \frac{1}{u''(c^{SS}(x))}\frac{\partial \omega}{\partial a}(x,z)\right)\right).$$

The structure of the stochastic FAME (14) mirrors that of the deterministic FAME (9). The main difference is simply the expression for the direct price impact of aggregate shocks Ω. As with the deterministic FAME, all the objects that enter into the stochastic FAME (14) are evaluated at the deterministic steady-state. They are thus immediately known given given the steady-state.

As with the deterministic FAME, the stochastic FAME (14) is a standard Bellman equation that may again be solved with standard methods. A crucial property is that the deterministic Impulse Value v is known by the time one solves the stochastic FAME. Thus, inspection of (14) reveals that the stochastic FAME is a linear equation in ω. By contrast, the deterministic FAME (9) is a quadratic equation in v.[14]

With both Impulse Values v, ω at hand, I turn back to the law of motion of the distribution and obtain the evolution of the impulse h_t in the distribution up to a first order:

$$dh_t(x) \quad = \quad \left\{\mathcal{L}^*(x)[h_t] + \mathcal{G}(x)[h_t] + S(x,z_t)\right\} dt. \tag{15}$$

Equation (15) is the linearized version of the SPDE in (12). Iterating forward on (15) for a given sequence of aggregate shocks z_t then delivers any desired impulse response function.

1.6 Dynamic discrete choice

I now provide a second example of the FAME. I illustrate its flexibility with a dynamic discrete choice setting in which there are multiple prices per locations. This example also clarifies the link between perturbations with respect to a continuous distribution and perturbations with respect to a finite-dimensional vector. For concreteness, I interpret the framework as a dynamic location choice setting with migration, but the framework may be more broadly interpreted as an industry, occupation or product choice problem.

A dynamic migration setup is a natural environment to use the FAME because there are typically several prices per location. Thus, the state-space approach at the heart of the Master Equation is well-suited to handle the complexity of such a framework.[15]

Consider a unit mass of individuals who choose in which location $i \in \{1, ..., I\}$ to live. Individuals have flow CRRA preferences over a consumption index $C(c,h)$: $u(C) = \frac{C^{1-\gamma}-1}{1-\gamma}$, where γ denotes relative risk aversion. The consumption index is a Cobb-Douglas aggregator of a freely traded final good c, used as the numeraire, and of housing h with share β: $C(c,h) = (c/(1-\beta))^{1-\beta}(h/\beta)^{\beta}$.

[14] This observation mirrors the linearization of the RBC model, in which the deterministic component solves a quadratic scalar equation, while the stochastic component solves a linear equation given the solution to the deterministic component.

[15] By contrast, sequence space methods that linearize with respect to a sequence of prices require higher-dimensional calculations. If there are I locations, n prices per location and T discretized time periods, the FAME requires solving for a $I \times I$ matrix, while sequence-space Jacobians have dimension $(nIT) \times (nIT)$.

Individuals discount the future at rate ρ. In every location, individuals work for a unit measure of identical firms that produce the final good using labor and are subject to decreasing returns to scale.[16] The local wage is denoted w_{it}, and the housing rental rate r_{it}. Individuals consume their income each period and face a flow budget constraint $c + r_{it}h = w_{it}$.

Individuals are allowed to move at rate μ, in which case they draw extreme-value distributed idiosyncratic preference shocks for potential destinations, with dispersion parameter ν. If they move, they pay a bilateral moving cost τ_{ij}. Locations are endowed with a local productivity Z_{it}, a fixed supply of land H_i whose rents are paid to absentee landlords, and a housing supply curve $H_{it} = H_i r_{it}^{\eta}$. Output in location i is thus $Y_{it} = Z_{it} N_{it}^{1-\alpha}$.

There is a common aggregate productivity shock z_t which follows a continuous-time AR(1) process in logs: $dz_t = -\theta z_t dt + \sigma dW_t$, where dW_t is a Brownian motion. Denote $\mathcal{A}(z) = -\theta z \partial_z + \frac{\sigma^2}{2}\partial_{zz}$ the associated generator. Each location is exposed to this aggregate shock with a location-specific slope χ_i: $Z_{it} = Z_i e^{\varepsilon \chi_i z_t}$.[17]

Maximizing out the housing and optimal location choices, and clearing labor and housing markets, individuals solve

$$\rho V_{it} = U_i(z_{it}, N_{it}) + L_i[V] + \mathbb{E}\left[\frac{dV_{it}}{dt}\right], \quad L_i[V] \equiv \mu \left\{ \frac{1}{\nu} \log \left(\sum_j e^{\nu(V_j - \tau_{ij})} \right) - V_i \right\} \tag{16}$$

I provide details for all derivations in Appendix D.1. The flow payoff is $U_i(z_{it}, N_{it}) = C_{0i} N_{it}^{-\xi}$, where $\xi = \alpha + \frac{\beta(1-\alpha)}{1+\eta}$, and C_{0i} is a location-specific constant. The expectation operator $L_i[V]$ is nonlinear because of the presence of idiosyncratic taste shocks. Crucially, its action on a small perturbation of the value dV still coincides with the adjoint (tranpose) of the operator $\mathcal{L}_i^*$: $L_i[V^{SS} + dV] = L_i[V^{SS}] + \mathcal{L}_i(V^{SS})[dV]$. Location decisions are given by the conditional choice probabilities

$$\pi_{ijt}(V) = \frac{e^{\nu(V_{it} - \tau_{ij})}}{\sum_k e^{\nu(V_{kt} - \tau_{ik})}}. \tag{17}$$

The population distribution evolves according to the law of motion:

$$\frac{\partial N_{jt}}{\partial t} = \mu \left(\sum_k \pi_{kjt}(V) N_{kt} - N_{jt} \right) \equiv \mathcal{L}_i^*(V)[N_t]. \tag{18}$$

In this discrete choice setting, the operator $\mathcal{L}_i^*(V)$ is simply the matrix $\mu(\pi_t(V)^T - \mathrm{Id})$. The Master Equation writes

$$\rho V_i(z, N) = U_i(z, N) + L_i[V] + \sum_j \frac{\partial V_i}{\partial N_j} \mathcal{L}_j^*(V)[N_t] + \mathcal{A}(z)[V_i] \tag{19}$$

[16] Profits are distributed to absentee owners.

[17] The block-recursivity and separability properties of the FAME allow to easily consider shocks that are location-specific: $Z_{it} = Z_i e^{\varepsilon z_{it}}$, where each z_{it} follows an AR(1) process.

I seek a first-order solution to the Master Equation of the form

$$V_i(z,N) = V_i^{SS} + \varepsilon \left\{ \sum_j v_{ij} n_j + \omega_i z \right\}.$$

In Appendix D.2, I show that a similar substitution to the Krusell and Smith (1998) example delivers the deterministic FAME, a matrix equation for v:

$$\rho v = -\bar{v} + Mv + vM^T + vGv, \tag{20}$$

where $\bar{v} = \xi \mathbf{diag}\left(u'(C_i^{SS})C_i^{SS}/N_i^{SS}\right)$, $M = \mathcal{L}(V^{SS}) = \mu(m - \mathrm{Id})$ where m denotes the steady-state matrix of migration shares, and $G = \mu\nu\Big(\mathbf{diag}(N^{SS}) - m^T \mathbf{diag}(N^{SS})m\Big)$.

The stochastic FAME takes the form of a vector-valued equation:

$$(\rho + \theta)\omega = \varpi + (M + vG)\omega, \tag{21}$$

where $\varpi_i = \zeta \left(u'(C_i^{SS})C_i^{SS}\right)\chi_i$, $\zeta = \frac{1+\eta-\beta}{1+\eta}$.[18]

Finally, the linearized law of motion of the distribution satisfies:

$$\frac{\partial n_t}{\partial t} = M^T n_t + G(n_t + \omega z_t). \tag{22}$$

The combination of the stochastic FAMEs (21) and the law of motion (22) reveal that transitional dynamics in this economy with one aggregate shock per location only require solving a linear system—given the deterministic Impulse Value v—and a running a single time iteration.

1.7 Connection with existing numerical methods

Having described the FAME, I relate to existing first-order approaches that build either on a state-space or a sequence-space representation.

1.7.1 State-space methods

The FAME is by nature a state-space approach. It provides the foundation for the computational state-space approach in Reiter (2009) and Ahn et al. (2018). Specifically, the distributional Impulse Value v is the analytic counterpart to the matrix $\mathbf{D}_{vg}$ in their notation. The aggregate shock Impulse Value ω is the analytic counterpart to the matrix $\mathbf{D}_{vZ}$ in their notation. Ahn et al. (2018) rely on automatic differentiation of the nonlinear discretized Bellman equation and law of motion of the distribution to obtain a linear rational expectations system that stacks the linearized Bellman equation and the linearized law of motion of the distribution. They then perform a Blanchard and Kahn (1980) stable root-finding procedure to extract the relevant matrices from the associated Schur decomposition.[19]

[18] When each location is subject to its own shock $Z_{it} = Z_i e^{\varepsilon z_{it}}$ instead, the stochastic FAME becomes $\rho\omega_{\bullet j} = \varpi_{\bullet j} + (M + vG)\omega_{\bullet j} - \theta_j \omega_{\bullet j}$, where $\varpi_{ij} = \zeta u'(C_i^{SS})C_i^{SS}$.

[19] Note that in such a high-dimensional setting, there is no *a priori* guarantee that the stable hyperplane coincides with the true value function.

By contrast, the FAME reveals that linearizing the Bellman equation and the law of motion of the distribution analytically is feasible. Linearizing the economy before discretizing it uncovers a systematic structure that delivers standard Bellman equations that may be solved using standard and fast recursive methods.

The FAME unveils a further block-recursivity property. The deterministic FAME can be solved first, independently from the stochastic FAME. The stochastic FAME can be solved independently in a second step, leveraging a much smaller linear system. This property is particularly useful for estimation of parameters that affect only the stochastic FAME, such as the process for aggregate shocks. In this case, moment-matching estimation only requires to solve the stochastic FAME repeatedly, bypassing the computational need to solve for the deterministic FAME.

1.7.2 Sequence-space methods

Auclert et al. (2021) show that sequence-space linearization is highly efficient in contexts in which the equilibrium is summarized by a few equilibrium prices. In general, depending on the structure of the model, a state-space or a sequence-space approach may be preferable. When many moments of the distribution enter individual decisions, such as in dynamic spatial models as in Section 1.6 or in search-and-matching models, a state-space approach is more natural and is likely to perform better numerically because it is lower-dimensional. When only a few prices enter individual decisions such as in the Krusell and Smith (1998) economy, a sequence-space approach is likely to be more efficient.

In a companion paper (Bilal, 2023), I derive a continuous-time analytic counterpart to Auclert et al. (2021) without relying on numerical differentiation of the individual value function. I show that the linearized individual decisions satisfy a simple Bellman equation that maps into steady-state in closed form. I characterize the properties of analytic sequence-space Jacobians, use them to establish an existence and uniqueness criterion, describe the connection with the FAME and provide efficient algorithms for direct computation.

This section developed the main ideas and benefits of the FAME in the context of the Krusell and Smith (1998) example and in a dynamic discrete choice example. In Sections 2 and 3, I generalize the approach to nest many possible economic settings.

2 The Master Equation

This section builds on Section 1 and develops a general formulation of the Master Equation.

2.1 State space

This section sets up the notation for the remainder of the paper, most of which is to handle mass points in the distribution symmetrically to a smooth density.

Time $t \geq 0$ is continuous and runs forever. The economy is populated by a unit measure of agents. Agents are characterized by their individual state vector $x \in \bar{X}_0 \subset \mathbb{R}^{D_X}$, where $X_0 = (\underline{x}_1, \overline{x}_1) \times ... \times (\underline{x}_{D_X}, \overline{x}_{D_X})$ is a D_X-dimensional hypercube.[20] $\bar{X}_0$ denotes its closure in the Euclidean norm. $\bar{X}_0$ is endowed with the Borel $\sigma-$algebra, and a base measure η. Individuals may choose a control variable $c \in \Gamma \subset \mathbb{R}^{D_C}$.[21]

The base measure η plays a key role in the sequel. It encodes a priori information about where the distribution is absolutely continuous, and where it may develop mass points. If the example of Section 1 was enriched with an occasionally binding borrowing constraint $a \geq \underline{a}$, one would define $d\eta(a, y) = (da + \delta_{\{\underline{a}\}}(da)) \otimes dy$, where $\delta_{\{\underline{a}\}}$ denotes the Dirac measure at $\underline{a}$, and $\otimes$ denotes the tensor product of measures. This definition then allows for the possibility of a mass point at the borrowing constraint $\underline{a}$.

The base measure allows to handle only densities with respect to that base measure, and thus treats mass points and smooth densities symmetrically.[22] In the sequel, I always impose that the base measure is a product measure of marginal measures. The marginal measure along dimension i in turn consists of the Lebesgue measure together with a countable set of possible mass points. These possible mass points are located on a $D_X - 1$-dimensional manifold B that may include parts of the boundary of X_0—for instance when there are credit constraints—and may also include points in the interior of X_0—for instance when there are kinks in the interest rate. I assume that this manifold intersects any direction i at a countable number of points $\{x_{in}\}$ only. I denote by $X = X_0 \cup B$ the domain of the state x as the union of the open domain X together with the set of possible mass points B.

Assumption 1. *(Base measure)*
$d\eta(x) = d\eta_1(x_1) \otimes ... \otimes d\eta_{D_X}(x_{D_X})$, where, for all $i = 1...D_X$ and all Borel-measurable subset $Z \subset [\underline{x}_i, \overline{x}_i]$, $\eta_i(Z) = \ell(Z_i) + \sum_n \mathbb{1}\{x_{in} \in Z\}$, where $\ell(Z)$ denotes the Lebesgue measure of Z.

2.2 Evolution of individual and aggregate states

Agents' state x_t evolves over time according to a controlled jump-diffusion process

$$dx_t = b_{0t}(x_t, c_t)dt + \sigma_{0t}(x_t, c_t) \cdot dW_t + \int (y - x_t) f_{0t}(x_t, c_t, y) d\eta(y) N(dt). \tag{23}$$

b_{0t} is the $\mathbb{R}^{D_X}$-valued drift of the process. It may depend directly on calendar time t, the current state x_t, as well as the control c_t. Similarly, σ_0 is the $\mathbb{R}^{D_X \times D_W}$-valued function volatility of the process.

[20] Working with a hypercube is not strictly necessary for most of the results below. However, it makes the notation much lighter—in a relative sense—to handle mass points. Without mass points, virtually all the results below go through for a general open domain X without additional notation.

[21] To keep the exposition as concise as possible, discussion of filtrations and adaptedness are omitted. See Carmona and Delarue (2018a) for an in-depth exposition.

[22] In principle, it is possible to work without a base measure. In that case, the law of motion of the distribution is set in the space of measures. Consequently, one needs to develop the formalism of derivatives with respect to general measures in the Wassertstein space. See Cardaliaguet et al. (2019) for details. Introducing this formalism is beyond the scope of this paper, and is also irrelevant for many economic applications of interest in which a priori knowledge of where mass points may develop is often available.

It governs how individual states respond to the $\mathbb{R}^{D_W}$-valued Brownian motion W. Importantly, W is independent across agents. The Poisson jump measure N encodes the frequency of jump increments. The density f_0 captures the density of increments and their frequency over the base measures η, N. In the notation is implicit that jumps are also independent across agents. I assume that the stochastic process (23) is a Feller process, which encompasses a wide range of continuous-time Markov processes typically found in applications.

The state x_t may include discrete indicators for different types of agents, for instance employed workers and unemployed workers, workers and firms, regions or countries. The process x_t is assumed to remain within X, either through reflection at the boundary of X, or through an appropriate combination of drift and volatility at the boundary.

I use the notion of weak derivatives to handle mass points in the distribution.[23] When a η-measurable function f is continuously differentiable, the weak derivative coincides with the classical derivative. When f has a jump in direction i at some $x_0 \in X$, then the weak derivative in the sense of generalized functions is a measure: it is a Dirac mass point multiplied by the size of the jump: $\frac{\partial f}{\partial x_i}(x_0)dx \equiv \Big(\lim_{\varepsilon\downarrow 0} f(x_0 + \varepsilon\tau_i) - \lim_{\varepsilon\downarrow 0} f(x_0 - \varepsilon\tau_i)\Big)\delta_{x_0}(dx)$, where τ_i is the unit vector pointing in direction i. In that case, f admits a weak derivative only if $x_0 \in M_i$. Then, the weak derivative as a η-measurable function is the size of the jump at x_0: $\lim_{\varepsilon\downarrow 0} f(x_0 + \varepsilon\tau_i) - \lim_{\varepsilon\downarrow 0} f(x_0 - \varepsilon\tau_i)$. Thus, I can work only with densities with respect to the base measure.

Two functional spaces are useful in the sequel. The first is the space of square-integrable functions with respect to the base measure η. The second is the second Sobolev space:

$$L^2 = \left\{ f : X \to \mathbb{R} \text{ is } \eta\text{-measurable} \;\Big|\; \int f(x)^2 d\eta(x) < \infty \right\}, \qquad H^2 = \left\{ f \in L^2 \;\Big|\; \frac{\partial f}{\partial x}, \frac{\partial^2 f}{\partial x^2} \in L^2 \right\}.$$

The notation L^2 should not be confused with the notation for the generator of the process, L. The second Sobolev space H^2 consists of all square-integrable functions that have square-integrable first and second weak derivatives.

I now define two functional operators related to the state process x_t. The first operator L_{0t} is the generator L_{0t} of the state process and encodes conditional expectations of functions of x_t. For any $V \in H^2$, define:[24]

$$L_{0t}(x,c)[V] = \sum_{i=1}^{D_X} b_{0,i,t}(x,c)\frac{\partial V}{\partial x_i}(x) + \frac{1}{2}\sum_{i,j=1}^{D_X} \Sigma_{0,ij}(x,c)\frac{\partial^2 V}{\partial x_i \partial x_j}(x) + \int f_{0t}(x,c,y)\big(V(y) - V(x)\big)d\eta(y)$$

where $\Sigma_{0t}(x,c) = \sigma_{0t}(x,c)\sigma_{0t}(x,c)^T$, and recall that the ^{T} superscript denotes the matrix transpose.

The second operator is the formal adjoint operator L^*_{0t}.[25] It encodes how the cross-sectional

[23] See Online Appendix E for details

[24] For points on the boundary of X, $x \in \partial X$, it is understood that derivatives are taken with respect to the interior directions to X, and all functions are extended by 0 outside of X.

[25] Appendix E.2 shows that L^*_{0t} is the actual adjoint of L_{0t} only when there are no mass points in the distribution.

probability distribution of the process x_t evolves over time. For any $g \in H^2$, define:

$$\begin{aligned} L_{0t}^*(x,c)[g] &= -\sum_{i=1}^{D_X} \frac{\partial}{\partial x_i}\Big(g(x) b_{0,i,t}(x,c)\Big) + \frac{1}{2}\sum_{i,j=1}^{D_X} \frac{\partial^2}{\partial x_i \partial x_j}\Big(\Sigma_{0,ij}(x,c) g(x)\Big) \\ &+ \int f_{0t}(y,c,x) g(y) d\eta(y) - g(x)\int f_{0t}(x,c,y) d\eta(y) \end{aligned}$$

Aggregate shocks take the form of a stationary and Feller Markov process z_t with values in $Z \subseteq \mathbb{R}^{D_Z}$ and generator $\mathcal{A}$. Z is compact and $\mathcal{A}$ is such that the process z_t has a unique invariant distribution that admits a density $\phi(z)$. Aggregate shocks are scaled by a scalar parameter $\varepsilon \geq 0$ when entering individual decisions and market clearing conditions.

2.3 Individual optimization and evolution of the distribution

Armed with the notation above, I define agents' decision problem. They solve the following time-dependent Bellman equation with possible constraints on the state variable:

$$\rho V_t(x) = \max_{c \in \Gamma} u_{0t}(x,c,V_t) + L_{0t}(x,c)[V_t] + \mathbb{E}_t\left[\frac{dV_t}{dt}(x)\right] \quad \text{s.t.} \quad C_{0t}\left(x, V_t(x), \frac{\partial V_t}{\partial x}(x)\right) \geq 0, \ x \in B. \quad (24)$$

In equation (24), the flow payoff u_{0t} may depend on time, the current state, but also the value function V directly. This formulation embeds recursive preferences such as Epstein-Zin, as well as bargaining models of the labor market.[26]

The function C_{0t} captures constraints on the state x_t when it lies at a possible mass point inside B.[27] For instance, consider adding a credit constraint $a_t \geq \underline{a}$ to the economy of Section 1. In that case, a mass point may develop at $\underline{a}$. Consumption c must be such that $r\underline{a} + wy - c \geq 0$, i.e. $c \leq r\underline{a} + wy$. This constraint on the control may equivalently be re-stated on the value function by $\frac{\partial V_t}{\partial a} \geq u'(r_t \underline{a} + wy)$. Therefore, $C_{0t}\left(a, y, V_t, \frac{\partial V_t}{\partial a}\right) = \frac{\partial V_t}{\partial a} - u'(r_t a + wy)$ for $a = \underline{a}$.

Similarly, consider adding a kink in the interest rate at a^* in the interior of X to the economy of Section 1. For instance, suppose $r_t(a) = r_t$ for $a \geq a^*$, and $r_t(a) = r_t + \bar{r} > r_t$ for $a < a^*$. In that case, a mass point may develop at a^*. In this case, consumption c must be such that $ra^* + wy - c = 0$, which delivers the same state constraint as for the credit constraint: $\frac{\partial V_t}{\partial a} - u'(r_t a + wy)$ for $a = a^*$.

It is useful to define the evolution of the distribution in terms of the density $g_t(x)$ with respect to the base measure η in order to accommodate the presence of possible mass points. g_t satisfies the law of motion:

$$\frac{\partial g_t}{\partial t}(x) = L_{0t}^*(x, \hat{c}_t(x))[g_t]. \quad (25)$$

The notion of weak derivative with respect to the base measure η is key to systematically handle mass

[26]For Epstein-Zin preferences, set $u_t(x,c,V) = \rho V(x) + \rho \frac{1-\gamma}{1-\psi} V(x)\left[\left(c/((1-\gamma)V(x))^{\frac{1}{1-\gamma}}\right)^{1-\frac{1}{\psi}} - 1\right]$ and set $\rho = 0$ in the left-hand-side of (24). Note also that it is straightforward to include shocks to the discount rate ρ, which I omit here for brevity.

[27]See Fleming and Soner (2006) for more details.

points in the evolution of the distribution. As long as mass points develop only where η allows them to, the weak derivative is well-defined in the space H^2: for all $g \in H^2$, $L^*_{0t}(x, \hat{c}_t(x))[g] \in L^2$.

When the economy is stationary, the Bellman equation (24) becomes

$$\rho V(x) \quad = \quad \max_{c \in \Gamma} u_0(x, c, V) + L_0(x, c)[V] \quad \text{s.t.} \quad \mathcal{C}_0 \left(x, V(x), \frac{\partial V}{\partial x}(x) \right) \geq 0, \; x \in B. \tag{26}$$

Similarly, the law of motion of the distribution (25) becomes

$$0 = L^*_0(x, \hat{c}(x))[g]. \tag{27}$$

2.4 General equilibrium

I now specify how the flow payoff u_t as well as the process for the productivity process L_{0t} depend on the underlying distribution g_t and aggregate shocks z_t.[28]

Assumption 2. *(Dependence on aggregates)*
There exist functionals u, b, σ, f, C such that u_0 and the coefficients of L_0 satisfy

$$u_{0t}(x, c, V_t) = u(x, c, \varepsilon z_t, g_t, V_t), \qquad b_{0t}(x, c) = b(x, c, \varepsilon z_t, g_t, V_t), \qquad \sigma_{0t}(x, c) = \sigma(x, c, \varepsilon z_t, g_t, V_t),$$

$$f_{0t}(x, c, y) = f(x, c, y, \varepsilon z_t, g_t, V_t), \quad C_{0t} \left(x, V, \frac{\partial V}{\partial x} \right) = C \left(x, V, \frac{\partial V}{\partial x}, \varepsilon z_t, g_t \right)$$

In addition, $u, b, \sigma, f, \mathcal{C}$ are continuously L^2-Fréchet-differentiable in g and V, and continuously differentiable in z.

I impose Assumption 2 in the rest of the paper. It captures how prices and other general equilibrium forces feed back into individual decisions. Assumption 2 is typically mediated through market clearing conditions. It is widely satisfied in applications as in Section 1, in which the dependence on calendar time is a shorthand for dependence on the the distribution of individual states g. The dependence on V_t captures typical dependences in labor market models, in which vacancy creation depends on the distribution of surpluses across jobs. For the flow payoff u, dependence on V captures dependence on the individual's own value (e.g. for Epstein-Zin preferences) and dependence on the distribution of values of other agents.

Importantly, Assumption 2 allows for a much more flexible dependence of the individual decision problem than only a few prices. There can be any arbitrary number of prices that matter, as in Section 1.6. The distribution can also matter directly for individual decisions, as in a search-and-matching model with job-to-job search. I now define an equilibrium of the economy.

Definition 1. *(Equilibrium in sequential form)*
An equilibrium in sequential form of the economy consists of a path of distributions $(g_t)_{t \geq 0} \in H^2$, a path of values $(V_t)_{t \geq 0} \in H^2$ such that (24) and (25) holds for all times $t \geq 0$, and g_0 is given.

[28] To keep the exposition minimal, I assume that time dependence runs only through the distribution g_t as in Section 1. It is not difficult to let time affects the economy deterministically. In that case, one need only treat time as another state variable.

