A Perturbational Approach for Approximating Heterogeneous Agent Models*

Anmol Bhandari
University of Minnesota

Thomas Bourany
University of Chicago

David