There is no need to require any market to clear, since market clearing is embedded in Assumption 2. Similarly, a steady-state equilibrium is defined as follows.

Definition 2. *(Steady-state equilibrium)*
A steady-state equilibrium of the economy consists of a distribution $g^{SS} \in H^2$, a value $V^{SS} \in H^2$ such that (26) and (27) hold.

2.5 From the time-dependent problem to the Master Equation

Now turn to the Master Equation. Under Assumption 2, the economy may be represented fully recursively as a value function defined on idiosyncratic states as well as the space of distributions $X \times H^2$. As in Section 1, the key step is to change variables from calendar time t to the distribution g_t by writing $V_t(x) \equiv V(x, g_t, z_t)$.

Fréchet derivatives are well-defined in L^2 because it is a Hilbert space when equipped with the inner product $\langle f, g\rangle = \int f(x)g(x)d\eta(x)$. Suppose for now that $g \mapsto V(x, g)$ admits a Fréchet derivative for η-almost all x. It consists of a linear bounded operator from L^2 onto itself. Using the Riesz representation theorem, it may in turn be represented by an L^2 function. Denote this L^2 function by $x' \mapsto \frac{\partial V}{\partial g}(x, x', g, z)$.

The same change of variables as in Section 1, using the chain rule for Fréchet derivatives and the law of motion of the distribution (25) in the Bellman equation (24), delivers the Master Equation.

Definition 3. *(Master Equation)*
The Master Equation is defined by

$$\begin{aligned}
\rho V(x, z, g) \quad &= \quad \max_{c \in \Gamma} u(x, \varepsilon z, c, V, g) + L(x, \varepsilon z, c, g)[V] \\
&+ \quad \int \frac{\partial V}{\partial g}(x, x', z, g) L^*(x', \varepsilon z, \hat{c}(x'), g)[g] d\eta(x') + \mathcal{A}(z)[V] \\
&s.t. \quad C\left(x, V(x, z, g), \frac{\partial V}{\partial x}(x, z, g), \varepsilon z, g\right) \geq 0, \quad x \in B,
\end{aligned}$$

for functions $X \times Z \times H^2 \times \mathbb{R}_+ \ni (x, z, g) \mapsto V(x, z, g)$ that are Fréchet-differentiable in g η-a.e. in x.

The Master Equation in Definition 3 delivers a natural definition of a recursive equilibrium.

Definition 4. *(Equilibrium in recursive form)*
An equilibrium in recursive form consists of a solution V to the Master Equation in Definition 3.

Definition 4 emphasizes that a value function that solves the Master Equation is the only object that is needed to describe the equilibrium. By construction, both notions of equilibrium coincide whenever defining a solution to the Master Equation is possible as shown in Proposition 1.

Proposition 1. *(Coincidence of recursive and sequential competitive equilibrium)*
Suppose that there exists an equilibrium in recursive form given by a solution $V(x, g)$ to the Master

Equation in Definition 3. Define $V_t(x) \equiv V(x, g_t)$, and let g_t solve (25). Then the pair (V_t, g_t) defines an equilibrium in sequential form.

A natural and important question is of course when does a solution to the Master Equation exist at all. Several set of assumptions have been proposed recently in the mathematics mean field games literature. Since this literature is still growing, the set of assumptions is still limited at the time of writing. In particular, typical assumptions exclude several key economics applications, such as the presence of a credit constraint or multiplicative interactions between prices and states in the savings rate. Therefore, I do not list those assumptions in this paper, and merely point to existing results.[29] These assumptions also guarantee that the value function $V(x, g, z)$ is L^2-Fréchet-differentiable in the distribution g up to second order. In addition, these assumptions are typically the same as those required for existence and uniqueness of an equilibrium in sequential form.

It is not the purpose of this paper to attempt expanding the set of assumptions leading to existence and uniqueness results for the fully non-linear Master Equation—though one may hope that such results will eventually become available for most setups of economics interest. Instead, this paper is concerned with the more practical question of local approximations to the Master Equation conditional on the existence of at least one isolated steady-state equilibrium. Therefore, I impose the following assumption in the sequel.

Assumption 3. *(Existence, local uniqueness and regularity)*
There exists one isolated steady-state equilibrium V^{SS}, g^{SS}. There exists a solution $V(x, z, g)$ to the Master Equation in a neighborhood of $(g^{SS}, z = 0)$ that is continuously Fréchet-differentiable in $g \in H^2$ around $g = g^{SS}$, and continuously differentiable in ε around $\varepsilon = 0$, η-almost everywhere in x. The $D_X - 2$-dimensional boundary of B such that the state constraint holds with equality is continuously Fréchet-differentiable in g around $g = g^{SS}$, and differentiable in z around $z = 0$, η-almost everywhere in x.

Assumption 3 paves the way for local perturbations of the the Master Equation, which is the subject of the next section.

3 The FAME

This section derives the FAME, highlights its properties and derives its implications for the local stability of the steady-state and for the stochastic steady-state. Finally, this section proposes an efficient numerical implementation to compute the solution to the FAME.

[29] See for instance Theorem 2.4.2. p. 39 in Cardaliaguet et al. (2019) or Theorem 4.2.1. p. 295 in Carmona and Delarue (2018b).

3.1 Impulse Values and optimal control

Start by fixing a locally isolated steady-state equilibrium g^{SS}, V^{SS}, which I will refer to as "the" steady-state for brevity in the rest of the paper. Consider a small perturbation of the distribution around the steady-state, $g = g^{SS} + \varepsilon h$. I seek a first-order solution to the nonlinear Master Equation in the form:

$$V(x, z, g^{SS} + \varepsilon h) = V^{SS}(x) + \varepsilon \left\{ \int v(x, x')h(x')d\eta(x') + \omega(x, z) \right\}. \tag{28}$$

As in Section 1, $v(x, x')$ is the deterministic Impulse Value, i.e. the directional derivative of the value function with respect to the distribution g, evaluated at steady-state g^{SS}: $v(x, x') = \frac{\partial V}{\partial g}(x, x', g^{SS}, 0)$. $\omega(x, z)$ is the stochastic Impulse Value: the derivative of the value function with respect to aggregate shocks.

The linearization of the value function in (28), leads to the linearization of the optimal control:

$$\hat{c}(x, g^{SS} + \varepsilon h, z) = c^{SS}(x) + \varepsilon \left\{ \int \mathcal{M}(x, x', v)h(x')d\eta(x') + \overline{\mathcal{M}}(x, z, \omega) \right\} \tag{29}$$

The kernel $\mathcal{M}(x, x', v)$ is the distributional Marginal Propensity to Control (dMPC). The component $\overline{\mathcal{M}}(x, z, \omega)$ is the stochastic MPC (sMPC). I express them below. They encode how individual controls for individuals at x respond to a small distributional impulse $h(x')$ and to a small aggregagets shock. I need the following regularity condition, satisfied in most applications of interest, e.g. when the individual decision problem is strictly concave.

Assumption 4. *(Control regularity)*
The first-order optimality condition for the optimal control $\hat{c}(x, g, z)$ holds with equality for $||g - g^{SS}||_{L^2}$ and ε small enough. $\mathcal{U}(x) \equiv u_{cc}^{SS}(x) + \mathcal{L}_{cc}(x)[V^{SS}]$ is an invertible matrix for η-a.a. x.

Crucially, in continuous time as opposed to discrete time, the first order optimality conditions typically hold with equality in applications, even with credit constraints (Achdou et al., 2021), so that Assumption 4 holds. Thus, it suffices to differentiate the first-oder optimality condition without carrying Lagrange multipliers. I impose Assumption 4 in the rest of the paper.

3.2 The deterministic FAME

In light of Section 1, it is natural to expect the deterministic Impulse Value to satisfy a Bellman equation—the deterministic FAME. To define that Bellman equation as concisely as possible, it is useful to denote objects evaluated at steady-state functions by a script letter. For instance, denote $\mathcal{L}(x) \equiv L(x, c^{SS}(x), V^{SS}, g^{SS})$. Similarly, denote partial usual or Fréchet derivatives by the corresponding subscript. For instance, denote $u_g(x, x') \equiv \frac{\partial u}{\partial g}(x, x', V^{SS}, g^{SS}, 0)$. Denote also by $\mathcal{B}$ the set of states x such that the state constraint holds with equality in the steady-state. Appendix A.2 details the remaining notation.

The first step towards the deterministic FAME is to express how the optimal control in equation (29) responds to changes in the distribution.

Proposition 2. *(Deterministic optimal control)*

$$\begin{aligned}\mathcal{M}(x,x',v) &= -\mathcal{U}(x)^{-1}\left(m_0(x,x') + \mathcal{L}_c(x)[v(\cdot,x')] + \int m_1(x,y)v(y,x')d\eta(y)\right)\\ m_0(x,x') &\equiv u_{cg}(x,x') + \mathcal{L}_{cg}(x,x')[V^{SS}] \quad , \quad m_1(x,x') \equiv u_{cV}(x,x') + \mathcal{L}_{cV}(x,x')[V^{SS}].\end{aligned}$$

Proof. See Appendix B.1. □

Proposition 2 characterizes how individuals' controls respond to a distributional impulse h. The dMPC depend on the concavity of the utility function as well as the concavity of the generator L. The dMPC also depends on the corresponding cross-derivatives. I provide more specific expressions for the dMPC for specific generators in Proposition 4, Appendix B.1.

With Proposition 2 at hand, the deterministic FAME obtains.

Theorem 1. *(Deterministic FAME)*
The deterministic Impulse Value v satisfies for all $x \in X\backslash\mathcal{B}$ and all $x' \in X$:

$$\begin{aligned}\rho v(x,x') &= \overbrace{u_g(x,x') + \int u_V(x,y)v(y,x')d\eta(y) + \mathcal{L}_g(x,x')[V^{SS}] + \int \mathcal{L}_V(x,y)[V^{SS}]v(y,x')d\eta(y)}^{\text{Direct impact}}\\ &+ \underbrace{\mathcal{L}(x)\big[v(\cdot,x')\big]}_{\substack{\text{Partial equilibrium:}\\ \text{continuation value from}\\ \text{shocks to } x}} + \underbrace{\mathcal{L}(x')\big[v(x,\cdot)\big]}_{\substack{\text{General equilibrium:}\\ \text{continuation value from}\\ \text{propagation of impulse at } x'}} + \underbrace{\int v(x,x'')G(x'',x',v)d\eta(x'')}_{\substack{\text{General equilibrium: weighted average}\\ \text{of changes in decisions of other agents } x''\\ \text{in response to impulse at } x'}},\end{aligned}$$

together with, for all $x \in \mathcal{B}$ and $x' \in X$:

$$0 = \mathcal{C}_g(x,x') + \mathcal{C}_V(x)v(x,x') + \mathcal{C}_{\partial V}(x) \cdot \frac{\partial v}{\partial x}(x,x'), \quad \textit{all } x \in \mathcal{B}, \ \textit{all } x',$$

and where

$$G(x'',x',v) = \mathcal{L}^*_g(x'',x')[g^{SS}] + \int v(y,x')\mathcal{L}^*_V(x'',y)[g^{SS}]\ d\eta(y) + \mathcal{L}^*_c\big(x'',\mathcal{M}(x'',x',v)\big)[g^{SS}].$$

Proof. See Appendix B.2. □

The structure of the FAME in Theorem 1 generalizes equation (9) in Section 1. The main addition is the linearization of the state constraint. The state constraint binds in the first-order approximation exactly at points where it binds in steady-state. It is natural to expect the state constraint to bind out of steady-state in a neighborhood of the points where it binds in steady-state. Perhaps surprisingly, these set of points turn out to coincide exactly. This conclusion arises because changes in where the state constraint binds in response to an impulse h result in a second-order contribution when interacted with the state constraint itself.

3.3 The stochastic FAME

With the deterministic FAME at hand, I turn to the stochastic FAME. As for the deterministic component, I first express how the optimal control in equation (29) responds to aggregate shocks.

Proposition 3. *(Stochastic optimal control)*

$$\begin{aligned}\overline{\mathcal{M}}(x,z,\omega) &= -\mathcal{U}(x)^{-1}\left(\overline{m_0}(x)\cdot z + \mathcal{L}_c(x)[\omega(\cdot,z)] + \int \overline{m_1}(x,y)\omega(y,z)d\eta(y)\right)\\ \overline{m_0}(x) &\equiv u_{cZ}(x) + \mathcal{L}_{cZ}(x)[V^{SS}] \quad , \quad \overline{m_1}(x,y) \equiv u_{cV}(x,y) + \mathcal{L}_{cV}(x,y)[V^{SS}].\end{aligned}$$

Proof. Follow the same steps as in the proof of Proposition 2. □

With Proposition 3, I obtain the stochastic FAME.

Theorem 2. *(Stochastic FAME)*
The stochastic Impulse Value ω satisfies for all $x \in X\backslash\mathcal{B}$ and all $x' \in X$:

$$\rho\omega(x,z) = z\cdot\left\{u_Z(x) + \mathcal{L}_Z(x)[V^{SS}]\right\} + \mathcal{L}(x)[\omega(\cdot,z)] + \mathcal{A}(z)[\omega(x,\cdot)] + \int v(x,x')S(x',\omega,z)d\eta(x'),$$

together with, for all $x \in \mathcal{B}$ and all $z \in Z$:

$$0 = \mathcal{C}_V(x)\omega(x,z) + \mathcal{C}_{\partial V}(x)\cdot\frac{\partial\omega}{\partial x}(x,z),$$

and where

$$S(x',z,\omega) = \mathcal{L}_c^*\big(y,\overline{\mathcal{M}}(y,z,\omega)\big)[g^{SS}]\Big) + \int \mathcal{L}_V^*\big(x',y,\omega(y,z)\big)[g^{SS}]d\eta(y)$$

Proof. Follow the same steps as in the proof of Theorem 1. □

As in Section 1, the economy is block-recursive. The deterministic Impulse Value v is independent from the stochastic Impulse Value ω. One only needs to solve for the stochastic Impulse Value ω after having solved for the deterministic Impulse Value v.

In many applications such as Section 1, the generator $\mathcal{A}$ of aggregate shocks scales as: $\mathcal{A}(z)[\varphi] = z\mathcal{B}[\varphi]$ when φ is linear. In this case, it is straightforward to guess and verify that the stochastic Impulse Value scales linearly in z: $\omega(x,z) = \omega_0(x)\cdot z$. In that case, ω_0 satisfies the simpler stochastic FAME

$$\rho\omega_0(x) = u_z(x) + \mathcal{L}_z(x)[V^{SS}] + \mathcal{L}(x)[\omega_0] + +\mathcal{B}[\omega_0] + \int v(x,x')S_0(x',\omega_0)d\eta(x'),$$

with S_0 defined analogously to S.

3.4 Impulse response functions

Equipped with the deterministic and stochastic Impulse Values, I characterize impulse response functions in this economy.

Theorem 3. *(Evolution of the distribution)*
To first order, following a path of aggregate shocks $\{z_t\}_{t=0}^{\infty}$*, the impulse in the distribution* h_t *follows the SPDE*

$$\frac{dh_t(x)}{dt} = \mathcal{L}^*(x)[h_t] + \mathcal{G}(x)[h_t] + \mathcal{S}(x, z_t),$$

where $\mathcal{G}(x)[h] = \int G(x, x', v)h(x')d\eta(x')$*,* $\mathcal{S}(x, z_t) = S(x, z_t, \omega)$*, and* h_0 *is given.*

Proof. Follow the steps in the proof of Theorems 1 and 2. □

Theorem 3 shows that the first-order dynamics of the distribution follow a linear law of motion. As expected, the evolution depends on the steady-state transition probabilities embedded in the generator $\mathcal{L}^*(x)$. This contribution to the law of motion represents the partial equilibrium response of aggregate dynamics to an impulse h in the distribution.

The evolution of the distribution also depends on the general equilibrium feedback of the economy, as highlighted by the integral operator $\mathcal{G}(x)$. The action of this operator on the distribution embeds the first-order response of individual controls $\hat{c}$ to an impulse in the distribution, through the dMPCs and the direct impact of a distributional impulse on transition probabilities. Finally, the evolution of the distribution depends on the impact of aggregate shocks, summarized in $\mathcal{S}(x, z_t)$. As aggregate shocks hit the economy, individuals change their decisions, which affects the law of motion of the distribution.

3.5 Dynamic stability

Theorem 3 is particularly useful to obtain stability and convergence conditions. I seek conditions under which $||h_t||_1 \to 0$ at an exponential rate, where $||h_t||_1 = \int |h_t(x)|d\eta(x)$. To that end, it is useful to first consider the partial equilibrium stochastic process x_t^{PE} given by the generator $\mathcal{L}$. The partial equilibrium process x_t^{PE} corresponds to the dynamics of the state x_t when prices are held constant at their steady-state values and controls are chosen accordingly. Denote by $P_t(x, \chi)$ the probability that this process reaches the set χ starting from x after time t.[30] Starting from an initial impulse in the distribution h_0, denote by h_t^{PE} the resulting impulse after time t. By Theorem 3, h_t^{PE} satisfies $\partial_t h_t^{PE} = \mathcal{L}^*(x)[h_t^{PE}]$.

Theorem 4. *(Partial equilibrium stability)*
Suppose that either one of the following condition holds:

(i) There exists a time $\tau > 0$ *and a constant* $\alpha > 0$ *such that* $P_\tau(x, \chi) \geq \alpha \int_\chi g^{SS}(x')d\eta(x')$*, for all* $x \in X$ *and* $\chi \subset X$*.* x_t^{PE} *is an aperiodic process.*

(ii) There exists a function $W : X \to [1, +\infty)$ *and a constant* $\beta > 0$ *such that for all* $x \in X$*,* $\mathcal{L}(x)[W] \leq -\beta W(x)$*.*

Then there exists $\gamma > 0, R > 0$ *such that* $||h_t^{PE}||_1 \leq Re^{-\gamma t}||h_0||_1$ *for all* h_0 *that integrates to 0.*

[30] P_t is called the semigroup with generator $\mathcal{L}$.

Proof. See Appendix B.3. □

Theorem 4 provides two possible sufficient conditions under which the steady-state, partial equilibrium process converges back to its steady-state distribution g^{SS} at an exponential rate. In both cases, the rate of convergence γ is related to the constants α or β, but in general no direct mapping is available.

Why is Theorem 4 useful? When the state space is finite dimensional (as in the RBC model, or the New Keynesian model), standard linear algebra techniques suffice to characterize convergence. In particular, if the second largest eigenvalue of the counterpart of $\mathcal{L}^*$—the transition matrix—has a negative real part in continuous time, then the partial equilibrium dynamics are stable.[31] This property continues to hold in numerically discretized frameworks that represent economies with an underlying continuum of agents, with an important caveat.

When the underlying state space is truly continuous, the theoretically exact dynamics are given by transition operators such as $\mathcal{L}^*$ that act in infinite-dimensional spaces. There, it turns out that much of the finite-dimensional spectral theory fails. The main difficulty consists in the possibility of a continuous spectrum: the possibility that $\mathcal{L}^*$ has a continuum of eigenvalues that includes 0. In that case, convergence would be slower than exponential. If this is the case, even finite-dimensional numerical discretizations of $\mathcal{L}^*$ and their eigenvalues are a poor indicator of convergence because they deliver a second largest eigenvalue very close to 0. Crucially, the value of this eigenvalue approaches zero as the discretization becomes finer.

Thus, simply checking the second largest eigenvalue of the numerically discretized economy is not enough to characterize exponential convergence back to steady-state in heterogeneous agent economies. To avoid this difficulty, one must ensure a spectral gap, that is, the property that the second-largest eigenvalue of $\mathcal{L}^*$ is strictly negative. This is what conditions (i) or (ii) do.

Condition (i) is often called Doeblin's condition. It ensures that the process is mixing enough, by stating that transition probabilities are uniformly bounded below by a multiple of the steady-state distribution. It is related to conditions D and M in Stokey et al. (1989), which also ensure convergence of Markov processes back to their invariant distribution. Note, however, that condition D only ensures convergence of the Cesaro means of h_t^{PE}. Condition M ensures convergence of the actual distribution but requires a uniform lower bound on transition probabilities that is unlikely to hold in settings in which transition probabilities to certain parts of the state space become arbitrarily low. Condition (i) in Theorem 4 allows for that possibility by scaling the lower bound by the invariant distribution itself. Because α can be arbitrarily small, condition (i) is likely to hold in many settings of interest.

When it is difficult to check condition (i), one can turn to condition (ii). The function W in condition (ii) is called a Lyapunov function, or an energy function. The inequality in condition (ii) guarantees that $W(x_t^{PE})$ decreases at a geometric rate on average, which in turn ensures exponential stability of the partial equilibrium process.

[31] In discrete time, if the second largest eigenvalue is inside the unit circle.

It may be tempting to use the value function V^{SS} as a candidate Lyapunov function. For it to satisfy the inequality, one needs $\beta \leq \inf_{x \in X} \frac{u^{SS}(x)}{V^{SS}(x)} - \rho$. In practice however, the infimum on the right-hand-side is often negative. To circumvent this difficulty, a mixture of conditions (i) and (ii) can also be used.[32] If condition (i) holds for $x \in \Delta \subset X$, and condition (ii) holds for $x \in X \backslash \Delta$, then the conclusions of Theorem 4 continue to hold. In particular, the steady-state value function is a Lyapunov function if it is strictly positive and $\beta \leq \inf_{x \in X \backslash \Delta} \frac{u^{SS}(x)}{V^{SS}(x)} - \rho$, where the infimum now excludes Δ. Thus, one can replace low values of the right-hand-side with enough mixing.

Theorem 4 characterizes stability of the partial equilibrium dynamics. The general equilibrium dynamics in Theorem 3 depend on the general equilibrium feedback encoded in the operator $\mathcal{G}$. The following result provides condition for stability of the general equilibrium dynamics.

Theorem 5. *(General equilibrium stability)*
Suppose that the conditions of Theorem 4 are satisfied, with R, γ controlling convergence.

(i) If $\mathcal{G}$ is bounded and is dissipative, i.e. for all $h \in H^2$, $\langle h, \mathcal{G}h \rangle \leq 0$, then the general equilibrium dynamics satisfy exponential convergence at rate γ when $z_t = 0$ for all t: $||h_t||_1 \leq Re^{-\gamma t}||h_0||_1$ for all h_0 that integrates to 0.

(ii) If the deterministic FAME depends continuously on $u_g + \mathcal{L}_g[V^{SS}]$ in $||\cdot||_1$, then for $||u_g + \mathcal{L}_g[V^{SS}]||_1$ small enough, the general equilibrium dynamics satisfy exponential convergence at rate $\gamma' \in (0, \gamma)$ when $z_t = 0$ for all t.

In both cases, h_t remains bounded in the presence of bounded aggregate shocks.

Proof. See Appendix B.4. □

Theorem 5 provides two conditions under which the general equilibrium dynamics are stable and lead to exponential convergence back to steady-state in the absence of aggregate shocks.

Condition (i) states that when general equilibrium forces stabilize the economy and pull it back to steady-state, the convergence happens at least as fast as in partial equilibrium. The stabilizing effect of general equilibrium forces is encoded in the condition that $\mathcal{G}$ is dissipative. In the Krusell and Smith (1998) economy of Section 1, it is possible to show that $\mathcal{G}$ is indeed dissipative, and so convergence back to steady-state is exponential.[33] In richer settings however, it can be challenging to check whether $\mathcal{G}$ is dissipative because it depends on the deterministic Impulse Value v. In some settings, general equilibrium forces may even be de-stabilizing. In that case, when can there be convergence back to steady-state?

Condition (ii) provides an answer when general equilibrium forces cannot be shown to be stabilizing, but can be shown to be small. When this is the case and the deterministic Impulse Value depends smoothly on the direct impact functions, then exponential convergence holds because the partial equilibrium stabilizing effect dominates.

[32] See Appendix B.3 for the proof.

[33] To see this, exploit the relationship between the constants $\mathcal{R}_i, \mathcal{W}_i$ and concavity of the utility function.

In the presence of aggregate shocks, the distributional impulse h_t remains bounded because the general solution to the law of motion in Theorem 3 is

$$h_t = Q_t h_0 + \int_0^t Q_{t-s}\mathcal{S}(z_s)ds, \tag{30}$$

where $Q_t h_0$ denotes the solution to $\partial_t k_t = (\mathcal{L}^* + \mathcal{G})k_t, k_0 = h_0$ and is called the semigroup associated with $\mathcal{L}^* + \mathcal{G}$. This solution remains bounded as soon as $||Q_t h_0||_1 \leq R' e^{-\gamma' t}||h_0||_1$ for some $R', \gamma' > 0$ as guaranteed by Theorem 5.

Having established conditions under which the distribution $h_t(x)$ of individuals across states x remains bounded, I characterize its own probability distribution across possible values h: the stochastic steady-state.

3.6 Stochastic steady-state

It is useful to view the linearized KF equation in Theorem 3 as a law of motion for the density h_t. Theorem 3 defines a SDE in the infinite-dimensional space of functions $h_t \in L^2$. Denote by $\mathcal{P}(dh, dz)$ the joint probability measure of the deviation in the distribution, h_t, and the aggregate shock, z_t, in the stochastic steady-state guaranteed by Theorem 5.

In general, $\mathcal{P}$ is a complex and high-dimensional object because it takes as argument a distribution, h. However, because it represents the density of the solution to a SDE, it turns out that $\mathcal{P}$ itself satisfies a 'meta' Kolmogorov Forward equation. I specify it in Appendix B.5. Attempting to solve this 'meta' KF equation runs into the curse of dimensionality, because $\mathcal{P}$ is a function of a distribution h.

Crucially however, the 'meta' KF equation aggregates because of the linearity of the law of motion elucidated in Theorem 3. Thus, it is possible to tightly determine the conditional average of h_t under $\mathcal{P}$ given z: $\bar{h}(z) = \mathbb{E}_{\mathcal{P}}[h|z]$, which is itself a L^2 function for each z. Hence, I also write $\bar{h}(z) \equiv \bar{h}(x, z)$. Denote $\mathcal{J}(x) = \mathcal{L}^*(x) + \mathcal{G}(x)$.

Theorem 6. *(First moments in the stochastic steady-state)*
Suppose that the conditions of Theorem 5 hold and that $\mathcal{S}$ is continuous in z. Then, there exists an invariant measure $\mathcal{P}(dh, dz)$ in the stochastic steady-state, and the conditional average of such a distribution in the stochastic steady-state is given by $\bar{h}(x, z) = \hat{h}(x, z)/\phi(z)$, where $\hat{h}$ satisfies:

$$0 = \mathcal{J}(x)[\hat{h}(\cdot, z)] + \mathcal{A}^*(z)[\hat{h}(x, \cdot)] + \mathcal{S}(x, z)\phi(z).$$

Proof. See Appendix B.5. □

Theorem 6 reveals that the conditional distribution $\bar{h}(x, z)$ satisfies a *standard* KF equation that only depends on konwn steady-state objects. Thus, it can be solved with standard techniques. In fact, once the KF equation in Theorem 6 is discretized, $\hat{h}(x, z)$ satisfies a standard Sylvester equation.

The same logic also holds for higher-order moments of the distribution in the stochastic steady-state. Denote by $\overline{H}(x, x', z) = \mathbb{E}_{\mathcal{P}}[h(x)h(x')|z]$.

Theorem 7. *(Second moments in the stochastic steady-state)*
Under the conditions of Theorem 6, $\overline{H}(x,x',z) = \widehat{H}(x,x',z)/\phi(z)$, where $\widehat{H}$ satisfies:

$$0 = \mathcal{J}(x)[\widehat{H}(\cdot,x',z)] + \mathcal{J}(x')[\widehat{H}(x,\cdot,z)] + \mathcal{A}^*(z)[\widehat{H}(x,x',\cdot)] + \mathcal{S}(x,z)\hat{h}(x',z) + \mathcal{S}(x',z)\hat{h}(x,z).$$

Proof. See Appendix B.5. □

Theorem 7 delivers a simple equation that determines second-order moments of the distribution $\mathcal{P}$. These second moments can thus be easily used as inputs in an estimation routine. They are also key to characterizing the welfare cost of aggregate risk in the SAME in Section 4.3.

3.7 Solution method

Three key properties of the FAME facilitate the computation of impulse responses. The first property is block-recursivity: first solve the deterministic FAME from Theorem 1, then solve the stochastic FAME of Theorem 2, and finally simulate an Impulse Response using Theorem 3.

The second property is that the FAMEs in Theorems 1 and 2 have the structure of a standard jump-diffusion Bellman equation in finite dimension. In particular, the general equilibrium effects enters the FAMEs in Theorems 1 and 2 just as a collection of standard diffusion terms would. Therefore, readily available and highly efficient discretization schemes apply.

The third property is that the FAME provides a closed-form mapping between steady-state objects and all the elements of the Bellman equation to solve. These observations lead to the following numerical scheme, described in pseudo-code below. I focus on the case $u_V, L_V = 0$ and no state constraint for simplicity, but it is straightforward to extend the scheme when these partial derivatives are not zero or the state constraint binds.

Corollary 1. *(Numerical implementation)*
Define grids $\{x_i\}_{i=1}^I$, $\{z_k\}_{k=1}^K$, $\{z_\ell\}_{\ell=1}^L$ and a time step Δ. Define the matrices $\mathrm{v}_{ij} = v(x_i,x_j) \in \mathbb{R}^{I\times I}$ and $\omega_{ik} = \omega(x_i,z_k) \in \mathbb{R}^{I\times K}$. Then:

1. ***Deterministic FAME.*** *Let $\mathrm{L}, \mathrm{u}_g, \mathrm{M}, \mathrm{N}$ and Pv discretize $\mathcal{L}, u_g, \mathcal{L}_g(x,x')[g^{SS}], \mathcal{L}^*(x,x')[g^{SS}]$ and $\mathcal{L}_c^*(x,\mathcal{M}(x'',x',v)[g^{SS}]$ respectively. Then:*
   - *Guess v^0*
   - *Given v^n, update v^{n+1} by solving the standard Sylvester equation:*

   $$\Big(\rho\mathrm{Id} - \mathrm{L}\Big)\mathrm{v}^{n+1} - \mathrm{v}^{n+1}\Big(\mathrm{L}^T + \mathrm{M} + \mathrm{N}\mathrm{v}^n\Big) \quad = \quad \mathrm{u}_g.$$

   - *Stop when v^{n+1} and v^n are close enough.*

2. ***Stochastic FAME.*** *Let $\mathrm{u}_Z, \mathrm{Q}, A$ discretize $zu_Z(x), z\mathcal{L}_Z(c)[V^{SS}], \mathcal{A}$ respectively. Then ω solves the Sylvester equation*

   $$\Big(\rho\mathrm{Id} - \mathrm{L} - \mathrm{vP}\Big)\omega - \omega\mathrm{A}^T = \mathrm{u}_Z + \mathrm{Q}$$

3. ***Impulse response functions.*** *Let* $\mathrm{G} = \mathrm{M} + \mathrm{N}\mathrm{v}$ *discretize the operator* $\mathcal{G}$. *Let* S_t *discretize* $S(\cdot, z_t)$ *for any* t. *Given* h_0 *and a time step* Δ, *the discretized distributional impulse* h_t *solves the recursion*

$$\mathrm{h}_{t+\Delta} = \mathrm{h}_t + \Delta\Big\{\mathrm{L}^T\mathrm{h}_t + \mathrm{G}\mathrm{h}_t + \mathrm{S}_t\Big\}.$$

4. ***Stochastic steady-state.*** *Let* $\bar{\mathrm{h}}, \hat{\mathrm{h}}, \phi, \mathrm{A}, \mathrm{S}$ *discretize* $\bar{h}, \hat{h}, \varphi, \mathcal{A}, S$. $\hat{\mathrm{h}}, \phi$ *solve the Sylvester and linear equations, respectively:*

$$0 = (\mathrm{L}^T + \mathrm{G})\hat{\mathrm{h}} + \hat{\mathrm{h}}A + S \quad , \quad 0 = A^T\phi,$$

and set $\bar{\mathrm{h}}_{i\ell} = \hat{\mathrm{h}}_{i\ell}/\phi_\ell$.

Corollary 1 provides a simple way of computing Impulse Values and impulse response functions to first order. The discretization of steady-state operators into matrices follows standard finite difference rules as described in Achdou et al. (2021).

At the heart of Corollary 1 lies the specific structure of the FAME in Theorem 1. Once discretized, the deterministic FAME becomes a modified Sylvester matrix equation. A standard Sylvester matrix equation is a linear system with a specific structure, so that it may be written as a function of a matrix unknown Y that satisfies $AY + YB = C$ for known matrices A, B, C. Of course, it is always possible to stack this linear system and solve it without exploiting the Sylvester structure. However, doing so would abstract from useful information about the structure of the linear system. Instead, standard routines such as Matlab's `sylvester.m` function solve the standard Sylvester equation more efficiently than the stacked system.

The deterministic FAME however leads to a Sylvester equation with a quadratic term. Thus, Corollary 1 proposes an iterative scheme leveraging a sequence of standard Sylvester equations. A key observation is to treat the quadratic component vNv consistently with implicit schemes. The first Impulse Value v in the quadratic component represents household's x own Impulse Value. Thus, it is natural to treat it as implicit—solve for it endogenously at every iterative step. The second Impulse Value v in the quadratic component represents the change in control of all other households x''. Thus, it is natural to treat it as explicit—exogenous from the perspective of a given iterative step.[34]

There is an alternative calculation of the deterministic Impulse Value. Rather than solving the deterministic FAME directly, one can view as a linear system in (v_t, h_t) and solve for v using linear rational expectation techniques by imposing stability as in Ahn et al. (2018). Theorem 5 provides sufficient conditions under which stability occurs. When these conditions are satisfied, both solution methods deliver the same outcome. I provide more details in Appendix B.6.

Once the solution to the deterministic Impulse Value is known, the stochastic Impulse Value satisfies a standard Sylvester equation from Theorem 2. When the generator for aggregate shocks scales $\mathcal{A}[\varphi] =$

[34] A formal proof of convergence is beyond the scope of this paper. Yet, this implicit-explicit structure leads to robust convergence in practice. By contrast, reversing which Impulse Value is treated as implicit or explicit leads to systematic divergence of the scheme.

$z\mathcal{B}[\varphi]$ for φ linear, then the stochastic FAME defines a linear vector equation, which solves more efficiently than the more general Sylvester equation. With both Impulse Values at hand, iterating forward on the linearized law of motion in Theorem 3 is straightforward.

In addition to providing a systematic approach to first-order perturbations, the Master Equation is uniquely suited to obtain higher-order perturbations.

4 The SAME

This section develops the Second-order Approximation to the Master Equation (SAME).

4.1 Setup

To obtain the SAME, the strategy is the same as for the FAME. When ε is small enough, the second-order approximation to the value function is

$$\begin{aligned} V(x,z,g^{SS}+\varepsilon h) \;=\;& \overbrace{V^{SS}(x)}^{\text{Steady-state}} + \varepsilon \overbrace{\left\{ \int v(x,x')h(x')d\eta(x') + \omega(x,z) \right\}}^{\text{First order}} \\ &+ \underbrace{\frac{\varepsilon^2}{2}\left\{ \iint \underbrace{\mathcal{V}(x,x',x'')}_{\substack{\text{2nd-order effect}\\ \text{of distribution}}} h(x')h(x'')d\eta(x')d\eta(x'') + 2\int \underbrace{\Gamma(x,x',z)}_{\substack{\text{Cross effect}\\ \text{of ag. shock.}\\ \text{\& distrib.}}} h(x')d\eta(x') + \underbrace{\Omega(x,z)}_{\substack{\text{2nd-order}\\ \text{effect of}\\ \text{ag. shock}}} \right\}}_{\text{Second order}}. \end{aligned} \tag{31}$$

The structure of the second-order approximation to the value function mirrors that of the first order approximation. The second-order deterministic Impulse Value $\mathcal{V}(x,x'x,'')$ encodes how deviations in the distribution affect values up to second-order. In contrast to the first order, pairwise deviations at x' and x'' now matter. Formally, $\mathcal{V}(x,x',x'') = \frac{\partial^2 V}{\partial^2 g}(x,x',x'',0,g^{SS})$ is the directional Hessian of the value function with respect to the distribution, understood as Fréchet derivatives.

Aggregate shocks matter up to second order as well, as encoded in the second-order stochastic Impulse Value $\Omega(x,z)$. Up to second order, the cross-effect between deviations in the distribution and aggregate shocks also enters in the cross component $\Gamma(x,x',z)$.

4.2 Values and law of motion

I follow the same strategy as in the FAME to characterize the unknown derivatives $\mathcal{V},\Gamma,\Omega$. I substitute the second-order approximation (31) into the nonlinear Master Equation in definition 3, and identify 'coefficients' that are again functions. To keep the exposition simple, I focus on the case of drift control only and no state constraint. It is straightforward to incorporate these additional features. It is useful to define $k(x) = \frac{1}{u''(c^{SS}(x))}$ and $k_p(x) = \frac{u'''(c^{SS}(x))}{u''(c^{SS}(x))^2}$.

To state the SAME as concisely as possible, I define by analogy to the matrix product, for any functions $\varphi(x,x'),\psi(x')$, the operator $\varphi(x,\cdot)[\psi] = \int \varphi(x,x'')\psi(x')dx'$. I further define $\tau(x,x') =$

$\partial_{x'}\Big(k(x')v_{x'}(x,x')g^{SS}(x')\Big)$, $\sigma(x,x') = -\partial_{x'}\left[g^{SS}(x')(b_g(x',x) - \mathcal{M}(x',x,v))\right]$, $\overline{\sigma}(z,x') = -\partial_{x'}\left[g^{SS}(x')(zb_z(x') - \overline{\mathcal{M}}(x',z,\omega))\right]$, as well as the operators $\mathcal{L}_\tau(x) = \mathcal{L}(x) + \tau(x,\cdot)$, $\mathcal{L}_\sigma(x) = \mathcal{L}(x) + \sigma(x,\cdot)$ and $\mathcal{L}_{\overline{\sigma}}(x,z) = \mathcal{L}(x) + \overline{\sigma}(z,\cdot)$.

Theorem 8. *(SAME)*
The deterministic SAME is:

$$\rho\mathcal{V}(x,x',x'') = T_{\mathcal{V}}(x,x',x'') + \mathcal{L}_\tau(x)[\mathcal{V}(\cdot,x',x'')] + \mathcal{L}_\sigma(x')[\mathcal{V}(x,\cdot,x'')] + \mathcal{L}_\sigma(x'')[\mathcal{V}(x,x',\cdot)].$$

The cross SAME is:

$$\rho\Gamma(x,x',z) = T_\Gamma(x,x',z) + \mathcal{L}_\sigma(x)[\Gamma(\cdot,x',z)] + \mathcal{A}(z)[\Gamma(x,x',\cdot)].$$

The stochastic SAME is:

$$\rho\Omega(x,z) \quad = \quad T_S(x,z) + \mathcal{L}_{2\tau}(x)[\Omega(\cdot,z)] + \mathcal{A}(z)[\Omega(x,\cdot)].$$

$T_{\mathcal{V}}, T_\Gamma$ and T_S are independent from $\mathcal{V}, \Gamma, \Omega$ respectively, and are given in equations (38),(39) and (40) respectively in Appendix C.1.

Proof. See Appendix C.1. □

The structure of the SAME for the deterministic Impulse Value in Theorem 8 is analogous to the one found in the deterministic FAME of Theorem 1. When the generator $\mathcal{A}$ scales, it is once more straightforward to guess and verify that the cross SAME scales in z, and hence that $\Gamma(x,x',z) \equiv \Gamma_0(x,x') \cdot z$, where Γ_0 satisfies a simpler cross SAME described in Appendix C.1.1.[35]

Given the second-order Impulse Values, the law of motion of distribution follows immediately.

Theorem 9. *(Law of motion to second order)*
To second order in ε,

$$\frac{dh_t}{dt}(x) = \mathcal{L}^*(x)[h_t] + \mathcal{G}(x)[h_t] + \mathcal{S}(z_t) + \varepsilon\mathcal{Q}(h_t,z_t)$$

where the quadratic form $\mathcal{Q}$ has a closed-form expression given in equation (41), Appendix C.2, that depends on the solutions to the FAME and the SAME.

Proof. See Appendix C.2. □

4.3 Welfare cost of aggregate risk

The SAME can be fruitfully combined with the characterization of the stochastic steady-state in Section 3.6. Together, they determine the welfare cost of aggregate risk.

[35] With some additional structure, the stochastic SAME scales quadratically in z, and so $\Omega(x,z) = z^T \cdot \Omega_0(x) \cdot z$, where Ω_0 satisfies a simpler stochastic SAME as well.

To characterize the welfare consequences of aggregate risk, I integrate the second-order expansion of the value function in equation (31) against the invariant distribution in the stochastic steady-state $\mathcal{P}(dh, dz)$. Denote by $\overline{V}(x,z) = \mathbb{E}_{\mathcal{P}}[V(x,z,\cdot)|z]$ the expected value of agents in the stochastic steady-state. The key observation is that I only need to keep track of the first- and second-order moments of the stochastic steady-state distribution to integrate the second-order Impulse Values. When integrating the first-order Impulse Values, I also need to characterize the second-order expansion of $\mathbb{E}_{\mathcal{P}}[h(x)|z] = \overline{h}(x,z) + \varepsilon\widetilde{h}(x,z)$. I do so using Theorem 9 and the same techniques as in Theorem 6.

Theorem 10. *(Welfare consequences of aggregate risk)*
Under the conditions of Theorems 7 and 8, to a second order,

$$\begin{aligned}\overline{V}(x,z) &= V^{SS}(x) + \varepsilon\left\{\int v(x,x')\overline{h}(x',z)dx' + \omega(x,z)\right\}\\ &+ \frac{\varepsilon^2}{2}\left\{\iint \mathcal{V}(x,x',x'')\overline{H}(x',x'',z)dx'dx'' + 2\int\Gamma(x,x',z)\overline{h}(x',z)dx' + 2\int v(x,x')\widetilde{h}(x',z)dx' + \Omega(z)\right\},\end{aligned}$$

where $\widetilde{h}(x,z)\phi(z)$ satisfies the same equation as in Theorem 6 after replacing $\mathcal{S}$ with $\check{\mathcal{S}}$ given in equation (42), Appendix C.3, which depends on $\mathcal{Q}, \overline{H}$ and $\overline{h}$.

Proof. See Appendix C.3 □

The welfare cost of aggregate risk is given by the first- and second-order changes in the aggregate shocks and in the distribution, weighted by the Impulse Values.

4.4 Solution method

The closed-form expressions in Theorem 8 also deliver an efficient algorithm to compute the solution to the SAME. The determinstic and cross SAMEs in $\mathcal{V}, \mathcal{G}$ are the only non-standard equations to solve in the SAME. The stochastic SAME is a standard matrix equation that can be solved with standard methods. I start by describing the solution method for the deterministic SAME, before turning to the cross and stochastic SAME and impulse response functions.

Denote by $\mathcal{V}_{ijk} \equiv \mathcal{V}(x_i, x_j, x_k)$ the discretization of the second-order deterministic Impulse Value $\mathcal{V}$ on a grid and into a tensor (a three-dimensional array). Inspection of the SAME reveals that $\mathcal{V}$ solves a linear Sylvester tensor equation after discretization:

$$\mathcal{V} \times_1 \mathrm{L}_\tau + \mathcal{V} \times_2 \mathrm{L}_\sigma + \mathcal{V} \times_3 \mathrm{L}_\sigma = \mathrm{T}_{\mathcal{V}}. \tag{32}$$

In equation (32), L_τ and L_σ denote standard square matrices. $\mathrm{T}_{\mathcal{V}}$ denotes a three-dimensional tensor, that is, a three-dimensional array. The matrices and tensor $\mathrm{L}_\tau, \mathrm{L}_\sigma$ and $\mathrm{T}_{\mathcal{V}}$ map into steady-state and first-order objects in closed form using Theorem 8. $\times_i$ denotes the standard tensor product along dimension i, which is a direct generalization of matrix products. For instance, $\left(\mathrm{V} \times_2 \mathrm{L}_\sigma\right)_{ijk} = \sum_\ell \mathrm{V}_{i\ell k}\mathrm{L}_{\sigma,\ell j}$. See Online Appendix F for details and properties of tensor operations.

I seek a solution $\mathcal{V}$ that is also symmetric in its last two coordinates. As a square matrix, L_τ is generically diagonalizable in complex numbers. Let $\mathrm{L}_\tau = P^{-1}\mathcal{D}_\tau P$, where $\mathcal{D}_\tau \equiv \mathbf{diag}(d_{\tau,i})$ is diagonal and possibly complex. P is invertible and also possibly complex. Using the results in Online Appendix F, I transform equation (32) into:

$$E = W \times_1 \mathcal{D}_\tau + W \times_2 \mathrm{L}_\sigma + W \times_3 \mathrm{L}_\sigma.$$

where I define $W = \mathcal{V} \times_1 P$ as the new unknown and $E = D \times_1 P$. W and E are still symmetric in their last two coordinates. Evaluating at $(i, \bullet, \bullet)$, I obtain:

$$E_{i\bullet\bullet} = C_i W_{i\bullet\bullet} + W_{i\bullet\bullet}(C_i') \quad , \quad C_i \equiv \frac{d_{\tau,i}}{2}\mathrm{Id} + \mathrm{L}_\sigma, \tag{33}$$

where C_i' denotes the matrix transpose of C_i. For each i, equation (33) defines a standard Sylvester matrix equation. Solving a Sylvester equation relies on an underlying Schur or diagonal decomposition of the left- and right-multiplying matrices, C_i and C_i'. Crucially, I only need to compute the diagonal form of L_σ once, and it delivers the diagonal form of each C_i.

Specifically, denote the diagonal form of L_σ by $\mathrm{L}_\sigma = Q^{-1}\mathcal{D}_\sigma Q$ if it exists, with $\mathcal{D}_\sigma = \mathbf{diag}(d_{\sigma,i})$ diagonal. Then $C_i = Q^{-1}\mathcal{C}_i Q$, where $\mathcal{C}_i = \mathcal{D}_\sigma + \frac{d_{\tau,i}}{2}\mathrm{Id}$ is diagonal, and $C_i' = Q'\mathcal{C}_i(Q')^{-1}$. The Sylvester equation (33) then rewrites in diagonal form:

$$F_i = \mathcal{C}_i X_i + X_i \mathcal{C}_i \Longrightarrow X_{i;jk} = \frac{F_{i;jk}}{d_{\tau,i} + d_{\sigma,j} + d_{\sigma,k}}. \tag{34}$$

where the unknown is a matrix $X_i = QW_{i\bullet\bullet}Q'$ and $F_i = QE_{i\bullet\bullet}Q'$ for each i, and last equality obtains after evaluating at (j,k). Equation (34) also provides a sufficient condition for the discretized deterministic SAME to have a solution: when $\mathrm{L}_\tau, \mathrm{L}_\sigma$ admit a diagonal form, and $d_{\tau,i} + d_{\sigma,j} + d_{\sigma,k} \neq 0$ for all i, j, k.

Corollary 2. *(Numerical implementation of the deterministic SAME)*
Define grids $\{x_i\}_{i=1}^{I}$. Define the matrices $\mathrm{L}_\tau, \mathrm{L}_\sigma$ and the tensor $\mathrm{T}_\mathcal{V}$ based on Theorem 8. Then:

1. *Compute the possibly complex diagonal form of L_τ, $\mathrm{L}_\tau = P^{-1}\mathcal{D}_\tau P$. Compute the possibly complex diagonal form of L_σ: $\mathrm{L}_\sigma = Q^{-1}\mathcal{D}_\sigma Q$.*

2. *For each i, obtain X_i by solving equation (34).*

3. *Recover $W_{i\bullet\bullet} = Q^{-1}X_i(Q')^{-1}$ and $\mathcal{V} = W \times_1 P^{-1}$.*

Having established how to compute the solution to the deterministic SAME, I turn to the cross and stochastic SAME. Theorem 8 reveals that, once discretized, the cross SAME also satisfies a Sylvester tensor equation:

$$\Gamma \times_1 \mathrm{L}_\tau + \Gamma \times_2 \mathrm{L}_\sigma + \Gamma \times_3 \mathrm{A} = \mathrm{T}_\Gamma.$$

The alogrithm in Corollary 2 does not rely on the symmetry of the deterministic SAME. It is straightforward to adapt it to solve for the cross SAME. When the cross SAME scales with z, $\Gamma(x, x', z) =$

$\Gamma_0(x,x')z$, the discretized cross SAME boils down to a linear Sylvester equation $(\mathrm{L}_\tau+\mathrm{B})\Gamma+\Gamma\mathrm{L}_\sigma=\mathrm{T}_\Gamma$, which can be solved directly. The discretized stochastic SAME takes the form of a standard Bellman equation and can thus also be solved with standard methods.

Together, Theorems 8 and 9 as well as Corollary 2 demonstrate that taking second-order perturbations of the Master Equation is conceptually no more difficult than taking first-order perturbations. This property of the Master Equation approach makes it particularly well-suited to studying settings with aggregate shocks in which nonlinearities and aggregate risk matter, as well as environments with asset pricing.

5 Applications

5.1 The welfare cost of business cycles with incomplete markets

In this section I solve the SAME in the Krusell and Smith (1998) model of Section 1. I enrich the economy with two key features: I introduce an occasionally binding borrowing contraint, and I add countercyclical income risk. I characterize the welfare cost of business cycles.

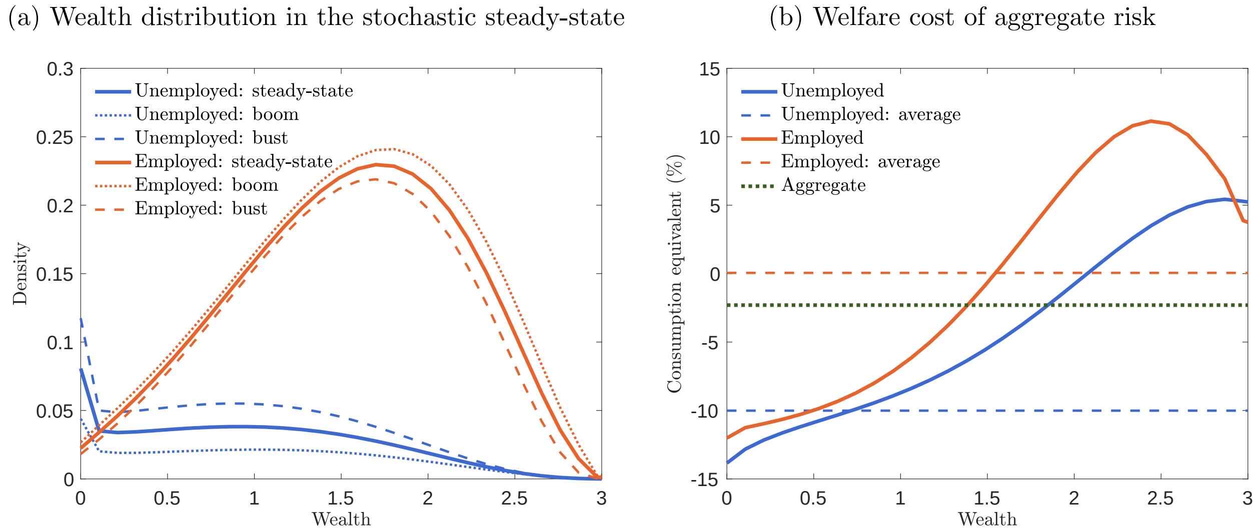
I derive the FAME using Theorems 1 and 2. I obtain the SAME using Theorem 8. The law of motion of the distribution follows from Theorems 3 and 9. I characterize the first and second moments of the stochastic steady-state distribution using Theorems 6 and 7. I calculate the welfare cost of business cycles using Theorem 10.

I calibrate the economy at an annual frequency. The discount rate is set to $\rho = 0.05$, and risk-aversion to $\gamma = 2$. The two-state income process captures employment and non-employment. Income during non-employment is half of income during employment. The baseline transition rate out of non-employment is $\lambda_1 = 0.5$, and the transition rate out of employment is $\lambda_2 = 0.1$, which leads to a non-employment rate of 17%. These rates are cyclical: $\lambda_{it} = \lambda_i + \lambda_i' z_t$: a 10% aggregate shock doubles the out-flow rate from non-employment, and divides in half the inflow rate into non-employment. The borrowing constraint is set to $\underline{a} = 0$. Aggregate shocks have a half-life of 1 year so that $\theta = 0.69$. The labor share is $\alpha = 0.3$. I also let capital depreciate at 10% annually, and introduce foreign capital owners that hold a quantity K^f of the domestic capital stock.

I target two key statistics to calibrate the volatility of aggregate shocks σ and foreign holdings of capital K^f. I depart from Krusell et al. (2009) and use a high-MPC, low-liquidity calibration. I target a steady-state average MPC of 0.2. I achieve this value by adjusting the stock of foreign-held capital to $K^f = 4.4$. Finally, I obtain $\sigma = 0.06$ by targeting the volatility of aggregate consumption of 0.032 used in Lucas (1987). The computational speed inherent to the FAME and the SAME facilitates calibration: solving the FAME takes 0.23 seconds and solving the SAME takes 4.3 seconds on a laptop with 100 grid points for assets.[36]

[36] All calculations are run on a MacBook Pro laptop with 2.4 GHz 8-Core Intel Core i9 processors and 64 GB 2667 MHz DDR4 memory.

Figure 1: Wealth distribution and welfare cost of risk



Note: Panel (a): invariant distribution in the steady-state g^{SS} (solid lines), and invariant distribution $\mathbb{E}_{\mathcal{P}}[h|z]$ in the stochastic steady-state (dashed and dotted lines). Wealth expressed in multilpes of steady-state average annual labor market income. Boom and bust correspond to a one standard deviation aggregate shock $z = \pm 1\sigma$. Panel (b): welfare cost of aggregate risk in consumption-equivalent percent, by wealth, income and in the aggregate.

Figure 1(a) shows the invariant distribution in the stochastic steady-state. Contrary to Krusell et al. (2009), there is no need to simulate the model for many periods to calculate it: Theorem 6 provides a direct formula. Due to countercyclical income risk, the share of employed individuals rises and the share of unemployed individuals falls in booms. Absent countercyclical income risk, the cyclical movements in the distribution are much less pronounced because the consumption-savings decision alone leads to slow internal transitional dynamics. With countercyclical income risk however, the distribution moves substantially over the business cycle and hence all the components in Theorem 10 are necessary to accurately evaluate the cost of business cycles.

Figure 1(b) displays the cost of aggregate risk by wealth and labor market status. Again, there is no need to simulate the model for many periods: Theorem 10 directly delivers the relevant expression.

The aggregate cost of business cycles is 2.3% of steady-state aggregate consumption, 23 times the value reported in Lucas (1987)'s seminal contribution. Both cyclical income risk and high MPCs play a role. Without cyclical income risk, the welfare cost of business cycles is roughly half, at 1.3%. The absence of Arrow-Debreu credit markets delivers the rest of the amplification. Quantitatively, these magnitudes are also twice as large than those in Krusell et al. (2009) because the high-MPC, low-liquidity calibration further amplifies the cost of business cycle risk. This amplification is particularly salient at the bottom of the wealth distribution.

Figure 1(b) shows that for low-wealth individuals, the cost of business cycles is between 12% and 14% in consumption-equivalent terms. These individuals face particularly high costs of business cycles because they have a limited ability to self-insure against income shocks. These costs are about halved without countercyclical income risk. Because non-employed individuals have lower wealth than employed individuals, as a group they face a cost of business cycles of 10%. By contrast, employed individuals do not gain or lose substantially on average, although there is substantial heterogeneity by wealth.

These results highlight how to use the FAME and the SAME to efficiently characterize the role of aggregate risk in the workhorse incomplete credit market heterogeneous agent model. The next section shows how to use the FAME and the SAME for a different class of heterogeneous agent models: dynamic spatial models.

5.2 A dynamic migration model with aggregate shocks

In this section I solve the SAME in the dynamic migration model of Section 1.6 and characterize the welfare cost of risk for each location in a calibrated version to 381 Metropolitan Statistical Areas (MSAs).

As in Section 4, I seek a solution to the second-order approximation of the value function, similarly to (31):

$$V = V^{SS} + \varepsilon(vn + z\omega) + \frac{\varepsilon^2}{2}\left(\mathcal{V} \times_2 n^* \times_3 n^* + 2z\Gamma n + \Omega(z)\right).$$

$\mathcal{V} \in \mathbb{R}^{I\times I\times I}$ denotes the second-order deterministic Impulse Value and is a 3-dimensional tensor. $\Gamma \in \mathbb{R}^{I\times I}$ denotes the second-order cross Impulse Value and is a square matrix. Ω denotes the second-order stochastic Impulse Value, and will take the form $\Omega(z) = z^2\Delta + \Lambda$, with $\Delta, \Lambda \in \mathbb{R}^I$ vectors.

As in Section 4, the SAME takes the form of a sequence of matrix equations. Appendix D.3 details the derivations and results. The deterministic SAME takes the form of a tensor-valued Sylvester equation in the unknown $\mathcal{V}$:

$$\mathcal{V} \times_1 A + \mathcal{V} \times_2 B + \mathcal{V} \times_2 B = D, \tag{35}$$

where A, B are known square matrices and D is a known three-dimensional tensor. I use the algorithm from Corollary 2 to solve the deterministic SAME.

Having solved the deterministic SAME, the cross SAME takes the form of a standard Sylvester matrix equation in the unknown Γ:

$$(A - \theta \mathrm{Id})\Gamma + \Gamma B^* = C,$$

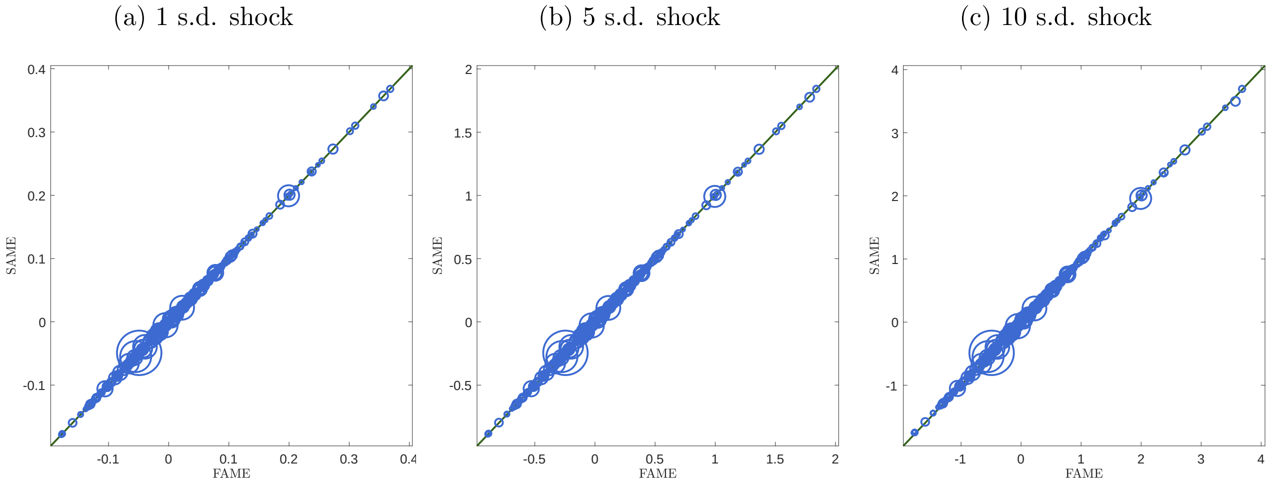
where C denotes a known matrix. As the FAME, the cross SAME can thus be solved with standard Sylvester equation routines.

Having solved the cross SAME, the stochastic SAMEs for Δ and Λ take the form of standard vector-valued equations that may be solved in sequence:

$$(A - 2\theta \mathrm{Id})\Delta = E \quad , \quad A\Lambda = -\sigma^2\Delta,$$

where E is a known vector. To compute an impulse response in this economy, I use the second-order

Figure 2: Population change 5 years out in the FAME and the SAME (%)



Note: Each blue circle represents an MSA. Both axes represent percent deviations from steady-state 5 years after the shock first occurs. The size of circles is proportional to steady-state population. Panel (a) compares the FAME and the SAME following a time-0 1 standard deviation shock $z_0 = \sigma$ that mean-reverts according to $dz_t = -\theta z_t dt$. Panel (b) increase the initial value to $z_0 = 5\sigma$, and panel (c) to $z_0 = 10\sigma$.

expansion of the law of motion that underlies the derivation of the SAME and is detailed in Appendix D.3.8.

I calibrate the model to 381 MSAs of the US economy.[37] To calibrate shifters Z_i and H_i, I follow the inversion procedure in Bilal and Rossi-Hansberg (2023) and use the same preference and technology parameters except risk-aversion which is set to $\gamma = 2$ as in the previous example. I estimate the local exposure coefficients χ_i by indirect inference: I match the regression coefficient of local output on aggregate output. I set the standard deviation of aggregate shocks to 0.06 as in the previous example, and set and its mean-reversion θ such that the half-life of a shock is 1 year.

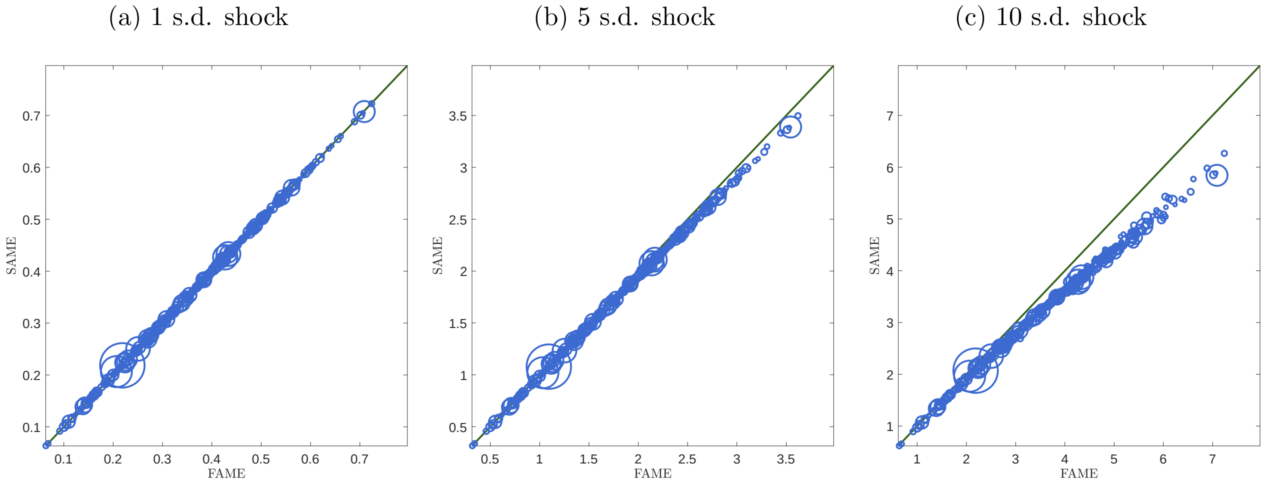
Solving for the steady-state takes 1.2 seconds on a laptop. Solving the FAME takes 0.2 seconds, and solving the SAME takes 2.1 seconds with the algorithm in Section 4.4. Computing an impulse response function with 100 periods in the FAME takes 0.2 seconds, and in the SAME it takes 4.2 seconds.

Figure 2 compares the population response in the FAME and the SAME for three different aggregate shocks in all MSAs. All the shocks follow the aggregate shock's stochastic process. The only difference is the initial magnitude of the shock. The response of population 5 years after the shock in all locations is strikingly similar in the FAME and the SAME for an aggregate shock of 5 standard deviations—an extreme event since it occurs less frequently than every million years. Even for an aggregate shock of 10 standard deviations (that has probability equal to 0 up to machine precision), the response of population in the SAME is very close to the FAME. Thus, for usual aggregate productivity shocks, the FAME provides a good approximation to the SAME for population responses.

Next, Figure 3 compares consumption-equivalent welfare changes in the FAME and the SAME for

[37] To include all of the US economy, I attribute any county initially not in an MSA to its nearest MSA.

Figure 3: Initial welfare impact in the FAME and the SAME (consumption-equivalent %)



Note: Each blue circle represents an MSA. Both axes represent percent deviations from steady-state when the shock first occurs. The size of circles is proportional to steady-state population. Panel (a) compares the FAME and the SAME following a time-0 1 standard deviation shock $z_0 = \sigma$ that mean-reverts according to $dz_t = -\theta z_t dt$. Panel (b) increase the initial value to $z_0 = 5\sigma$, and panel (c) to $z_0 = 10\sigma$.

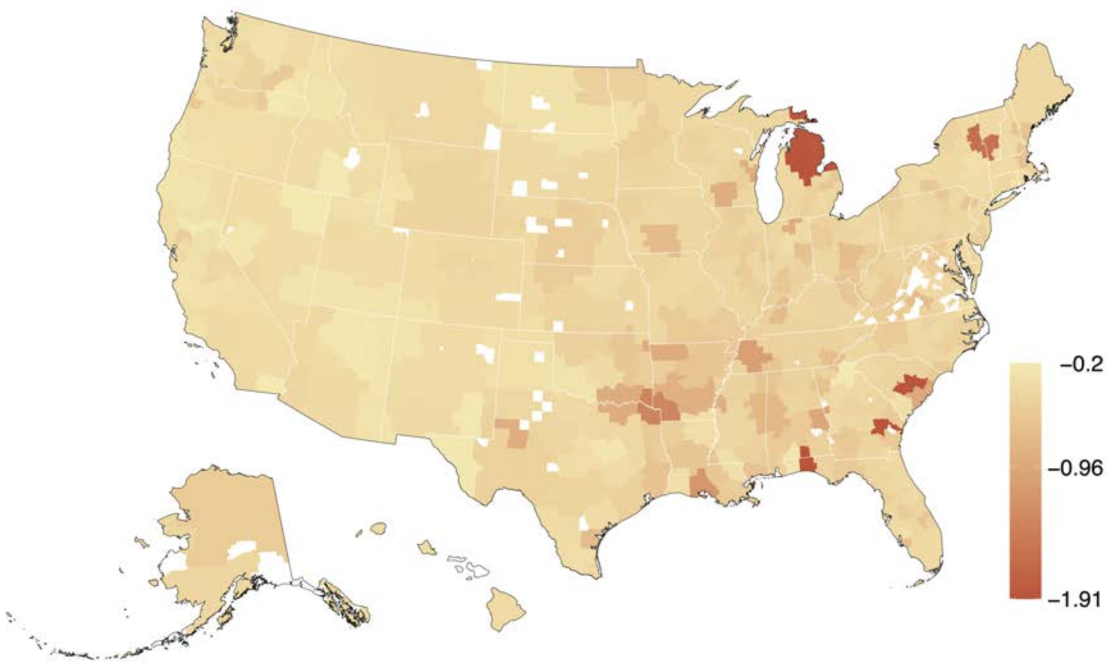
the same three aggregate shocks at impact. The welfare impact of aggregate shocks is remarkably similar for aggregate shocks of 1 standard deviation. Even for an extreme 1-in-a-million-years, 5 standard deviation shock, the welfare change in the SAME remains very close to the FAME. For a 10 standard deviation shock, the welfare change in the SAME starts to deviate markedly from the FAME. However, such a large shock would virtually never materialize in the stochastic steady-state of the economy.

Hence, is the FAME always sufficient to understand this economy for usual aggregate productivity shocks? Figures 2 and 3 indicate that for population and welfare changes in response to a particular shock, the FAME may be enough given the calibration. However, the FAME always misses an important time-invariant margin: the value of aggregate risk. The value of aggregate risk for location i is encapsulated in the stochastic SAME through the intercept Λ_i.

Figure 4 displays the consumption-equivalent value of aggregate risk for every MSA in the US. To benchmark this value, I map it as a fraction of the welfare impact of an 1 standard deviation aggregate productivity shock. Throughout the US, the cost of aggregate risk ranges from 20% to 191% of the value of a 1 standard deviation positive aggregate shock. Thus, the SAME is crucial to accurately assess the welfare cost of aggregate risk.[38]

[38] Compared to the previous example, this spatial economy lacks individuals whose consumption can fluctuate by as much as 30% over the business cycle. Here, consumption fluctuations are closer to 5% despite the lack of a savings technology. Thus, the welfare cost of risk is closer to the magnitude found in Lucas (1987)'s calculation.

Figure 4: Welfare cost of aggregate risk



Note: Welfare cost of aggregate risk Λ_i, represented in consumption-equivalent percent, as a fraction of the consumption-equivalent welfare impact of a 1 standard deviation aggregate shock for the same MSA i. Counties in white: missing data. Cost of risk winsorized at 1% and 99% for display.

Conclusion

This paper proposes a new representation of dynamic general equilibrium economies with cross-sectional heterogeneity. By treating the underlying distribution as an explicit state variable in decision-makers' problem, the economy becomes fully recursive and is characterized by the Master Equation. I show that local perturbations of the Master Equation in aggregates deliver interpretable, block-recursive and easily computable representations of equilibrium: the FAME and the SAME. They further deliver stability and convergence results, together with a characterization of the stochastic steady-state.

I highlighted the versatility of the FAME and the SAME in two heterogeneous agent economies: an incomplete credit market model and dynamic spatial model. The FAME and the SAME apply to many more economic settings. For instance, the FAME could be used to study the impact of aggregate shocks in job ladder models of the labor market. The SAME could be used to investigate asset pricing with cross-sectional heterogeneity, as well as the cost of climate change risk in multi-location settings.

References

Achdou, Yves, Han, Jiequn, Lasry, Jean-Michel, Lions, Pierre-Louis, and Moll, Benjamin (2021). "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach". *Forthcoming, Review of Economic Studies*.

Ahn, SeHyoun, Kaplan, Greg, Moll, Benjamin, Winberry, Thomas, and Wolf, Christian (2018). "When Inequality Matters for Macro and Macro Matters for Inequality". *NBER Macroeconomics Annual* 32, pp. 1–75.

Aiyagari, S. Rao (1994). "Uninsured Idiosyncratic Risk and Aggregate Saving". *The Quarterly Journal of Economics* 109.3, pp. 659–684.

Aliprantis, Charalambos D. and Border, Kim C. (2006). *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer.

Alvarez, Fernando E and Lippi, Francesco (2021). *The Analytic Theory of a Monetary Shock*. Working Paper 28464. National Bureau of Economic Research.

Auclert, Adrien, Bardóczy, Bence, Rognlie, Matthew, and Straub, Ludwig (2021). "Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models". *Econometrica* 89.5, pp. 2375–2408.

Azinovic, Marlon, Gaegauf, Luca, and Scheidegger, Simon (2022). "Deep Equilibrium Nets". *International Economic Review* 63.4, pp. 1471–1525.

Bhandari, Anmol, Bourany, Thomas, Evans, David, and Golosov, Mikhail (2023). "A Perturbational Approach for Approximating Heterogeneous-Agent Models". *Working Paper*.

Bhandari, Anmol, Evans, David, Golosov, Mikhail, and Sargent, Thomas J. (2021). "Inequality, Business Cycles, and Monetary-Fiscal Policy". *Econometrica* 89.6, pp. 2559–2599.

Bilal, Adrien (2023). "Some Pleasant Sequence-Space Arithmetic in Continuous Time". *Working Paper*.

Bilal, Adrien and Rossi-Hansberg, Esteban (2023). "Anticipating Climate Change Across the United States". *National Bureau of Economic Research Working Paper Series 31323*.

Blanchard, Olivier Jean and Kahn, Charles M. (1980). "The Solution of Linear Difference Models under Rational Expectations". *Econometrica* 48.5, pp. 1305–1311.

Bogachev, Vladimir I., Krylov, Nicolai V., Rockner, Michael, and Shaposhnikov, Stanislav V. (2015). *Fokker–Planck–Kolmogorov Equations*. American Mathematical Society.

Boppart, Timo, Krusell, Per, and Mitman, Kurt (2018). "Exploiting MIT shocks in heterogeneous-agent economies: the impulse response as a numerical derivative". *Journal of Economic Dynamics and Control* 89. Fed St. Louis-JEDC-SCG-SNB-UniBern Conference, titled: "Fiscal and Monetary Policies"., pp. 68–92.

Brumm, Johannes and Scheidegger, Simon (2017). "Using Adaptive Sparse Grids to Solve High-Dimensional Dynamic Models". *Econometrica* 85.5, pp. 1575–1612.

Cardaliaguet, Pierre, Delarue, Francois, Lasry, Jean-Michel, and Lions, Pierre-Louis (2019). *The Master Equation and the Convergence Problem in Mean Field Games.* ISBN: 9780691190709.

Carmona, René and Delarue, Francois (2018a). *Probabilistic Theory of Mean Field Games with Applications I: Mean Field FBSDEs, Control and Games.* Springer.

— (2018b). *Probabilistic Theory of Mean Field Games with Applications II: Mean Field Games with Common Noise and Master Equations.* Springer.

Childers, David (2018). "Solution of Rational Expectations Models with Function Valued States". *Working Paper.*

Fleming, Wendell H. and Soner, Halil Mete (2006). *Controlled Markov Processes and Viscosity Solutions.* Springer.

Gabaix, Xavier, Lasry, Jean-Michel, Lions, Pierre-Louis, and Moll, Benjamin (2016). "The Dynamics of Inequality". *Econometrica* 84.6, pp. 2071–2111.

Hairer, 2021 (2021). "Convergence of Markov Processes". *Lecture Notes.*

Huang, Minyi, Malhamé, Roland P., and Caines, Peter E. (2006). "Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle". *Communications in Information & Systems* 6.3, pp. 221 –252.

Kahou, Mahdi Ebrahimi, Fernández-Villaverde, Jesús, Perla, Jesse, and Sood, Arnav (2021). *Exploiting Symmetry in High-Dimensional Dynamic Programming.* Working Paper 28981. National Bureau of Economic Research.

Krusell, Per, Mukoyama, Toshihiko, Şahin, Ayşegül, and Smith, Anthony A. (2009). "Revisiting the welfare effects of eliminating business cycles". *Review of Economic Dynamics* 12.3, pp. 393–404.

Krusell, Per and Smith Jr., Anthony A. (1998). "Income and Wealth Heterogeneity in the Macroeconomy". *Journal of Political Economy* 106.5, pp. 867–896.

Lasry, Jean-Michel and Lions, Pierre-Louis (2006). "Jeux à champ moyen. I – Le cas stationnaire". *Comptes Rendus Mathematique* 343.9, pp. 619–625.

Lions, Pierre-Louis (2011). "Cours au Collège de France, 2007-2012".

Liu, Ernest and Tsyvinski, Aleh (2020). *Dynamical Structure and Spectral Properties of Input-Output Networks.* Working Paper 28178. National Bureau of Economic Research.

Lucas, Robert E (1987). *Models of business cycles.* Oxford: Basil Blackwell.

Mertens, Thomas M. and Judd, Kenneth L. (2017). *Solving an Incomplete Markets Model with a Large Cross-Section of Agents.* Working Paper. SSRN.

Meyn, Sean P. and Tweedie, R. L. (1993a). “Stability of Markovian Processes III: Foster-Lyapunov Criteria for Continuous-Time Processes”. *Advances in Applied Probability* 25.3, pp. 518–548. (Visited on 09/27/2023).

Meyn, Sean P. and Tweedie, Richard L. (1993b). *Markov Chains and Stochastic Stability*. Springer.

Pazy, Amnon (1983). *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer.

Preston, Bruce and Roca, Mauro (2007). *Incomplete Markets, Heterogeneity and Macroeconomic Dynamics*. Working Paper 13260. National Bureau of Economic Research.

Reiter, Michael (2009). “Solving heterogeneous-agent models by projection and perturbation”. *Journal of Economic Dynamics and Control* 33.3, pp. 649–665.

Schaab, Andreas (2021). “Micro and Macro Uncertainty”.

Stokey, Nancy L., Lucas, Robert E., and Prescott, Edward C. (1989). *Recursive Methods in Economic Dynamics*. Harvard University Press. ISBN: 9780674750968. (Visited on 09/24/2023).

Appendix

A Notation

A.1 Additional notation for Section 1

Define $Y = \iint y g^{SS}(a,y) da dy$ and $A = \iint a g^{SS}(a,w) da dy$, and:

$$\mathcal{R}_0 = -\alpha(1-\alpha)\bar{Z}\left(\frac{Y}{A}\right)^{1-\alpha} A^{-1} \qquad \mathcal{R}_1 = \alpha(1-\alpha)\bar{Z}\left(\frac{Y}{A}\right)^{1-\alpha} Y^{-1}$$
$$\mathcal{W}_0 = \alpha(1-\alpha)\bar{Z}\left(\frac{A}{Y}\right)^{\alpha} A^{-1} \qquad \mathcal{W}_1 = -\alpha(1-\alpha)\bar{Z}\left(\frac{A}{Y}\right)^{\alpha} Y^{-1}.$$

Define also

$$\mathcal{R}_2 = \alpha A^{\alpha-1} Y^{1-\alpha} \qquad \mathcal{W}_2 = (1-\alpha) A^{\alpha} Y^{-\alpha}.$$

A.2 Additional notation for the FAME

Denote

$$\mathcal{L}(x) = L(x, c^{SS}(x), V^{SS}, g^{SS}), \qquad \mathcal{C}(x) = \mathcal{C}\left(x, V^{SS}, \frac{\partial V^{SS}}{\partial x}, g^{SS}\right))$$
$$u_g(x,y) = \frac{\partial u}{\partial g}(x, y, V^{SS}, g^{SS}), \qquad u_V(x,y) = \frac{\partial u}{\partial V}(x, y, V^{SS}, g^{SS})$$
$$\mathcal{L}_g(x,y) = \frac{\partial L}{\partial g}(x, y, c^{SS}(x), V^{SS}, g^{SS}), \qquad \mathcal{L}_V(x,y) = \frac{\partial L}{\partial V}(x, y, c^{SS}(x), V^{SS}, g^{SS})$$
$$\mathcal{C}_g(x,y) = \frac{\partial C}{\partial g}(x, y, V^{SS}, \frac{\partial V^{SS}}{\partial x}, g^{SS}), \qquad \mathcal{C}_V(x,y) = \frac{\partial C}{\partial V}(x, y, V^{SS}, \frac{\partial V^{SS}}{\partial x}, g^{SS})$$
$$\mathcal{C}_{\partial V}(x,y) = \frac{\partial C}{\partial p}(x, y, V^{SS}, p^{SS}, g^{SS})\Big|_{p=\frac{\partial V^{SS}}{\partial x}}$$

the Fréchet differentials of the flow payoff u, the generator L and the state constraint $\mathcal{C}$ with respect to either the distribution or the value. $\partial/\partial p$ denotes the derivative of $\mathcal{C}$ with respect to $\partial V/\partial x$.

B Proofs for Section 3

B.1 Proof of Proposition 2

B.1.1 Special case

To gain intuition, it is useful to specify additional structure that is often satisfied in applications. Denote by L^{diff} the diffusion part of the generator L, and L^{int} the integral part.

Assumption 5. *(Payoff and generator structure)*
There are two sets of indices E_1, E_2 that partition $\{1, ..., D_X\}$ such that for $i \in E_1$, $b_i(x, g, c, V) =$

$b_i(x,g,V)-c_i$, *and for* $i\in E_2$, $L_i^{int}(x,c,g,V)=c_iL^{int}(x,g,V)$. *In addition,* $u(x,c,g,V)=u_0(x,g,V)+\sum_{i\in E_1}u_{1i}(c_i)-\sum_{i\in E_2}u_{2i}(c_i)$. u_{ki} *are strictly concave and satisfy Inada conditions, and* $\bar\Gamma_i=[0,+\infty)$.

Assumption 5 ensures that Assumption 4 holds, and that

$$\forall i\in E_1,\quad \hat c_i=(u'_{1i})^{-1}\left(\frac{\partial V}{\partial x_i}(x)\right),\qquad\qquad \forall i\in E_2,\quad \hat c_i=(u'_{2i})^{-1}\left(L_i^{int}(x,g,V)[V]\right).$$

Denote $U_{1i}(x)=\frac{1}{u''_{1i}(\partial_{x_i}V^{SS}(x))}$ and $U_{2i}(x)=\frac{1}{u''_{2i}(L_i^{int}(x,g^{SS},V^{SS}[V^{SS}])}$. Armed with this notation, I obtain the first-order perturbation in individuals' optimal control in response to a small distributional impulse.

Proposition 4. *(Optimal control)*
Under Assumption 5,

$$\mathcal{M}_i(x,x',v)\quad=\quad\begin{cases}U_{1i}(x)\partial_{x_i}v(x,x'), & i\in E_1\\ U_{2i}(x)\left[\mathcal{L}_i^{int}(x)[v(\cdot,x')]+\mathcal{L}_{g,i}^{int}(x,x')[V^{SS}]+\int\mathcal{L}_{V,i}^{int}(x,y)[V^{SS}]\ v(y,x')d\eta(y)\right], & i\in E_2.\end{cases}$$

Proof. The linearized FOCs are

$$\begin{aligned}\forall i\in E_1,\quad dc_i(x)\quad&=\quad\frac{1}{u''_{1i}(c_i^{SS}(x))}\partial_{x_i}dv(x)\\ \forall i\in E_2,\quad dc_i(x)\quad&=\quad\frac{1}{u''_{2i}(c_i^{SS}(x))}\left(\int h(x')d\eta(x')\partial_g\mathcal{L}_i^{int}(x,x')+\int dv(y)d\eta(y)\partial_V\mathcal{L}_i^{int}(x,y)\right)\end{aligned}$$

where now c_i denote steady-state controls. Substituting out dv yields the result. □

B.1.2 General case

I now prove Proposition 2 in the general case. The first-order optimality condition is

$$u_c(x,\hat c,V,g)+L_c(x,\hat c,g,V)[V]=0.$$

Totally differentiating this condition, I obtain

$$\langle u_{cg},h\rangle+u_{cc}\hat c+\langle u_{cV},\langle v,h\rangle\rangle+\mathcal{L}_{cc}[V^{SS}]\hat c+\langle\mathcal{L}_c[v],h\rangle+\left\langle\mathcal{L}_{cg}[V^{SS}],h\right\rangle+\left\langle\mathcal{L}_{cV}[V^{SS}],\langle v,h\rangle\right\rangle$$

Hence,

$$\hat c(x)=\mathcal{U}(x)^{-1}\left(\langle u_{cg},h\rangle+\left\langle\mathcal{L}_{cg}[V^{SS}],h\right\rangle+\langle\mathcal{L}_c[v],h\rangle+\langle u_{cV},\langle v,h\rangle\rangle+\left\langle\mathcal{L}_{cV}[V^{SS}],\langle v,h\rangle\right\rangle\right)$$

Writing $\hat c(x)=\int\mathcal{M}(x,x',v)h(x')d\eta(x')$ and identifying coefficients delivers the equation in Proposition 2.

B.2 Proof of Theorem 1

I focus on the deterministic FAME, since the stochastic FAME follows from the same derivations. Since none of the forcing terms in u,L depend directly on the derivatives of g, it is enough to look for a Fréchet derivative of V that only loads on h. To ease notation, denote partial derivatives by $\partial_g,\partial_c,\partial_V$, or directly with the corresponding subscript. Also denote $X \cdot Y \equiv \langle X, Y \rangle$. Since c is always interior, taking the FOC and substituting back into the Master Equation in Definition 3, I obtain

$$\underbrace{\rho V(x,g)}_{\equiv M_0} = \underbrace{u(\hat{c}(x,V,g),x,g,V)}_{\equiv M_1} + \underbrace{L(x,\hat{c}(x,V,g),g,V)[V]}_{\equiv M_2} + \underbrace{\int \partial_g V(x,x'',g) L^*(x'',\hat{c}(x'',V,g),g,V)[g] d\eta(x'')}_{\equiv M_3}$$

I will sometimes omit arguments of function to ease notation. In that case, they are evaluated at $x, g, V, \hat{c}(x,V,g)$. I sometimes denote by $dv(x) = \int v(x,x')h(x')d\eta(x')$ and $dc(x)$ the first-order change in controls. I now expand each component of the Master Equation up to first order.

M_0. The first-order contribution of an impulse in h to M_0 is

$$M_0 = \rho \int v(x,x')h(x')d\eta(x').$$

M_1. The flow gain is, up to first order,

$$\begin{aligned} M_1 &= \int u_g(x,x')h(x')d\eta(x') + u_c(x) \cdot dc(x) + \int u_V(x,y)dv(y)d\eta(y) \\ &= \int u_g(x,x')h(x')d\eta(x') + u_c(x) \cdot dc(x) + \int \left(\int u_V(x,y)v(y,x')d\eta(y) \right) h(x')d\eta(x'). \end{aligned}$$

Now note that u depends on V partly through c. So the FOC implies, for all g, V,

$$0 = u_c(x) + L_c(x, dc(x))[V]$$

M_2. The continuation value term is

$$ME2 = L_c(x,dc(x))[V] + \int L_g(x,x')[V]h(x')d\eta(x') + \int L_V(x,y)[V]dv(y)d\eta(y) + L(x)[dv].$$

Passing the operator $\mathcal{L}(x)$ inside the integral for the last term,

$$L(x)[dv] = \int L(x)[v(\cdot,x')]h(x')d\eta(x')$$

Similarly,

$$\int L_V(x,y)[V]dv(y)d\eta(y) = \int \left(\int L_V(x,y)[V]v(y,x')d\eta(y) \right) h(x')d\eta(x')$$

$M_1 + M_2$. An envelope argument obtains where the contributions of ∂_c cancel out:

$$\begin{aligned} M_1 + M_2 = \int \Big\{ & u_g(x,x')h(x') + \int u_V(x,y)v(y,x')d\eta(y) + L_g(x,x')[V] \\ & + \int L_V(x,y)[V]v(y,x')d\eta(y) + L(x)[v(\cdot,x')] \Big\} h(x')d\eta(x') \end{aligned}$$

***M*₃.** To expand the continuation value from changes in the distribution, I use Proposition 2. Hence:

$$
\begin{aligned}
M_3 \;=\; & \int \Big(v(x,x'') + \mathcal{O}(||h||)\Big)\bigg(L^*(x'')[g] + \int L_g^*(x'',x')[g]h(x')d\eta(x') \\
& \qquad + \iint L_V^*(x'',y)[g]v(y,x')h(x')d\eta(y)d\eta(x') \\
& \qquad + L_c^*\left(x'', \int \mathcal{M}(x,x',v)h(x')d\eta(x')\right)[g] + L^*(x'')[h]\bigg)d\eta(x'')
\end{aligned}
$$

So far, I have considered a local perturbation around any point. To make progress, I now make use of the observation that, in steady-state, $\mathcal{L}^*(x'')[g] = 0$ for all x'' by definition. Thus, I can negelect the $\mathcal{O}(||h||)$ term to first order.

I also make use of the adjoint property between L and L^* developed in Online Appendix E.2: $\int L(x)[\phi]\psi(x)d\eta(x) = \int \phi(x)L^*(x)[\psi]d\eta(x) + \bar{B}$, where $\bar{B}$ is an integral over a measure $\bar{\eta}$ that only loads on the mass points in $\mathcal{B}$. Thus, I can express the last term as an integral over h. Together, these observations imply that to first order around a steady-state:

$$
\begin{aligned}
M_3 \;=\; & \int \bigg\{\mathcal{L}(x')[v(x,\cdot)] + \int v(x,x'')\bigg(\mathcal{L}_g^*(x'',x')[g^{SS}] + \int \mathcal{L}_V^*(x'',y)[g^{SS}]v(y,x')d\eta(y) \\
& \qquad + \mathcal{L}_c^*(x'', \mathcal{M}(x'',x',v))[g^{SS}]\bigg)d\eta(x'')\bigg\}h(x')d\eta(x') + \bar{B}
\end{aligned}
$$

Deriving the FAME. Equate $M_0 = M_1 + M_2 + M_3$ at $g = g^{SS} + h$. This equality must hold for all functions h conditional on x. Hence, the associated vector that integrates against $h(x')$ must be zero in L^2 ("identifying coefficients"). This identity delivers the FAME in Theorem 1. Crucially, one can only "identify coefficients" for $x \in X\backslash\mathcal{B}$: on $\mathcal{B}$, the additional components in $\bar{B}$ imply that the associated linear form is not an integral over $hd\eta$.

State constraint. Expanding the state constraint for $x \in B$, I obtain for all h:

$$
\mathcal{C}(x) + \int \Gamma(x,x')h(x')d\eta(x') \geq 0 \quad , \quad \Gamma(x,x') = \mathcal{C}_g(x,x') + \mathcal{C}_V(x)v(x,x') + \mathcal{C}_{\partial V}(x,x')\frac{\partial v}{\partial x}(y,x').
$$

If $\mathcal{C}(x) = 0$, then $\Gamma(x,x') = 0$ for the state constraint inequality to hold for all h. When $\mathcal{C}(x) > 0$, the state constraint inequality holds as long as $||h||$ is small enough. Thus, the set of x where the state constraint holds with equality may in principle depend on h.

Consider a point x a point on the boundary of $\{x \in B : \mathcal{C}(x) > 0\}$ when seen as a manifold in $\mathbb{R}^{D_X - 2}$. Under Assumption 3, I can parametrize the boundary as the set of points $X(h) = x + \int \psi(x,x')h(x')\eta(x')$ for a function $\psi(x,x') \in \mathbb{R}^{D_X}$. Recall that x is on the boundary of the domain. Thus, $\psi_j(x,x') = 0$ for some direction j that depends on which boundary x is located on. Without loss of generality, assume that this coordinate is $j = 1$. Evaluate the state constraint at x and expand in h to a first order:

$$\int \left(\sum_{i=2}^{D_X} \mathcal{C}_{x_i}(x)\psi_i(x,x') + \Gamma(x,x') \right) h(x')d\eta(x') \geq 0.$$

Therefore, the boundary of the constrained set changes as per

$$\Gamma(x,x') + \partial_x \mathcal{C}(x) \cdot \psi(x,x') = 0$$

where $\cdot$ denotes here the inner product in $\mathbb{R}^{D_X}$. Because x is on the border of the constrained set $\{y : \mathcal{C}(y) = 0\}$, by continuity $\mathcal{C}_x(x) = 0$ and $\Gamma(x,x') = 0$. Thus, the first-order approximation does not place additional restrictions on ψ nor Γ, and thus v.

B.2.1 Scaled stochastic FAME

When the generator of aggregate shocks $\mathcal{A}$ scales, I guess and verify $\omega(x,z) = \omega_0(x)z$. Then Proposition 3 becomes

$$\overline{\mathcal{M}}(x,z,\omega) \quad = \quad -\mathcal{U}(x)^{-1} \left(\overline{m_0}(x) + \mathcal{L}_c(x)[\omega_0] + \int \overline{m_1}(x,y)\omega_0(y)d\eta(y) \right) \cdot z \equiv \overline{\mathcal{M}}_0(x,\omega_0) \cdot z.$$

Then

$$S(x,z,\omega) = \left\{ \mathcal{L}_c^*\big(y, \overline{\mathcal{M}}_0(y,\omega_0)\big)[g^{SS}]\Big) + \int \mathcal{L}_V^*\big(x', y, \omega_0(y)\big)[g^{SS}]d\eta(y) \right\} \cdot z \equiv S_0(x,\omega_0) \cdot z.$$

Hence, the stochastic FAME becomes:

$$\rho\omega_0(x) = u_z(x) + \mathcal{L}_z(x)[V^{SS}] + \mathcal{L}(x)[\omega_0] + \mathcal{B}[\omega_0] + \int v(x,x')S_0(x',\omega_0)d\eta(x').$$

B.3 Proof of Theorem 4

Condition (i) implies condition (iv) in Theorem 16.0.2. in Meyn and Tweedie (1993b). Condition (ii) implies condition (viii) in Theorem 16.0.2. in Meyn and Tweedie (1993b). Thus, under either condition (i) or condition (ii), condition (ii) in Theorem 16.0.2. in Meyn and Tweedie (1993b) holds, which ensures geometric convergence of the total variation norm of $g_t^{PE}d\eta - g^{SS}d\eta$, where $g_t^{PE} = g^{SS} + h_t^{PE}$. To conclude, note that given the base measure η, the total variation norm is equal to the L^1 norm of the density h_t^{PE}.

To use the steady-state value function as a Lyapunov function, recall that $\mathcal{L}(x)[V^{SS}] = \rho V^{SS}(x) - u^{SS}(x)$. If there is β such that V^{SS} satisfies the Lyapunov inequality, then $\rho V^{SS} - u^{SS} \leq -\beta V^{SS}$, which implies the stated inequality. For the relaxed Lyapunov inequality, see conditition (viii) in Theorem 16.0.2. in Meyn and Tweedie (1993b) with $C = \Delta$, and note that condition (i) holding for $x \in \Delta$ implies that Δ is a petite set in the terminology of Meyn and Tweedie (1993b).

As shown in Meyn and Tweedie (1993a) and Hairer (2021), these results do not depend on the discrete vs. continuous time setting.

B.4 Proof of Theorem 5

Part (i) of Theorem 5 is a direct application of Corollary 3.3 p. 82 in Pazy (1983). To see this, consider the semigroup associated with $\mathcal{L}^*$ defined on $Y = L^1(\eta) \cap \{h \in L^1 : \int h d\eta = 0\}$, which is a Banach space when equipped with $||\cdot||_1$. By Theorem 4, on Y, the semigroup associated with $\mathcal{L}^* + \gamma \text{Id}$ is a contraction semigroup. Thus, we can apply Corollary 3.3 p. 82 in Pazy (1983), and so the semigroup associated with $\mathcal{L}^* + \gamma \text{Id} + \mathcal{G}$ is a contraction semigroup. Hence, the semigroup associated with $\mathcal{L}^* + \mathcal{G}$ has a growth bound $-\gamma$.

Part (ii) of Theorem 5 is a direct application of Theorem 1.1 p. 76 in Pazy (1983). To see this, consider the same semigroup as above. Its growth bound is $(R, -\gamma)$. Thus, the semigroup associated with $\mathcal{L}^* + \mathcal{G}$ has growth bound $-\gamma + R|||\mathcal{G}|||$, where $|||\mathcal{G}|||$ is the operator norm of $\mathcal{G}$. Under the assumptions (ii), as $||u_g + \mathcal{L}_g[V^{SS}]||_1 \to 0$, $||v||_1 \to 0$ and $|||\mathcal{G}||| \to 0$. Hence, when $||u_g + \mathcal{L}_g[V^{SS}]||_1 \to 0$ is small enough, the growth bound $-\gamma + R|||\mathcal{G}|||$ is strictly positive.

In the presence of aggregate shocks, use the solution to the law of motion in Theorem 3:

$$h_t = Q_t h_0 + \int_0^t Q_{t-s} \mathcal{S}(z_s) ds,$$

where Q_t is semigroup associated with $\mathcal{L}^* + \mathcal{G}$. Theorem 5 ensures that $||Q_t h_0||_1 \leq R' e^{-\gamma' t} ||h_0||_1$ for some R', g'. So as long as aggregate shocks remain bounded, and so $||\mathcal{S}(z_s)||_1 \leq \bar{s} < \infty$, I have

$$||h_t||_1 \quad \leq \quad R' e^{-\gamma' t} ||h_0||_1 + R' \int_0^t e^{-\gamma'(t-s)} ||\mathcal{S}(z_s)||_1 ds \leq R' ||h_0||_1 + \frac{R' \bar{s}}{\gamma'} < \infty.$$

B.5 Proof of Theorem 6

B.5.1 Existence

The existence proof leverages Brouwer's fixed point theorem.

I first argue that the space of distributions in the stochastic steady-state is compact. To see this, start from the observation that the state space X is compact. Hence, Prokhorov's theorem ensures that the space of probability measures on X with respect to the measure η, $\mathcal{M}(X)$, is also compact for the Prokhorov metric d that metrizes weak convergence. I now consider the space of probability measures on the space of probability measures, $\mathcal{M}(\mathcal{M}(X) \times Z)$. Since $(\mathcal{M}(X), d)$ is compact, the same argument applies and $(\mathcal{M}(\mathcal{M}(X) \times Z), D)$ is compact, where D is Prokhorov's metric when the underlying space is $\mathcal{M}(X) \times Z$.

I next argue that the law of motion of (z_t, h_t) defines a continuous operator. I start by defining this operator. The SPDE in Theorem 3 defines a SDE for $h_t d\eta$ in $\mathcal{M}(X)$. Denote by $\mathcal{P}_t \in \mathcal{M}(\mathcal{M}(X \times Z))$ the probability measure of $h_t d\eta$. Fix some $\tau > 0$, and consider the operator $T : \mathcal{P}_0 \to \mathcal{P}_\tau$. T defines a linear map from $(\mathcal{M}(\mathcal{M}(X) \times Z), D)$ into itself. Associate to T the map $Q : \mathcal{M}(X) \times Z \to \mathcal{M}(\mathcal{M}(X) \times Z)$ that maps an initial condition (h_0, z_0) to the probability measure $\mathcal{P}_\tau$ over possible (h_τ, z_τ).

I now show continuity of the operator T. By Theorem 19.14 p. 631 in Aliprantis and Border (2006), continuity of T is equivalent to continuity of Q. To establish continuity of Q, consider two starting distributions h_0, h'_0. Denote by $g_0 = g^{SS} + h_0$, $g'_0 = g^{SS} + h'_0$. Consider two corresponding sequences z_t, z'_t of aggregate shocks that start at different points z_0, z'_0. Equation (30) and Theorem 5 imply that

$$||g_\tau - g'_\tau||_1 \quad \leq \quad C_1||g_0 - g'_0||_1 + C_2 \int_0^\tau ||\mathcal{S}(z_s) - \mathcal{S}(z'_s)||_1 ds \tag{36}$$

where the C_i are positive constants. Equation (36) defines a bound on the $g_\tau - g'_\tau$ gap that is stochastic because it depends on the realization of the paths $\{z_s, z'_s\}_{0<s\leq\tau}$. The probability distribution of the path $\{z_s\}_{0<s\leq\tau}$ is continuous in z_0 because of the Feller property. Hence, equation (36) implies that the probability distribution Q of (g_τ, z_τ) is continuous in (g_0, z_0), and so T is continuous.

T being continuous and mapping into a compact convex space, Brouwer's theorem (Corollary 17.56 p. 583 in Aliprantis and Border, 2006) ensures that it has a fixed point. A fixed point of T is an invariant stochastic steady-state measure $\mathcal{P}(dh, dz)$.

B.5.2 Conditional average

Denote by $\Delta(x, h, z) = \mathcal{L}^*(x)[h] + \mathcal{G}(x)[h] + S(x, \omega, z)$ the 'drift' in the 'meta'-KF equation, so that

$$dh_t = \Delta(\cdot, h_t, z_t)dt.$$

Finite-dimensional case. To build intuition, consider first a finite-dimensional case. Here, I assume that the state space is finite and index $x \equiv x_i$. In that case, $\mathcal{L}^*$ and $\mathcal{G}$ are matrices. Then

$$dh_{it} = \Delta(x_i, h_t, z_t)dt.$$

The 'meta'-KF equation for the density $\mathcal{P}(h, z)$ over the vector (h_t, z_t) is then

$$0 = -\sum_i \frac{\partial}{\partial h_i}\Big(\Delta(x_i, h, z)\mathcal{P}(h, z)\Big) + \mathcal{A}^*(z)[\mathcal{P}].$$

Denote by $\hat{h}_k(z) = \int h_k \mathcal{P}(h, z)dh$. Multiply the KF equation by h_k, integrate over h, and integrate by parts the first component. All terms drop out except the one for $i = k$:

$$0 = \int \Delta(x_k, h, z_t)\mathcal{P}(h, z)dh + \mathcal{A}^*(z)[\hat{h}_k].$$

Exploiting that the drift Δ is linear in h, I obtain

$$0 \quad = \quad \int \Big(\mathcal{L}^*_{k\bullet}h + \mathcal{G}_{k\bullet}h + S_k(z)\Big)\mathcal{P}(h, z)dh + \mathcal{A}^*(z)[\hat{h}_k] = \mathcal{L}^*_{k\bullet}\hat{h}(z) + \mathcal{G}_{k\bullet}\hat{h}(z) + S_k(z)\phi(z) + \mathcal{A}^*(z)[h_k]$$

Hence, I have obtained

$$0 \quad = \quad \mathcal{L}^*\hat{h}(z) + \mathcal{G}\hat{h}(z) + S(z)\phi(z) + \mathcal{A}^*(z)[\hat{h}(\cdot)]$$

Infinite-dimensional case. To handle the infinite-dimensional case, I work with the dual form of the 'meta'-KF equation. Denote by $\mathcal{P}(dh, dz)$ an invariant measure associated with the (h_t, z_t)

process. The dual form of the 'meta'-KF equation is (see Bogachev et al., 2015) writes:

$$\int \left\{ \int \Delta(x,h,z) \frac{\partial f(h,z)}{\partial h(x)} d\eta(x) + \mathcal{A}(z)[f(h,\cdot)] \right\} \mathcal{P}(dh,dz) = 0 \quad , \quad \forall h, z,$$

for all smooth functions $f(h,z)$. Fix some x_0 and evaluate this equation at $f(h,z) = \varphi(z)\Psi(h)$, $\Psi(h) = \int h(x)\psi(x)d\eta(x)$ for some smooth functions φ, ψ. I obtain:

$$\int \left\{ \int \Delta(x,h,z)\varphi(z)\psi(x)d\eta(x) + \Psi(h)\mathcal{A}(z)[\varphi] \right\} d\mathcal{P}(h,z) = 0.$$

Although there is no Lebesgue measure in infinite-dimensional settings, if the aggregate shock process has a density, I may write $\mathcal{P}(dh,dz) = \mathcal{Q}(dh,z)dz$. Define $\hat{h}(z) = \int h\mathcal{Q}(dh,z) \in L^2$, and so I will also write $\hat{h}(x,z)$. I then obtain

$$\begin{aligned} 0 &= \int \left\{ \int \Big[(\mathcal{L}^* + \mathcal{G})[\hat{h}(\cdot,z)] + \mathcal{S}(z)\phi(z) \Big] \varphi(z)\psi(x)d\eta(x) + \left(\int \hat{h}(x,z)\psi(x)d\eta(x) \right) \mathcal{A}(z)[\varphi] \right\} dz \\ &= \iint \left\{ \Big[(\mathcal{L}^* + \mathcal{G})[\hat{h}(\cdot,z)] + \mathcal{S}(z)\phi(z) \Big] \varphi(z) + \hat{h}(x,z)\mathcal{A}(z)[\varphi] \right\} \psi(x)d\eta(x)dz \\ &= \iint \left\{ (\mathcal{L}^* + \mathcal{G})[\hat{h}(\cdot,z)] + \mathcal{S}(z)\phi(z) + \mathcal{A}^*(z)[\hat{h}(x,\cdot)] \right\} \varphi(z)\psi(x)d\eta(x)dz \end{aligned}$$

where the last line integrates by parts the term in $\mathcal{A}(z)$ over z. Since this equation must hold for all test functions φ, ψ, I obtain for all x, z:

$$0 = (\mathcal{L}^* + \mathcal{G})[\hat{h}(\cdot,z)] + \mathcal{S}(z)\phi(z) + \mathcal{A}^*(z)[\hat{h}(x,\cdot)].$$

B.5.3 Second moments

I provide a proof only in the finite-dimensional case. The infinite-dimensional case follows the same steps, with notations as in the previous section. Here, I assume that the state space is finite and index $x \equiv x_i$. In that case, $\mathcal{L}^*$ and $\mathcal{G}$ are matrices. Start again from the 'meta'-KF equation for the density:

$$0 = -\sum_i \frac{\partial}{\partial h_i} \Big(\Delta(x_i,h,z_t)\mathcal{P}(h,z) \Big) + \mathcal{A}^*(z)[\mathcal{P}].$$

Denote by $\hat{H}_{k\ell}(z) = \int h_k h_\ell \mathcal{P}(h,z)dh$.

Consider first $k \neq \ell$. Multiply the KF equation by $h_k h_\ell$, integrate over h, and integrate by parts the first component. All terms drop out except the ones for $i = k$ and $i = \ell$:

$$0 = \int \Delta(x_k,h,z_t)h_\ell\mathcal{P}(h,z)dh + \int \Delta(x_\ell,h,z_t)h_k\mathcal{P}(h,z)dh + \mathcal{A}^*(z)[\bar{h}_{k\ell}].$$

Exploiting that the drift Δ is linear in h, I obtain

$$\begin{aligned} 0 &= \int \Big(\mathcal{L}^*_{k\bullet}h + \mathcal{G}_{k\bullet}h + S_k(z) \Big) \mathcal{P}(h,z)h_\ell dh + \int \Big(\mathcal{L}^*_{\ell\bullet}h + \mathcal{G}_{\ell\bullet}h + S_\ell(z) \Big) \mathcal{P}(h,z)h_k dh + \mathcal{A}^*(z)[\bar{h}_{k\ell}] \\ &= (\mathcal{L}^*_{k\bullet} + \mathcal{G}_{k\bullet})\hat{H}_{\bullet,\ell}(z) + (\mathcal{L}^*_{\ell\bullet} + \mathcal{G}_{\ell\bullet})\hat{H}_{\bullet,k}(z) + S_k(z)\hat{h}_\ell(z) + S_\ell(z)\hat{h}_k(z) + \mathcal{A}^*(z)[\hat{H}_{k\ell}] \end{aligned}$$

Using that $\hat{H}$ is symmetric, and denoting $\mathcal{J} = \mathcal{L}^* + \mathcal{G}$, I obtain, for $x \neq x'$

$$0 = \mathcal{J}(x)[\hat{H}(\cdot, x', z)] + \mathcal{J}(x')[\hat{H}(x, \cdot, z)] + \mathcal{A}^*(z)[\hat{H}(x, x', \cdot)] + S(x, z)\hat{h}(x', z) + S(x', z)\hat{h}(x, z),$$

which, once discretized, defines a Sylvester tensor equation.

Now consider $k = \ell$. Multiply the KF equation by h_k^2, integrate over h, and integrate by parts the first component. All terms drop out except the ones for $i = k$:

$$0 = 2\int \Delta(x_k, h, z_t)h_k\mathcal{P}(h, z)dh + \mathcal{A}^*(z)[\bar{h}_{kk}].$$

The same arguments as above imply that, for $x = x'$,

$$0 = 2\mathcal{J}(x)[\hat{H}(\cdot, x', z)] + \mathcal{A}(z)[\hat{H}(x, x', \cdot)] + 2S(x, z)\hat{h}(x, z),$$

which concludes the proof.

B.6 Linear system approach

Consider the setup of Corollary 1. Denote $v_t(x) = \int v(x, x')h_t(x')d\eta(x')$. Then the FAME and the KF equation write:

$$\begin{aligned} \frac{\partial v_t(x)}{\partial t} &= \rho v_t(x) - \int u_g(x, x')h_t(x)d\eta(x) - \mathcal{L}(x)[v_t] \\ \frac{\partial h_t(x)}{\partial t} &= \mathcal{L}^*(x)[h_t] + \mathcal{G}(x, v)[h_t] \equiv \mathcal{L}^*(x)[h_t] + \mathcal{G}_1(x)[h_t] + \mathcal{G}_2(x)[v_t], \end{aligned}$$

where I denoted $\mathcal{G}(x, v) = \mathcal{G}_1(x) + \mathcal{G}_2(x)[v]$. Stacking this linear system, I obtain

$$\frac{\partial}{\partial t}\begin{pmatrix} v_t \\ h_t \end{pmatrix} = \begin{pmatrix} \rho\text{Id} - \mathcal{L}(x) & u_g(x, \cdot) \\ \mathcal{G}_2(x) & \mathcal{L}^*(x) + \mathcal{G}_1(x) \end{pmatrix}\begin{pmatrix} v_t \\ h_t \end{pmatrix}.$$

Following standard linear rational expectations techniques, consider the discretized version of the system. Imposing dynamic stability, the deterministic Impulse Value v must be such that the system is stable, and so $v_t = \int v(x, x')h_t(x')d\eta(x')$ must lie in the vector space spanned by the stable roots of the system. In particular this must hold at time 0 for all possible starting distributions h_0, which determines v.

C Proofs for Section 4

C.1 Proof of Theorem 8

To shorten notation, I denote $h(dx) \equiv h(x)d\eta(x)$ in the sequel. I use the same notation $M_0 = M_1 + M_2 + M_3$ as in Appendix B.2.

I also allow for the state process to be affected by the aggregate shock through other channels than through the drift. Namely, I write $L[V] = b\partial_x V + L_0[V] + \varepsilon z\Lambda[V]$, with Λ, L_0 exogenous.

$M_1 + M_2$. To first order,

$$d[M_1+M_2] = \left(\int b_g(x,x',g,z)h(dx') + b_z(x,g,z)z\right) V_x(x) + z\Lambda[V] + L(x,\hat{c},g,z)[dV]$$

To second order and evaluating at steady-state,

$$\begin{aligned}
d^2[M_1+M_2]^{SS} &= V_x^{SS}(x)\left(\iint b_{gg}(x,x',x'')\ h(dx')h(dx'') + 2\int b_{gz}(x,x')zh(dx') + b_{zz}(x)z^2\right) \\
&+ 2\left(\int b_g(x,x',g,z)h(dx') + b_z(x,g,z)z\right) dV_x(x) - dc(x)dV_x(x) + \mathcal{L}(x)[d^2V] + 2z\Lambda[dV]
\end{aligned}$$

To leading order,

$$\begin{aligned}
dV(x) &= \int v(x,x'')h(dx'')' + \omega(x,z) \\
d^2V(x) &= \iint \mathcal{V}(x,x',x'')h(dx')h(dx'') + 2\int \Gamma(x,x',z)h(dx') + \Omega(x,z)
\end{aligned}$$

Therefore, to second order,

$$\begin{aligned}
d^2[M_1+M_2]^{SS} &= V_x^{SS}(x)\left(\iint b_{gg}(x,x',x'')\ h(dx')h(dx'') + 2\int b_{gz}(x,x')zh(dx') + b_{zz}(x)z^2\right) \\
&+ 2\left(\int b_g(x,x')h(dx') + b_z(x)z\right)\left(\int v_x(x,x'')h(dx'') + \omega_x(x,z)\right) \\
&- \left(\int \mathcal{M}(x,x',v)h(dx') + \overline{\mathcal{M}}(x,z,\omega)\right)\left(\int v_x(x,x'')h(dx'') + \omega_x(x,z)\right) \\
&+ \iint \mathcal{L}(x)[\mathcal{V}(\cdot,x',x'')]h(dx')h(dx'') + 2\int \mathcal{L}(x)[\Gamma(\cdot,x',z)]h(dx') + \mathcal{L}(x)[\Omega(\cdot,z)] \\
&+ 2z\int \Lambda[v(\cdot,x')]h(dx') + 2z\Lambda[\omega(\cdot,z)]
\end{aligned}$$

Re-arranging and changing integration indices to symmetrize the second-order terms in h only, I obtain

$$\begin{aligned}
d^2[M_1+M_2]^{SS} &= V_x^{SS}(x)\left(\iint b_{gg}(x,x',x'')\ h(dx')h(dx'') + 2\int b_{gz}(x,x')z\ h(dx') + b_{zz}(x)z^2\right) \\
&+ \int \Big(b_g(x,x')v_x(x,x'') + b_g(x,x'')v_x(x,x')\Big)\ h(dx')h(dx'') + 2b_z(x)z\omega_x(x,z) \\
&+ 2\int \Big(b_g(x,x')\omega_x(x,z) + b_z(x)zv_x(x,x')\Big)\ h(dx') \\
&- u''(c^{SS}(x))\Bigg\{\iint \mathcal{M}(x,x',v)\mathcal{M}(x,x'',v)\ h(dx')h(dx'') \\
&\qquad + \int \overline{\mathcal{M}}(x,z,\omega)\mathcal{M}(x,x',v)\ h(dx'') + \overline{\mathcal{M}}(x,z,\omega)^2\Bigg\} \\
&+ \iint \mathcal{L}(x)[\mathcal{V}(\cdot,x',x'')]\ h(dx')h(dx'') + 2\int \mathcal{L}(x)[\Gamma(\cdot,x',z)]\ h(dx') + \mathcal{L}(x)[\Omega(\cdot,z)] \\
&+ 2z\int \Lambda[v(\cdot,x')]h(dx') + 2z\Lambda[\omega(\cdot,z)]
\end{aligned}$$

***M*₃.** To first order,

$$M_3 = \int \Big(v(x,x'') + dV_g(x,x'')\Big)\left(L^*(x'')[g] + d\big(L^*(x'')[g]\big)\right) d\eta(x'')$$

where

$$\begin{aligned} d\big(L^*(x'')[g]\big) &= -\partial_{x''}\left(\left(\int b_g(x'',x')h(dx') + zb_z(x'') - dc(x'')\right) g(x'')\right) + L^*(x'')[h] \\ &+ \ z\Lambda^*(x'')[g] \end{aligned}$$

To second order and evaluating at steady-state,

$$\begin{aligned} d^2\big[L^*(x'')[g]\big]^{SS} &= -2\partial_{x''}\left(\left(\int b_g(x'',x')h(dx') + zb_z(x'') - dc(x)\right) h(x'')\right) \\ &- \ \partial_{x''}\left(\left(\iint b_{gg}(x'',x',y)h(dx')h(dy) + 2\int zb_{gz}(x'',x')h(dx') + z^2 b_{zz}(x'') - d^2c(x)\right) g^{SS}(x'')\right) \\ &+ \ 2z\Lambda^*(x'')[h] \end{aligned}$$

Using that $u'''(c^{SS}(x))(dc(x))^2 + u''(c^{SS}(x))d^2c(x) = d^2V_x(x)$, I obtain

$$\begin{aligned} d^2\big[L^*(x'')[g]\big]^{SS} &= -2\partial_{x''}\left(\left(\int \big(b_g(x'',x') - \mathcal{M}(x'',x',v)\big)h(dx') + zb_z(x'') - \overline{\mathcal{M}}(x'',z,\omega)\right) h(x'')\right) \\ &- \ \partial_{x''}\left(\left(\iint b_{gg}(x'',x',y)h(dx')h(dy) + 2\int zb_{gz}(x'',x')h(dx') + z^2 b_{zz}(x'')\right) g^{SS}(x'')\right) \\ &+ \ \partial_{x''}\Bigg(\frac{g^{SS}(x'')}{u''(c^{SS}(x''))}\Bigg(\iint \mathcal{V}_{x''}(x'',x',y)\ h(dx')h(dy) + 2\int \Gamma_{x''}(x'',x',z)\ h(dx') + \Omega_{x''}(x'',z) \\ &\qquad - \frac{u'''(c^{SS}(x''))}{\big(u''(c^{SS}(x''))\big)^2}\Bigg[\iint v_{x''}(x'',x')v_{x''}(x'',y)h(dx')h(dy) \\ &\qquad + 2\int v_{x''}(x'',x')\omega_x(x'',z)h(dx') + \omega_x(x'',z)^2\Bigg]\Bigg)\Bigg) \\ &+ \ 2z\Lambda^*(x'')[h] \end{aligned} \tag{37}$$

Finally, to leading order,

$$dV_g(x,x'') = \int \Big(\mathcal{V}(x,x',x'') + \mathcal{V}(x,x'',x')\Big)h(dx') + 2\Gamma(x,x'',z)$$

Now, to a leading order,

$$d[M_3] = \underbrace{\int v(x,x'') \times d^2\big(L^*(x'')[g]\big)\ d\eta(x'')}_{\equiv M_{31}} + \underbrace{\int dV_g(x,x'') \times d\big(L^*(x'')[g]\big)\ d\eta(x'')}_{\equiv M_{32}}.$$

Starting with the first component,

$$
\begin{aligned}
M_{31} &= 2\iint v_{x''}(x,x'')\big(b_g(x'',x') - \mathcal{M}(x'',x',v)\big)\ h(dx')h(dx'') + 2\int v_{x'}(x,x')\big(zb_z(x'') - \overline{\mathcal{M}}(x',z,\omega)\big)\ h(dx') \\
&+ \iint \left(\int v_y(x,y) b_{gg}(y,x',x'') g^{SS}(dy)\right) h(dx')h(dx'') + \int \left(2\int v_y(x,y) z b_{gz}(y,x') g^{SS}(dy)\right) h(dx') \\
&+ \int v_y(x,y) z^2 b_{zz}(y) g^{SS}(dy) \\
&- \iint \left(\int v_y(x,y) g^{SS}(dy) k(y) \left[\mathcal{V}_y(y,x',x'') + k_p(y) v_y(y,x') v_y(y,x'')\right]\right)\ h(dx')h(dx'') \\
&- 2\int \left(\int v_y(x,y) g^{SS}(dy) k(y) \left[\Gamma_y(y,x',z) + k_p(y) v_y(y,x') \omega_y(y,z)\right]\right)\ h(dx') \\
&- \int v_y(x,y) g^{SS}(dy) k(y) \left[\Omega_y(y,z) + k_p(y) \omega_y(y,z)^2\right] \\
&+ 2z \int \Lambda[v(x,\cdot)] h(dx'')
\end{aligned}
$$

where I denoted

$$
k(y) = \frac{1}{u''(c^{SS}(y))} \quad , \quad k_p(y) = -\frac{u'''(c^{SS}(y))}{\big(u''(c^{SS}(y))\big)^2}.
$$

The second component is, to leading order,

$$
\begin{aligned}
M_{32} &= \int \left(\int \mathcal{V}(x,x',x'') h(dx') + 2\Gamma(x,x'',z)\right) \\
&\qquad \times \Bigg\{ -\partial_{x''} \left(\left(\int \big(b_g(x'',x') - \mathcal{M}(x'',x',v)\big) h(dx') + zb_z(x'') - \overline{\mathcal{M}}(x'',z,\omega) \right) g^{SS}(x'') \right) \\
&\qquad + \mathcal{L}^*(x'')[h] \Bigg\} d\eta(x'') \\
&= \iint \left(\mathcal{L}(x'')[\mathcal{V}(x,x',\cdot)] + \mathcal{L}(x')[\mathcal{V}(x,\cdot,x'')] + \int \mathcal{V}_y(x,y,x'')\big(b_g(y,x') - \mathcal{M}(y,x',v)\big) g^{SS}(dy) \right) h(dx')h(dx'') \\
&+ \int \Bigg\{ 2\mathcal{L}(x')[\Gamma(x,\cdot,z)] + 2\int \Bigg[\Gamma_y(x,y,z)\big(b_g(y,x') - \mathcal{M}(y,x',v)\big) \\
&\qquad + \big(zb_z(y) - \overline{\mathcal{M}}(y,z,\omega)\big) \left(\mathcal{V}_y(x,x',y) + \mathcal{V}_y(x,y,x')\right) \Bigg] g^{SS}(dy) \Bigg\} h(dx') \\
&+ 2\int \Gamma_y(x,y,z) \left(zb_z(y) - \overline{\mathcal{M}}(y,z,\omega)\right) g^{SS}(dy)
\end{aligned}
$$

Putting these equations together and identifying coefficients,[39] I obtain the SAMEs.

[39] In the case of the second-order expansion, 'identifying coefficients' corresponds to the results stating that if a quadratic form defined by a symmetric operator is equal to another quadratic form defined by a symmetric operator, then both operators must be equal. When either one of the operators is not symmetric, then only their symmetric parts are equal.

Deterministic SAME.

$$\begin{aligned}
\rho\mathcal{V}(x,x',x'') &= u'(c^{SS}(x)b_{gg}(x,x',x'') + b_g(x,x')v_x(x,x'') + b_g(x,x'')v_x(x,x') \\
&- u''(c^{SS}(x))\mathcal{M}(x,x',v)\mathcal{M}(x,x'',v) + \mathcal{L}(x)[\mathcal{V}(\cdot,x',x'')] \\
&+ v_{x''}(x,x'')\big(b_g(x'',x') - \mathcal{M}(x'',x',v)\big) + v_{x'}(x,x')\big(b_g(x',x'') - \mathcal{M}(x',x'',v)\big) \\
&+ \int v_y(x,y)\Big\{b_{gg}(y,x',x'') - k(y)\Big[\mathcal{V}_y(y,x',x'') + k_p(y)v_y(y,x')v_y(y,x'')\Big]\Big\}g^{SS}(dy) \\
&+ \mathcal{L}(x'')[\mathcal{V}(x,x',\cdot)] + \mathcal{L}(x')[\mathcal{V}(x,\cdot,x'')] \\
&+ \int\Big[\mathcal{V}_y(x,y,x'')\big(b_g(y,x') - \mathcal{M}(y,x',v)\big) + \mathcal{V}_y(x,x',y)\big(b_g(y,x'') - \mathcal{M}(y,x'',v)\big)\Big]g^{SS}(dy)
\end{aligned}$$

I define

$$\begin{aligned}
T_{\mathcal{V}}(x,x',x'') &= \underbrace{b_{gg}(x,x',x'')u'(c^{SS}(x))}_{\text{Direct price}} + \underbrace{b_g(x,x')v_x(x,x'') + b_g(x,x'')v_x(x,x')}_{\text{Cross price-continuation value}} + \underbrace{u''(c^{SS}(x))\mathcal{M}(x,x',v)\mathcal{M}(x,x'',v)}_{\text{Cross consumption-continuation value}} \\
&+ \underbrace{\Big[v_{x'}(x,x')\big(b_g(x',x'') - \mathcal{M}(x',x'',v)\big) + v_{x''}(x,x'')\big(b_g(x'',x') - \mathcal{M}(x'',x',v)\big)\Big]}_{\text{GE: cross: others' savings-impulse} \equiv \text{change in propagation of impulse due to change in savings}} \\
&+ \underbrace{\int v_y(x,y)g^{SS}(y)\Big[b_{gg}(y,x',x'') - k(y)k_p(y)v_y(y,x')v_y(y,x'')\Big]dy}_{\text{GE: 1}^{\text{st}}\text{-order valuation of 2}^{\text{nd}}\text{-order changes in others' savings}} \qquad (38)
\end{aligned}$$

and

$$\sigma(y,x) = -\partial_y\Big(\big(b_g(y,x) - \mathcal{M}(y,x,v)\big)g^{SS}(y)\Big) \quad , \quad \tau(x,y) = \partial_y\Big(v_y(x,y)k(y)g^{SS}(y)\Big)$$

so that the deterministic SAME re-writes

$$\begin{aligned}
\rho\mathcal{V}(x,x',x'') &= T(x,x',x'') + \mathcal{L}(x)[\mathcal{V}(\cdot,x',x'')] + \mathcal{L}(x')[\mathcal{V}(x,\cdot,x'')] + \mathcal{L}(x'')[\mathcal{V}(x,x',\cdot)] \\
&+ \int\Big(\mathcal{V}(x,y,x'')\sigma(y,x') + \mathcal{V}(x,x',y)\sigma(y,x'')\Big)dy + \int\mathcal{V}(y,x',x'')\tau(x,y)dy.
\end{aligned}$$

Cross SAME. The cross SAME writes

$$\begin{aligned}
\rho\Gamma(x,x',z) &= u'(c^{SS}(x))b_{gz}(x,x')z + b_g(x,x')\omega_x(x,z) + b_z(x)zv_x(x,x') - u''(c^{SS}(x))\overline{\mathcal{M}}(x,z,\omega)\mathcal{M}(x,x',v) \\
&+ v_{x'}(x,x')\big(zb_z(x') - \overline{\mathcal{M}}(x',z,\omega)\big) + \int v_y(x,y)zb_{gz}(y,x')g^{SS}(dy) \\
&+ z\Lambda[v(\cdot,x')] \\
&+ \mathcal{L}(x)[\Gamma(\cdot,x',z)] + \mathcal{L}(x')[\Gamma(x,\cdot,z)] + \mathcal{A}(z)[\Gamma(x,x',\cdot)] \\
&- \int v_y(x,y)g^{SS}(dy)k(y)\Big[\Gamma_y(y,x',z) + k_p(y)v_y(y,x')\omega_y(y,z)\Big] \\
&+ \int\Big[\Gamma(x,y,z)\sigma(y,x') + \mathcal{V}(x,x',y)\overline{\sigma}(y,z)\Big]dy.
\end{aligned}$$

where

$$\overline{\sigma}(y,z) = -\partial_y\Big(g^{SS}(y)\left(zb_z(y) - \overline{\mathcal{M}}(y,z,\omega)\right)\Big)$$

Thus, define

$$\begin{aligned}
T_\Gamma(x,x',z) &= u'(c^{SS}(x))b_{gz}(x,x')z + b_g(x,x')\omega_x(x,z) + b_z(x)zv_x(x,x') - u''(c^{SS}(x))\overline{\mathcal{M}}(x,z,\omega)\mathcal{M}(x,x',v) \\
&+ v_{x'}(x,x')\big(zb_z(x') - \overline{\mathcal{M}}(x',z,\omega)\big) + \int g^{SS}(y)v_y(x,y)\big[zb_{gz}(y,x') - k(y)k_p(y)v_y(y,x')\omega_y(y,z)\big] \\
&+ \int \mathcal{V}(x,x',y)\overline{\sigma}(y,z)dy \\
&+ z\Lambda[v(\cdot,x')]. \qquad (39)
\end{aligned}$$

The cross SAME then becomes

$$\begin{aligned}
\rho\Gamma(x,x',z) &= T_\Gamma(x,x',z) + \mathcal{L}(x)[\Gamma(\cdot,x',z)] + \mathcal{L}(x')[\Gamma(x,\cdot,z)] + \mathcal{A}(z)[\Gamma(x,x',\cdot)] \\
&+ \int \Big(\tau(x,y)\Gamma(y,x',z) + \Gamma(x,y,z)\sigma(y,x')\Big)dy.
\end{aligned}$$

Stochastic SAME. The stochastic SAME writes

$$\begin{aligned}
\rho\Omega(x,z) &= u'(c^{SS}(x))b_{zz}(x)z^2 + 2b_z(x)z\omega_x(x,z) - u''(c^{SS}(x))\overline{\mathcal{M}}(x,z,\omega)^2 + \int \Gamma(x,y,z)\overline{\sigma}(y,z)dy + 2z\Lambda[\omega(\cdot,z)] \\
&+ 2z\Lambda[\omega(\cdot,z)] \\
&+ \mathcal{L}(x)[\Omega(\cdot,z)] + \mathcal{A}(z)[\Omega(x,\cdot)] \\
&+ 2\int v_y(x,y)z^2 b_{zz}(y)g^{SS}(dy) \\
&- 2\int v_y(x,y)g^{SS}(dy)k(y)\left[\Omega_y(y,z) + k_p(y)\omega_y(y,z)^2\right] \\
&+ 2\int \Gamma_y(x,y,z)\left(zb_z(y) - \overline{\mathcal{M}}(y,z,\omega)\right)g^{SS}(dy).
\end{aligned}$$

Define

$$\begin{aligned}
T_S(x,z) &= u'(c^{SS}(x))b_{zz}(x)z^2 + 2b_z(x)z\omega_x(x,z) - u''(c^{SS}(x))\overline{\mathcal{M}}(x,z,\omega)^2 \\
&+ 2\int \Gamma(x,y,z)\overline{\sigma}(y,z) + \int v_y(x,y)g^{SS}(y)\big(b_{zz}(y,z) - k(y)k_p(y)\omega_y(y,z)^2\big)dy \\
&+ 2z\Lambda[\omega(\cdot,z)]. \qquad (40)
\end{aligned}$$

The stochastic SAME becomes

$$\rho\Omega(x,z) = T_S(x,z) + \mathcal{L}(x)[\Omega(\cdot,z)] + \mathcal{A}(z)[\Omega(x,\cdot)] + 2\int \tau(x,y)\Omega(y,z)dy.$$

C.1.1 Scaled cross SAME

When $\mathcal{A}[\varphi] = z\cdot\mathcal{B}[\varphi]$ for φ linear, it is straightforward to guess and verify that Γ_0 solves:

$$\rho\Gamma_0(x,x') = T_{\Gamma 0}(x,x') + \mathcal{L}_\tau(x)[\Gamma_0(\cdot,x')] + \mathcal{L}_\sigma(x')[\Gamma_0(x,\cdot)] + \mathcal{B}[\Gamma_0(x,x')],$$

where

$$
\begin{aligned}
T_{\Gamma 0}(x,x') &= u'(c^{SS}(x))b_{gz}(x,x') + b_g(x,x')\omega_{0,x}(x) + b_z(x)v_x(x,x') - u''(c^{SS}(x))\overline{\mathcal{M}}_0(x,\omega_0)\mathcal{M}(x,x',v) \\
&+ v_{x'}(x,x')\big(b_z(x') - \overline{\mathcal{M}}_0(x',\omega_0)\big) + \int g^{SS}(y)v_y(x,y)\big[b_{gz}(y,x') - k(y)k_p(y)v_y(y,x')\omega_{0,y}(y)\big] \\
&+ \int \mathcal{V}(x,x',y)\overline{\sigma}_0(y)dy,
\end{aligned}
$$

and

$$
\overline{\sigma}_0(y) = -\partial_y\Big(g^{SS}(y)\left(b_z(y) - \overline{\mathcal{M}}_0(y,\omega_0)\right)\Big).
$$

C.2 Proof of Theorem 9

I may read off the perturbation of the law of motion of the distribution from the derivation of the SAME, in particular equation (37). I denote:

$$
\begin{aligned}
\sigma^*(x,x') &= \sigma(x',x) \\
\theta(x,x',x'') &= -\partial_x\left(g^{SS}(x)\Big[b_{gg}(x,x',x'') - k_p(x)v_x(x,x')v_x(x,x'') - k(x)\mathcal{V}_x(x,x',x'')\Big]\right) \\
\overline{\theta}(x,x',z) &= -\partial_x\left(g^{SS}(x)\Big[zb_{zg}(x,x') - k(x)\Gamma_x(x,x',z_t) - k(x)k_p(x)v_x(x,x')\omega_x(x,z_t)\Big]\right) \\
\Theta(x,z) &= -\partial_x\left(g^{SS}(x)\Big(z_t^2 b_{zz}(x) - k(x)\Omega_x(x,z_t) - k(x)k_p(x)\omega_x(x,z_t)^2\Big)\right)
\end{aligned}
$$

Then:

$$
\begin{aligned}
\frac{dh_t(x)}{dt} &= \mathcal{L}^*(x)[h_t] + \mathcal{G}(x)[h_t] + \mathcal{S}(z_t) \\
&+ \varepsilon \times \Bigg\{2\partial_x\left(h_t(x)\sigma^*(x,\cdot)[h_t]\right) + \iint \theta(x,x',x'')h_t(dx')h_t(dx'') \\
&\qquad + \partial_x\Big(\overline{\sigma}(x,z_t)h_t(x)\Big) + 2\int \overline{\theta}(x,x',z_t)h_t(dx') + \Theta(x,z_t) \\
&\qquad + 2z\Lambda^*(x)[h]\Bigg\}
\end{aligned}
$$

I define the quadratic form in (h,z):

$$
\begin{aligned}
\mathcal{Q}(h,z) &= 2\partial_x\left(h(x)\sigma^*(x,\cdot)[h]\right) + \iint \theta(x,x',x'')h(dx')h(dx'') \\
&+ \partial_x\Big(\overline{\sigma}(x,z_t)h(x)\Big) + 2\int \overline{\theta}(x,x',z)h(dx') + \Theta(x,z) \\
&+ 2z\Lambda^*(x)[h]. \qquad (41)
\end{aligned}
$$

Then, I obtain

$$
\frac{dh_t(x)}{dt} = \mathcal{L}^*(x)[h_t] + \mathcal{G}(x)[h_t] + \mathcal{S}(z_t) + \frac{\varepsilon}{2}\mathcal{Q}(h_t,z_t).
$$

C.3 Proof of Theorem 10

C.3.1 Characterizing $\widetilde{h}$

The first step is to use Theorem 9 to characterize $\widetilde{h}$. Using Dirac measures, one can write

$$\mathcal{Q}(x,h,z) = \iint \mathcal{Q}^{hh}(x,x',x'')h(dx')h(dx'') + \int \mathcal{Q}^{hz}(x,x',z)h(dx') + \mathcal{Q}^{zz}(x,z).$$

I follow the same steps as in Appendix B.5 and use a finite-dimensional notation once more for brevity. Denote by $\widehat{h}_{k,\varepsilon} = \int h_k \mathcal{P}(dh,dz)$ in the (second-order) stochastic steady-state. Similarly, denote $\widehat{H}_{k\ell,\varepsilon} = \int h_k h_\ell \mathcal{P}(dh,dz)$. I obtain

$$\begin{aligned} 0 &= \mathcal{J}_{k\bullet}\widehat{h}_\varepsilon + \mathcal{S}_k\phi + \varepsilon\left\{\sum_{j,\ell}\mathcal{Q}^{hh}_{kj\ell}\widehat{H}_{j\ell,\varepsilon} + \widehat{h}_\varepsilon^T\mathcal{Q}^{hz}_k + \mathcal{Q}^{zz}_k\right\} + \mathcal{A}^*(z)[\widehat{h}_{k,\varepsilon}] \\ &= \mathcal{J}_{k\bullet}\widehat{h}_\varepsilon + \mathcal{S}_k\phi + \varepsilon\left\{\sum_{j,\ell}\mathcal{Q}^{hh}_{kj\ell}\widehat{H}_{j\ell} + \widehat{h}^T\mathcal{Q}^{hz}_k + \mathcal{Q}^{zz}_k\right\} + \mathcal{A}^*(z)[\widehat{h}_{k,\varepsilon}] \end{aligned}$$

where the second equality replaces $\widehat{h}_\varepsilon, \widehat{H}_\varepsilon$ by $\widehat{h}, \widehat{H}$ to leading order. Hence, $\widehat{h}_\varepsilon$ satisfies the linear equation

$$0 = \mathcal{J}(x)[\widehat{h}_\varepsilon(\cdot,z)] + \mathcal{A}^*(z)[\widehat{h}_\varepsilon(x,\cdot)] + \mathcal{S}(x,z)\phi + \varepsilon\check{\mathcal{S}}(x,z),$$

where

$$\check{\mathcal{S}}(x,z) = \iint \mathcal{Q}^{hh}(x,x',x'')\widehat{H}(x',x'',z)dx'dx'' + \int \mathcal{Q}^{hz}(x,x',z)\widehat{h}(x',z)dx' + \mathcal{Q}^{zz}(x,z)\phi(z)\Big\}. \tag{42}$$

Hence, denoting $\widehat{h}_\varepsilon = \widehat{h} + \varepsilon\check{h}_\varepsilon$, I obtain

$$0 = \mathcal{J}(x)[\check{h}_\varepsilon(\cdot,z)] + \mathcal{A}^*(z)[\check{h}_\varepsilon(x,\cdot)] + \check{\mathcal{S}}(x,z).$$

C.3.2 Welfare cost of aggregate risk

With $\widetilde{h}(x,z) = \check{h}(x,z)/\phi(z)$ in hand, the derivation is a direct second-order approximation and is omitted for brevity.

D Details and proofs for examples and applications

D.1 Derivations for Section 1.6: Master Equation

D.1.1 Static equilibrium

The wage w_{it} in location i satisfies the firm's first-order condition $w_{it} = (1-\alpha)Z_{it}N_{it}^{-\alpha}$. Local demand for housing clears the housing market $\beta w_{it}N_{it} = r_{it}H_{it}$. The rental rate clears the housing market:

$\beta(1-\alpha)Z_{it}N_{it}^{1-\alpha} = H_i r_{it}^{1+\eta}$, and so $r_{it} = \left(\frac{\beta(1-\alpha)Z_{it}}{H_i}N_{it}^{1-\alpha}\right)^{\frac{1}{1+\eta}}$. Summarizing,

$$w_{it} = (1-\alpha)Z_{it}N_{it}^{-\alpha} = \underbrace{(1-\alpha)Z_i}_{\equiv w_{0i}} e^{\varepsilon\chi_i z_t} N_{it}^{-\alpha}$$

$$r_{it} = \left(\frac{\beta(1-\alpha)Z_{it}}{H_i}N_{it}^{1-\alpha}\right)^{\frac{1}{1+\eta}} = \underbrace{\left(\frac{\beta(1-\alpha)Z_i}{H_i}\right)^{\frac{1}{1+\eta}}}_{\equiv r_{0i}} e^{\frac{1}{1+\eta}\chi_i z_t} N_{it}^{\frac{1-\alpha}{1+\eta}}. \tag{43}$$

The consumption index (the real wage) in location i is

$$C_{it} = \frac{w_{it}}{r_{it}^{\beta}} = \frac{(1-\alpha)Z_{it}N_{it}^{-\alpha}}{\left(\frac{\beta(1-\alpha)Z_{it}}{H_i}N_{it}^{1-\alpha}\right)^{\frac{\beta}{1+\eta}}} = \underbrace{\beta^{-\frac{\beta}{1+\eta}}[(1-\alpha)Z_i]^{\zeta}H_i^{\frac{\beta}{1+\eta}}}_{\equiv C_{0i}} e^{\zeta\chi_i z_t} N_{it}^{-\xi} \tag{44}$$

where $\zeta = \frac{1+\eta-\beta}{1+\eta}$ and $\xi = \alpha + \frac{\beta(1-\alpha)}{1+\eta}$. Flow utility in location i at time t:

$$u\left(\frac{w_{it}}{r_{it}^{\beta}}\right) = u\left(C_{0i}e^{\zeta\chi_i z_t}N_{it}^{-\xi}\right) \equiv U_i(z_t, N_{it}) \tag{45}$$

D.1.2 Migration

Individuals solve

$$\rho V_{it} - \frac{\mathbb{E}[\partial_t V_{it}]}{dt} - U_i(z_t, N_{it}) = \mu\left\{\frac{1}{\nu}\log\left(\sum_j e^{\nu(V_{jt}-\tau_{ij})}\right) - V_{it}\right\} \equiv \mathcal{M}_i[V] \tag{46}$$

$\mathcal{M}_i[V]$ is the nonlinear continuation value operator. Crucially, while this operator is nonlinear, its action on a small perturbation of the value dV still coincides with the adjoint (tranpose) of the operator M^*: $\mathcal{M}[V^{SS}+dV] = \mathcal{M}[V^{SS}] + M(V^{SS})\cdot dV$. Location decisions are given by the conditional choice probabilities (migration shares):

$$m_{ijt}(V) = \frac{e^{\nu(V_j-\tau_{ij})}}{\sum_k e^{\nu(V_k-\tau_{ik})}}. \tag{47}$$

The population distribution evolves according to the KF equation

$$\frac{\partial N_{it}}{\partial t} = \mu\left(\sum_k m_{ki}(V_t)N_{kt} - N_{it}\right) \equiv \mu\Big(\big(m^*(V_t) - \mathrm{Id}\big)N_t\Big)_i. \tag{48}$$

In equation (48), m^* denotes the matrix transpose of the matrix m. m denotes the matrix of migration shares $m_{ij}(V_t)$. Denote by $M^*(V) = \mu\big(m^*(V) - \mathrm{Id}\big)$ the matrix that collects migration shares and the identity matrix, so that (48) becomes, in matrix notation, $\frac{\partial N_t}{\partial t} = M^*(V)N_t$. The Master Equation (19) follows.

D.2 Derivations for Section 1.6: FAME

To solve the FAMEs, I expand around the initial steady-state. I look for a solution

$$V_i \approx V_i^{SS} + \varepsilon \left\{ \sum_j v_{ij} n_j + \omega_i z \right\}$$

where ε is small, and I denote $N = N^{SS} + \epsilon n$.

D.2.1 Flow payoffs

The flow payoffs become, to first order,

$$\frac{U_{it} - U_i^{SS}}{\varepsilon} = u'(C_i^{SS}) C_i^{SS} \cdot \left(\zeta \chi_i z_t - \xi \frac{n_i}{N_i^{SS}} \right) \implies \varepsilon^{-1}(U_t - U^{SS}) = z_t \varpi - \bar{v} n_t.$$

D.2.2 Continuation value from migration

The continuation value from migration becomes to first order $\frac{\mathcal{M}[V] - \mathcal{M}[V^{SS}]}{\varepsilon} = M v^N n + z M \omega$. where henceforth M denotes the steady-state matrix $L(V^{SS})$.

D.2.3 Continuation value from changes in population distribution

To linearize the law of motion of population, first note that

$$\frac{\partial m_{ji}}{\partial V_k} = \begin{cases} -\nu m_{ji} m_{jk} & \text{if} \quad i \neq k \\ \nu m_{ji}(1 - m_{ji}) & \text{if} \quad i = k \end{cases}$$

and so

$$M_{ij}^*(V + dV) \quad = \quad M_{ij}^*(V) + \mu \sum_k \frac{\partial m_{ji}}{\partial V_k} dV_k = M_{ij}^*(V) + \mu \nu m_{ji} \Big\{ dV_i - \big[m \cdot dV \big]_j \Big\}$$

Then

$$\begin{aligned} M_i^*(V + dV)[N^{SS}] \quad &= \quad M_i^*(V)[N^{SS}] + \mu \nu \sum_j m_{ji} \Big\{ dV_i - \big[m \cdot dV \big]_j \Big\} N_j^{SS} \\ &= \quad M_i^*(V)[N^{SS}] + \nu \mu \Big[\Big(\mathbf{diag}(m^* N^{SS}) - m^* \mathbf{diag}(N^{SS}) m \Big) \cdot dV \Big]_i \\ &\equiv \quad M_i^*(V)[N^{SS}] + (G \cdot dV)_i \end{aligned}$$

where

$$G \quad = \quad \nu \mu \Big(\mathbf{diag}(m^* N^{SS}) - m^* \mathbf{diag}(N^{SS}) m \Big). \tag{49}$$

In vector notation,

$$\varepsilon^{-1} \left(\sum_j \frac{\partial V_i}{\partial N_j} (M^*(V) N)_j \right)_{i=1....I} \quad = \quad v^N M^* n + v G \big(v n + z \omega \big) \tag{50}$$

D.2.4 FAME equation

Combining the previous elements, the linearized master equation for the worker value becomes, in vector notation,

$$\rho\big(vn + z\omega\big) = \varpi z - \bar{v}n + M\big(vn + z\omega\big) + \mathcal{A}(z)[z\omega] + vM^{*}n + vG\big(vn + z\omega\big)$$

Identifying coefficients on n and z, and using expression for the generator of the AR(1) process, I obtain the deterministic and stochastic FAMEs:

$$\begin{aligned} \rho v &= -\bar{v} + Mv + vM^{*} + vGv \\ (\rho+\theta)\omega &= \varpi + (M + vG)\omega \end{aligned} \tag{51}$$

D.2.5 Linearized KFEs

I obtain, in vector notation

$$\frac{dn_t}{dt} = M^{*}n_t + G\big(vn_t + z_t\omega\big) \tag{52}$$

D.3 Derivations for Section 5.2: SAME

For the second order approximation, I look for a solution $V = V^{SS} + \varepsilon v_1 + \frac{\varepsilon^2}{2}v_2$, where $N = N^{SS} + \varepsilon n$. The first-order term for V, v_1, is given as before by $v_{1i} = \sum_j v_{ij}n_j + z\omega$. The second-order term, v_2, is given by:

$$v_{2i} = \textstyle\sum_j \sum_s \mathcal{V}_{ijs}n_j n_s + 2\sum_j \Gamma_{ij} z n_j + \Delta_i z^2 + \Lambda_i,$$

where $\mathcal{V}, \Gamma, \Delta, \Lambda$ are the unknowns of interest. To ease calculations, it will be convenient to define $\Omega(z) = \frac{1}{z^2}\left\{\Delta z^2 + \Lambda\right\}$ so that $v_{2i} = \sum_j \sum_k \mathcal{V}_{ijk}n_j n_k + 2\sum_j \Gamma_{ij} z n_j + \Omega_i(z)z^2$. Written in matrix form:

$$v_1 = vn + z\omega \quad , \quad v_2 = \mathcal{V} \times_3 n^{*} \times_2 n^{*} + 2z\Gamma n + z^2\Omega,$$

where $\times_3$ and $\times_2$ are the 3-node and 2-node tensor products, respectively, and the super index * denotes the transpose operator. Tensor products simply generalize matrix products to higher-dimensional arrays. I recall basic properties of tensor operations in Appendix F.

D.3.1 Flow payoffs

The second-order effects are

$$\begin{aligned} \left.\frac{\partial^2 U_{it}}{\partial N_{it}^2}\right|_{SS} &= \frac{\xi C_i^{SS}\left[\xi u''\left(C_i^{SS}\right)C_i^{SS} + (1+\xi)u'\left(C_i^{SS}\right)\right]}{(N_i^{SS})^2} \\ \left.\frac{1}{\varepsilon}\frac{\partial^2 U_{it}}{\partial z \partial N_{it}}\right|_{SS} &= -\frac{\xi\zeta\chi_i C_i^{SS}\left[u''\left(C_i^{SS}\right)C_i^{SS} + u'\left(C_i^{SS}\right)\right]}{N_i^{SS}} \\ \left.\frac{1}{\varepsilon^2}\frac{\partial^2 U_{it}}{\partial z^2}\right|_{SS} &= \Big\{u'\left(C_i^{SS}\right)C_i^{SS} + u''\left(C_i^{SS}\right)\left(C_i^{SS}\right)^2\Big\}(\zeta\chi_i)^2. \end{aligned}$$

Therefore, I obtain the second-order term of the flow payoff:

$$\boxed{\frac{1}{\varepsilon^2}\left(U_{it}-U_i^{SS}\right)\Big|_2 = \frac{1}{2}\mathbb{D}_{UNN}\times_3 n^*\times_2 n^* - \mathrm{D}_{UNZ}n + \frac{1}{2}z^2\mathrm{D}_{UZZ},}$$

where $\mathbb{D}_{UNN}$ is a diagonal tensor defined as:

$$\mathbb{D}_{UNN} = \xi\mathbf{diag}_3\left(\frac{C_i^{SS}\left[\xi u''\left(C_i^{SS}\right)C_i^{SS}+(1+\xi)u'\left(C_i^{SS}\right)\right]}{(N_i^{SS})^2}\right),$$

and where $\mathbf{diag}_3$ is the 3-dimensional diagonal operator. Moreover, the matrix D_{UNZ} and the vector D_{UZZ} are defined as:

$$\mathrm{D}_{UNZ} = \mathbf{diag}\left(\frac{\xi\zeta\chi_i C_i^{SS}\left[u''\left(C_i^{SS}\right)C_i^{SS}+u'\left(C_i^{SS}\right)\right]}{N_i^{SS}}\right)$$

$$\mathrm{D}_{UZZ} = \mathbf{vec}\left(\left\{u'\left(C_i^{SS}\right)C_i^{SS}+u''\left(C_i^{SS}\right)\left(C_i^{SS}\right)^2\right\}(\zeta\chi_i)^2\right)$$

D.3.2 Continuation value from migration

The second-order expansion of the continuation value from migration is:

$$L_i(V_t)-L_i(V^{SS}) = \nabla[L_i](V^{SS})(V-V^{SS})+\frac{1}{2}(V-V^{SS})^T H\left[L_i\right](V^{SS})(V-V^{SS}),$$

where $\nabla[L_i](\cdot)$ and $H\left[L_i\right](\cdot)$ are the gradient and the Hessian of L_i, respectively. Up to a second order, $V-V^{SS}=\epsilon v_1+\frac{\epsilon^2}{2}v_2$. Therefore, omitting the dependence on V^{SS}, the expression writes:

$$\begin{aligned} L_i(V_t)-L_i(V^{SS}) &= \varepsilon\left[Mv_1\right]_i+\frac{\varepsilon^2}{2}\left[Mv_2\right]_i+\frac{1}{2}(\varepsilon v_1+\frac{\varepsilon^2}{2}v_2)^T H\left[\mathcal{M}_i\right](V^{SS})(\epsilon v_1+\frac{\epsilon^2}{2}v_2) \\ &= \varepsilon\left[Mv_1\right]_i+\frac{\varepsilon^2}{2}\left[Mv_2\right]_i+\frac{\varepsilon^2}{2}(v_1^*Hv_1)_i \end{aligned} \quad (53)$$

Focus on the second-order term $Mv_2+\frac{1}{2}v_1^*Hv_1$ and consider the first component:

$$Mv_2 = M\left(\mathcal{V}\times_3 n^*\times_2 n^*+2z\Gamma n+z^2M\Omega\right) = \mathcal{V}\times_1 M\times_3 n^*\times_2 n^*+2zM\Gamma n+z^2M\Omega$$

For the second component of the second-order term, I first express the Hessian matrix. It is straightforward to see that:

$$H\left[L_i\right](V^{SS}) = \mu\left[\frac{\partial m_{ik}(V^{SS})}{\partial V_h}\right]_{hk},$$

which we have already calculated in the first-order analysis in Section D.2.3. In fact, we have

$$H_{ikh} = \mu\nu\begin{cases} -m_{ik}m_{ih} & \text{if } k\neq h \\ m_{ik}(1-m_{ik}) & \text{if } k=h \end{cases} = \mu\nu(m_{ik}\mathbb{1}_{k=h}-\mathbb{M}_{ikh})\equiv\overline{\mathbb{H}}, \quad (54)$$

where $\mathbb{M}$ is a tensor defined as $\mathbb{M}_{ikh}=m_{ik}m_{ih}$. I obtain

$$(v_1^*Hv_1)_{i=1\ldots I} = \overline{\mathbb{H}}\times_3 v_1^*\times_2 v_1^*.$$

Using Lemma 1 and Lemma 2 in the appendix, I obtain:

$$\overline{\mathbb{H}} \times_3 v_1^* \times_2 v_1^* = \overline{\mathbb{H}} \times_3 v^* \times_2 v^* \times_3 n^* \times_2 n^* + 2z\overline{\mathbb{H}} \times_2 v^* \times_3 \omega^* \times_2 n^* + \overline{\mathbb{H}} \times_3 \omega^* \times_2 \omega^*$$

Collecting all terms, the second-order term of the continuation value of migration is given by

$$\boxed{\begin{aligned}\frac{1}{\varepsilon^2}\big(\mathcal{M}(V_t) - \mathcal{M}(V^{SS})\big)\Big|_2 = &\frac{1}{2}\left[\mathcal{V} \times_1 M + \overline{\mathbb{H}}\right] \times_3 n^* \times_2 n^* \\ &+ z\left[M\Gamma + \overline{\mathbb{H}} \times_2 v^* \times_3 \omega^*\right] \times_2 n^* \\ &+ \frac{1}{2}z^2\left[M\Omega + \overline{\mathbb{H}} \times_3 \omega^* \times_2 \omega^*\right]\end{aligned}} \tag{55}$$

D.3.3 Continuation value from aggregate shocks

Recall that $\mathcal{A}(z) = -\theta z \partial_z + \frac{\sigma^2}{2}\partial_{zz}$ and that $z^2\Omega = z^2\Delta + \Lambda$. I have:

$$\begin{aligned}\mathcal{A}(z)[dv] &= \mathcal{A}(z)\left[\varepsilon z\omega + \frac{\varepsilon^2}{2}(2z\Gamma n + z^2\Delta + \Lambda)\right] = -\varepsilon\theta z\omega + \frac{\varepsilon^2}{2}\Big(-\theta z(2\Gamma n + 2z\Delta) + \frac{\sigma^2}{2}(2\Delta)\Big) \\ &= \varepsilon\Big\{-\theta z\omega\Big\} + \frac{\varepsilon^2}{2}\Big\{\sigma^2\Delta - 2\theta z\Gamma n - 2\theta z^2\Delta\Big\}.\end{aligned}$$

D.3.4 Continuation value from changes in the population distribution

The continuation value for workers from changes in the population distribution is given by:

$$\sum_j \frac{\partial V_i}{\partial N_j}\mu\left(\sum_k m_{kj}(V)N_k - N_j\right) = \sum_j \frac{\partial V_i}{\partial N_j}(M(V)^*N)_j$$

I approximate individual terms and then take the product. First, note that:

$$\left(\frac{\partial V_i}{\partial N_j}\right)_{ij} = v + \frac{\epsilon}{2}\left(\mathcal{V} \times_2 n^* + \mathcal{V} \times_3 n^* + 2z\Gamma\right).$$

Now, fixing j and k, obtain:

$$M_{jk}^*(V^{SS} + dv) = M_{jk}(V^{SS})^* + \nabla[M_{jk}^*](V^{SS})dv + \frac{1}{2}dv^* H\left[M_{jk}^*\right](V^{SS})dv.$$

I have already derived the gradient component in equation (54). To ease notation, define $\mathbb{G}_{jkh} = \overline{\mathbb{H}}_{kjh}$, so that $\Big(\nabla[M_{jk}^*](V^{SS})dv\Big)_{jk} = \mathbb{G} \times_3 dv^*$. For the Hessian, I have $H\left[M_{jk}^*\right]_{hs} = \mu\frac{\partial^2 m_{kj}}{\partial V_h \partial V_s}$. It is dtraightforward to check that

$$\begin{aligned}\frac{\partial^2 m_{kj}}{\partial^2 V_h} &= \nu^2 m_{kj}\left(2m_{kh}m_{kh} - m_{kh}\right) \qquad \text{for } h \neq j \\ \frac{\partial^2 m_{kj}}{\partial V_h \partial V_s} &= \nu^2 m_{kj}\left(2m_{kh}m_{ks}\right) \qquad \text{for } h \neq s \text{ and } h, s \neq j \\ \frac{\partial^2 m_{kj}}{\partial V_h \partial V_j} &= \nu^2 m_{kj}\left(2m_{kj}m_{kh} - m_{kh}\right) \qquad \text{for } h \neq j \\ \frac{\partial^2 m_{kj}}{\partial^2 V_j} &= \nu^2 m_{kj}(1 - m_{kj})\left(1 - 2m_{kj}\right).\end{aligned}$$

Define

$$\mathbb{H}_{jkhs} = \mu \frac{\partial^2 m_{kj}}{\partial V_h \partial V_s} \tag{56}$$

a four-dimensional tensor. Using these expressions, one obtains:

$$\Big(dv^* H \left[M^*_{jk}\right] dv\Big)_{jk} = \mathbb{H} \times_3 dv^* \times_4 dv^*.$$

Then, to leading order (the 0th order term drops out in steady-state):

$$\begin{aligned}
(M(V)^*N)_j &= \sum_k \left(M^*_{jk} + \mathbb{G}_{jk\bullet} \times_3 dv^* + \frac{1}{2}\mathbb{H}_{jk\bullet\bullet} \times_3 dv^* \times_4 dv^* \right)(N_k + \varepsilon n_k) \\
&= \mathbb{G}_{j\bullet\bullet} \times_2 N^* \times_3 dv^* + \varepsilon (M^* n)_j \\
&+ \varepsilon \mathbb{G}_{j\bullet\bullet} \times_2 n^* \times_3 dv^* + \frac{1}{2}\mathbb{H}_{j\bullet\bullet\bullet} \times_2 N \times_3 dv^* \times_4 dv^* \\
&= \varepsilon \Big\{ \mathbb{G}_{j\bullet\bullet} \times_2 N^* \times_3 v_1^* + (M^* n)_j \Big\} \\
&+ \varepsilon^2 \Big\{ \frac{1}{2}\mathbb{G}_{j\bullet\bullet} \times_2 N^* \times_3 v_2^* + \mathbb{G}_{j\bullet\bullet} \times_2 n^* \times_3 v_1^* + \frac{1}{2}\mathbb{H}_{j\bullet\bullet\bullet} \times_2 N \times_3 v_1^* \times_4 v_1^* \Big\} \\
&\equiv \varepsilon \mathcal{M}_j + \varepsilon^2 \mathcal{N}_j
\end{aligned}$$

where 3rd order terms drop out to get to the second line and to the third line. I further obtain

$$\mathcal{M} = \mathbb{G} \times_2 N^* \times_3 v_1^* + M^* n = (\mathbb{G} \times_2 N^* \times_3 v^*) n + z(\mathbb{G} \times_2 N^* \times_3 \omega^*) \equiv \mathbb{F}_1 n + z\mathbb{F}_2,$$

where $\mathbb{F}_1 = \mathbb{G} \times_2 N^* \times_3 v^* + M^*$ and $\mathbb{F}_2 \equiv \mathbb{G} \times_2 N^* \times_3 \omega^*$. Thus, again to leading order,

$$\sum_j \frac{\partial V_i}{\partial N_j}(M(V)^*N)_j = \sum_j \Big\{ v_{ij} + \frac{\varepsilon}{2}\Big[\mathcal{V}_{i\bullet j} \times_2 n^* + \mathcal{V}_{ij\bullet} \times_3 n^* + 2z\Gamma_{ij}\Big]\Big\}\Big\{\varepsilon \mathcal{M}_j + \varepsilon^2 \mathcal{N}_j\Big\}$$

Hence, the second-order term is

$$\frac{1}{\varepsilon^2}\sum_j \frac{\partial V_i}{\partial N_j}(M(V)^*N)_j\Big|_2 = \frac{1}{2}\sum_j \Big[\mathcal{V}_{i\bullet j} \times_2 n^* + \mathcal{V}_{ij\bullet} \times_3 n^* + 2z\Gamma_{ij}\Big]\mathcal{M}_j + \sum_j v_{ij}\mathcal{N}_j$$

The first component satisfies

$$\begin{aligned}
&\sum_j \Big[\mathcal{V}_{i\bullet j} \times_2 n^* + \mathcal{V}_{ij\bullet} \times_3 n^* + 2z\Gamma_{ij}\Big]\mathcal{M}_j = \mathcal{V} \times_2 n^* \times_3 \mathcal{M}^* + \mathcal{V} \times_3 n^* \times_2 \mathcal{M}^* + 2z\Gamma\mathbb{F}_1 n + 2z^2\Gamma\mathbb{F}_2 \\
=\;& \Big\{\mathcal{V} \times_2 \mathbb{F}_1^* + \mathcal{V} \times_3 \mathbb{F}_1^*\Big\} \times_2 n^* \times_3 n^* + 2z\left\{\frac{1}{2}(\mathcal{V} \times_2 \mathbb{F}_2^* + \mathcal{V} \times_3 \mathbb{F}_2^*) + \Gamma\mathbb{F}_1\right\} n + 2z^2\Gamma\mathbb{F}_2
\end{aligned}$$

The second component satisfies

$$
\begin{aligned}
\sum_j v_{ij}\mathcal{N}_j \equiv v\mathcal{N} &= \underbrace{\left(\frac{1}{2}\mathbb{G}\times_1 v\times_2 N^*\right)}_{\equiv\mathbb{X}} v_2 + \underbrace{\left(\mathbb{G}\times_1 v\right)}_{\equiv\mathbb{Y}}\times_2 n^*\times_3 v_1^* + \underbrace{\left(\frac{1}{2}\mathbb{H}\times_1 v\times_2 N^*\right)}_{\equiv\mathbb{Z}}\times_2 v_1^*\times_3 v_1^* \\
&= (\mathcal{V}\times_1\mathbb{X})\times_2 n^*\times_3 n^* + 2z\mathbb{X}\Gamma n + z^2\mathbb{X}\Omega \\
&+ (\mathbb{Y}\times_3 v^*)\times_2 n^*\times_3 n^* + z(\mathbb{Y}\times_3\omega^*)n \\
&+ (\mathbb{Z}\times_2 v^*\times_3 v^*)\times_2 n^*\times_3 n^* + 2z\big(\mathbb{Z}\times_2 v^*\times_3\omega^*\big)n + z^2\mathbb{Z}\times_2\omega^*\times_3\omega^* \\
&= \Big\{\mathcal{V}\times_1\mathbb{X} + \mathbb{Y}\times_3 v^* + \mathbb{Z}\times_2 v^*\times_3 v^*\Big\}\times_2 n^*\times_3 n^* \\
&+ 2z\Big\{\mathbb{X}\Gamma + \frac{1}{2}\mathbb{Y}\times_3\omega^* + \mathbb{Z}\times_2 v^*\times_3\omega^*\Big\}n \\
&+ z^2\Big\{\mathbb{X}\Omega + \mathbb{Z}\times_2\omega^*\times_3\omega^*\Big\}
\end{aligned}
$$

where for the second-to-last equality I have used the symmetry of $\mathbb{H}$ in its last two coordinates, which implies the symmetry of $\mathbb{Z}$ in its last two coordinates.

Collecting all terms, the second-order term of the continuation value from changes in the population distribution is given by:

$$
\boxed{
\begin{aligned}
\frac{1}{\varepsilon^2}\sum_j \frac{\partial V_i}{\partial N_j}(M(V)^*N)_j\Big|_2 &= \Big\{\mathcal{V}\times_1\mathbb{X} + \mathcal{V}\times_2\mathbb{F}_1^* + \mathcal{V}\times_3\mathbb{F}_1^* + \mathbb{Y}\times_3 v^* + \mathbb{Z}\times_2 v^*\times_3 v^*\Big\}\times_2 n^*\times_3 n^* \\
&+ 2z\left\{\frac{1}{2}(\mathcal{V}\times_2\mathbb{F}_2^* + \mathcal{V}\times_3\mathbb{F}_2^*) + \Gamma\mathbb{F}_1 + \mathbb{X}\Gamma + \frac{1}{2}\mathbb{Y}\times_3\omega^* + \mathbb{Z}\times_2 v^*\times_3\omega^*\right\}n \\
&+ z^2\Big\{2\Gamma\mathbb{F}_2 + \mathbb{X}\Omega + \mathbb{Z}\times_2\omega^*\times_3\omega^*\Big\}
\end{aligned}
}
$$

D.3.5 Deterministic SAME

I now apply the method of identifying coefficients. I obtain the following equations for the SAME. Collecting terms,

$$
\rho\frac{1}{2}\mathcal{V} = \frac{1}{2}\mathbb{D}_{UNN} + \frac{1}{2}\Big[\mathcal{V}\times_1 M + \overline{\overline{\mathbb{H}}}\Big] + \Big\{\mathcal{V}\times_1\mathbb{X} + \mathcal{V}\times_2\mathbb{F}_1^* + \mathcal{V}\times_3\mathbb{F}_1^* + \mathbb{Y}\times_3 v^* + \mathbb{Z}\times_2 v^*\times_3 v^*\Big\}.
$$

Re-arranging, and denoting the unknown deterministic Impulse Value $\boldsymbol{\mathcal{V}}$ in bold orange,

$$
\boldsymbol{\mathcal{V}}\times_1(\rho\mathrm{Id} - M - 2\mathbb{X}) - \boldsymbol{\mathcal{V}}\times_2(2\mathbb{F}_1)^* - \boldsymbol{\mathcal{V}}\times_3(2\mathbb{F}_1)^* = \mathbb{D}_{UNN} + \overline{\overline{\mathbb{H}}} + 2(\mathbb{Y}\times_3 v^* + \mathbb{Z}\times_2 v^*\times_3 v^*),
$$

which is a standard tensor Sylvester equation. Re-arranging,

$$
\boldsymbol{\mathcal{V}}\times_1 A + \boldsymbol{\mathcal{V}}\times_2 B + \boldsymbol{\mathcal{V}}\times_3 B = D, \tag{57}
$$

where

$$
A = M + 2\mathbb{X} - \rho\mathrm{Id}, \qquad B = 2\mathbb{F}_1^*, \qquad -D = \mathbb{D}_{UNN} + \overline{\overline{\mathbb{H}}} + 2(\mathbb{Y}\times_3 v^* + \mathbb{Z}\times_2 v^*\times_3 v^*).
$$

D.3.6 Cross SAME

Collecting terms,

$$(\rho+\theta)\Gamma = -D_{UNZ} + M\Gamma + \overline{\mathbb{H}} \times_2 v^* \times_3 \omega^*) + (\mathcal{V} \times_2 \mathbb{F}_2^* + \mathcal{V} \times_3 \mathbb{F}_2^*) + 2\Gamma\mathbb{F}_1 + 2\mathbb{X}\Gamma + \mathbb{Y} \times_3 \omega^* + 2\mathbb{Z} \times_2 v^* \times_3 \omega^*.$$

Re-arranging, and denoting the unknown cross Impulse Value $\boldsymbol{\Gamma}$ in bold orange and the now known deterministic Impulse Value $\boldsymbol{\mathcal{V}}$ in bold blue,

$$\begin{aligned} & ((\rho+\theta)\mathrm{Id} - M - 2\mathbb{X})\boldsymbol{\Gamma} - \boldsymbol{\Gamma}(2\mathbb{F}_1) \\ = \; & -D_{UNZ} + \overline{\mathbb{H}} \times_2 v^* \times_3 \omega^* + (\boldsymbol{\mathcal{V}} \times_2 \mathbb{F}_2^* + \boldsymbol{\mathcal{V}} \times_3 \mathbb{F}_2^*) + \mathbb{Y} \times_3 \omega^* + 2\mathbb{Z} \times_2 v^* \times_3 \omega^*, \end{aligned}$$

which is a standard matrix Sylvester equation. Re-arranging,

$$(A - \theta\mathrm{Id})\boldsymbol{\Gamma} + \boldsymbol{\Gamma}B^* = C,$$

where

$$-C = -D_{UNZ} + \overline{\mathbb{H}} \times_2 v^* \times_3 \omega^* + (\boldsymbol{\mathcal{V}} \times_2 \mathbb{F}_2^* + \boldsymbol{\mathcal{V}} \times_3 \mathbb{F}_2^*) + \mathbb{Y} \times_3 \omega^* + 2\mathbb{Z} \times_2 v^* \times_3 \omega^*.$$

D.3.7 Stochastic SAME

I first identify coefficients for the combined term Ω.

$$\rho\frac{1}{2}\Omega \;=\; \frac{1}{2}D_{UZZ} + \frac{1}{2}\left[M\Omega + \overline{\mathbb{H}} \times_2 \omega^* \times_3 \omega^*)\right] + 2\Gamma\mathbb{F}_2 + \mathbb{X}\Omega + \mathbb{Z} \times_2 \omega^* \times_3 \omega^* + \frac{1}{2}\left\{\frac{\sigma^2\Omega}{z^2} - 2\theta\Omega\right\}.$$

Now I separately identify coefficients for Δ and Λ. I start with Δ. Collecting terms,

$$\rho\frac{1}{2}\Delta \;=\; \frac{1}{2}D_{UZZ} + \frac{1}{2}\left[M\Delta + \overline{\mathbb{H}} \times_2 \omega^* \times_3 \omega^*)\right] + 2\Gamma\mathbb{F}_2 + \mathbb{X}\Delta + \mathbb{Z} \times_2 \omega^* \times_3 \omega^* - \theta\Delta.$$

Re-arranging, and denoting the unknown stochastic Impulse Value $\boldsymbol{\Delta}$ in bold orange and the now known cross Impulse Value $\boldsymbol{\mathcal{G}}$ in bold blue,

$$((\rho+2\theta)\mathrm{Id} - M - 2\mathbb{X})\boldsymbol{\Delta} \;=\; D_{UZZ} + 4\boldsymbol{\Gamma}\mathbb{F}_2 + \mathbb{Z} \times_2 \omega^* \times_3 \omega^*,$$

which is a standard vector equation. Re-arranging,

$$(A - 2\theta\mathrm{Id})\boldsymbol{\Delta} = E, \qquad\qquad -E = D_{UZZ} + 4\boldsymbol{\Gamma}\mathbb{F}_2 + \mathbb{Z} \times_2 \omega^* \times_3 \omega^*.$$

Now for Λ, collecting terms,

$$\rho\frac{1}{2}\Lambda \;=\; \frac{1}{2}\sigma^2\Delta + \frac{1}{2}M\Lambda + \mathbb{X}\Lambda.$$

Re-arranging, and denoting the unknown (intercept) stochastic Impulse Value $\boldsymbol{\Lambda}$ in bold orange and the now known (slope) stochastic Impulse Value $\boldsymbol{\Delta}$ in bold blue,

$$A\boldsymbol{\Lambda} \;=\; -\sigma^2\boldsymbol{\Delta},$$

which is again a standard vector equation.

D.3.8 KFE

The law of motion for population is given by (48). I already computed its second-order expansion in Section D.3.4. I obtain:

$$\frac{dn_t}{dt} = \overbrace{M^* n_t + G v_{1,t}}^{\text{first-order term}} + \overbrace{\frac{\varepsilon}{2}\mathcal{N}_t}^{\text{second-order term}},$$

$$\mathcal{N}_t = \frac{1}{2}\mathbb{G} \times_2 N^* \times_3 v^*_{2,t} + \mathbb{G} \times_2 n^*_t \times_3 v^*_{1,t} + \frac{1}{2}\mathbb{H} \times_2 N^* \times_3 v^*_{1,t} \times_4 v^*_{1,t},$$

where we construct the first- and second-order components of the value function vector $v_{1,t}$ and $v_{2,t}$ in each period.

To interpret the impulse response of the sole effect of a shock, it is important to adjust the steady-state welfare by Λ, which is a constant and does not depend on the time path of the shock. Otherwise, the interpretation of the impulse response is that of a news shock on top of the actual shock, whereby individuals learn of aggregate risk at the same time as a shock hits.

Online Appendix

E Weak derivatives and duality

E.1 Weak derivatives

Let f be a η-measurable function. Its weak derivative $\frac{\partial f}{\partial x_i}$ is, when it exists, defined by duality. Suppose that there exists a η-measurable function w_i such that

$$\int f(x)\frac{\partial \varphi}{\partial x_i}(x)d\eta(x) = -\int w_i(x)\varphi(x)d\eta(x)$$

for all continuously differentiable functions φ that vanish on $\widehat{\partial X}$. In that case, define its weak derivative $\frac{\partial f}{\partial x_i}$ to be w_i.

Unidimensional case. Consider a domain $(-1,1)$ with a possible mass point at 0. Compute, for a smooth function φ such that $\varphi(-1) = \varphi(1) = 0$,

$$\langle f, \varphi' \rangle = \int_{-1}^{0^-} f(x)\varphi'(x)dx + f_0\varphi_0' + \int_{0^+}^{1} f(x)\varphi'(x)dx$$

where I denote evaluation at the mass point by subscripts to emphasize the role of (non-)smoothness at the possible mass points. Assuming f is differentiable on $(-1,0)$ and $(0,1)$, obtain $\langle f, \varphi' \rangle = f(0^-)\varphi(0^-) - \int_{-1}^{0^-} f'(x)\varphi(x)dx + f_0\varphi_0' - f(0^+)\varphi(0^+) - \int_{0^+}^{1} f'(x)\varphi(x)dx$. Finally, make $\langle f', \varphi \rangle$ appear: $\langle f, \varphi' \rangle = f(0^-)\varphi(0^-) + f_0'\varphi_0 + f_0\varphi_0' - f(0^+)\varphi(0^+) - \langle f', \varphi \rangle \equiv J_0 - \langle f', \varphi \rangle$. The key object of interest is therefore $J_0 \equiv f(0^-)\varphi(0^-) + f_0'\varphi_0 + f_0\varphi_0' - f(0^+)\varphi(0^+)$. The duality property requires that the sum of terms around 0, J_0, is equal to zero.

Smooth φ. When φ is continuously differentiable on $(-1,1)$, and in particular is smooth around 0, then $J_0 = \Big[f_0' - \big(f(0^+) - f(0^-)\big)\Big]\varphi(0) + f_0\varphi_0'$. So clearly, for J_0 to be zero and f to have a weak derivative, one needs

$$f_0' \equiv f(0^+) - f(0^-) \quad , \quad \varphi_0' \equiv 0 = \varphi(0^+) - \varphi(0^-).$$

Thus, the definition of the weak derivative w.r.t. the base measure η imposes that the value of the derivative at possible mass points is equal to the jump there. In particular, if a function is continuous at a possible mass point, then the derivative there that enters into the inner product (but not around) is endogenously 0. It need not be a requirement.

Multidimensional case. This argument generalizes straightforwardly to multiple dimensions.

E.2 Borrowing constraints and duality

For concreteness, consider the Aiyagari (1994) economy with a borrowing constraint $a \geq \underline{a}$. The asset domain is $[\underline{a}, \overline{a}]$. For simplicity, focus only on the asset dimension. Consider the operators associated with the asset drift $L \equiv s(a)\partial_a$ and $L^* \equiv -\partial_a s(a)$, where s denotes the savings rate. Consider a base measure $d\eta(a) = \delta_{\underline{a}}(a) + da$. A distribution is given by a density $\{g(\underline{a}); g(a), a > \underline{a}\}$, so that the probability measure that represents the distribution of individuals is $g(\underline{a})\delta_{\underline{a}}(a) + g(a)da$. For any function f, I denote by $f(\underline{a}^+) = \lim_{a \downarrow \underline{a}} f(a)$.

A mass point arises at $\underline{a}$ if $s(a) \leq 0$ in a neighborhood of $\underline{a}$. The borrowing constraint imposes $s(\underline{a}) = 0$. The presence of a mass point requires that $g(\underline{a}^+) = +\infty$ so that the (asset-induced) inflow into the mass point, $(sg)(\underline{a}^+)$, is finite and non-zero. Then the weak derivative of sg at $\underline{a}$ (with respect to the base measure η) is:

$$\partial_a(s(a)g(a)) = \begin{cases} (sg)(\underline{a}^+) & \text{if} \quad a = \underline{a} \\ \text{the classical derivative } \partial_a(s(a)g(a)) & \text{if} \quad a > \underline{a} \end{cases}$$

Thus, for any smooth test function and any g, g' that vanish at $a = \overline{a}$:

$$\begin{aligned} \int_{\underline{a}}^{\overline{a}} \varphi(a)L^*(a)[g]d\eta(a) &= -\varphi(\underline{a})s(\underline{a})g(\underline{a}) - \int_{\underline{a}^+}^{\overline{a}} \varphi(a)\partial_a(s(a)g(a))da \\ &= \int_{\underline{a}^+}^{\overline{a}} s(a)g(a)\partial_a\varphi(a)da + (sg)(\underline{a}^+)\varphi(\underline{a}) - \varphi(\underline{a})s(\underline{a})g(\underline{a}) \\ &= \int_{\underline{a}^+}^{\overline{a}} L(a)[\varphi]g(a)da + (sg)(\underline{a}^+)\varphi(\underline{a}) - \varphi(\underline{a})s(\underline{a})g(\underline{a}) \\ &= \int_{\underline{a}}^{\overline{a}} L(a)[\varphi]g(a)d\eta(a) + (sg)(\underline{a}^+)\varphi(\underline{a}) - \underbrace{s(\underline{a})g(\underline{a})\big(\varphi(\underline{a}^+) + \varphi(\underline{a})\big)}_{=0 \text{ because } s(\underline{a})=0} \\ &= \int_{\underline{a}}^{\overline{a}} L(a)[\varphi]g(a)d\eta(a) + (sg)(\underline{a}^+)\varphi(\underline{a}) \\ &= \int_{\underline{a}}^{\overline{a}} L(a)[\varphi]g(a)d\eta(a) + \int_{\underline{a}}^{\overline{a}} (sg)(\underline{a}^+)\varphi(\underline{a})\delta_{\underline{a}}(da), \end{aligned}$$

where the first line uses the definition of $d\eta(a)$ and the expression for the weak derivative of sg. The second line integrates the second term by parts. The third line uses the definition of L and of the weak derivative of φ at $\underline{a}$ with respect to η. The fourth line uses the definition of η. The last line uses the identity $1 = \int_{\underline{a}}^{\overline{a}} \delta_{\underline{a}}(da)$.

A similar derivation delivers the duality formula when there are kinks in the interior of the domain. In that case, $(sg)(\underline{a}^+)$ must be replaced with $(sg)(\underline{a}^+) - (sg)(\underline{a}^-)$.

F Tensor Algebra

F.1 Conventions

We choose the convention that when we mutliply a tensor by a matrix, the summation is with respect to the second entry of the matrix. In addition, we put the remaining index of the matrix in the position that corresponds to the product index. That is,

$$\begin{aligned}
(\mathbb{T} \times_1 A)_{ijk} &= \sum_\ell A_{i\ell} \mathbb{T}_{\ell jk} \\
(\mathbb{T} \times_2 A)_{ijk} &= \sum_\ell A_{j\ell} \mathbb{T}_{i\ell k} \\
(\mathbb{T} \times_3 A)_{ijk} &= \sum_\ell A_{k\ell} \mathbb{T}_{ij\ell}
\end{aligned}$$

F.2 Basic results

Lemma 1. *Let $\mathbb{T} \in \mathbb{R}^{I\times I\times I}$ be a tensor, and let x and y be vectors in $\mathbb{R}^I$. Then:*

$$\mathbb{T} \times_3 (x+y)^* \times_2 (x+y)^* = \mathbb{T} \times_3 x^* \times_2 x^* + (\mathbb{T} + \mathbb{T}^*) \times_3 x^* \times_2 y^* + \mathbb{T} \times_3 y^* \times_2 y^*,$$

where $\mathbb{T}^*$ is the transpose of $\mathbb{T}$ with respect to its last two coordinates. Moreover, if $\mathbb{T}$ is symmetric in its last two coordinates we will have that:

$$\mathbb{T} \times_3 x^* \times_2 y^* = \mathbb{T} \times_3 y^* \times_2 x^*, \qquad \mathbb{T} \times_3 x^* \times_2 y^* = \mathbb{T} \times_2 x^* \times_3 y^*$$

and

$$\mathbb{T} \times_3 (x+y)^* \times_2 (x+y)^* = \mathbb{T} \times_3 x^* \times_2 x^* + 2\mathbb{T} \times_3 x^* \times_2 y^* + \mathbb{T} \times_3 y^* \times_2 y^*,$$

Lemma 2. *Let $\mathbb{T} \in \mathbb{R}^{I\times J\times K}$ be a tensor. Let $A \in \mathbb{R}^{P\times I}$ be a matrix and $x \in \mathbb{R}^J$ and $y \in \mathbb{R}^K$ be vectors. Then:*

$$A\,(\mathbb{T} \times_3 x^* \times_2 y^*) = (\mathbb{T} \times_1 A) \times_3 x^* \times_2 y^*$$

Lemma 3. *Let $\mathbb{T} \in \mathbb{R}^{I\times J\times K}$ be a tensor. Let $A \in \mathbb{R}^{K\times P}$, $B \in \mathbb{R}^{J\times L}$ be a matrices and $x \in \mathbb{R}^P$ and $y \in \mathbb{R}^L$ be vectors. Then:*

$$\mathbb{T} \times_3 (Ax)^* \times_2 (By)^* = \mathbb{T} \times_3 A^* \times_2 B^* \times_3 x^* \times_2 y^*$$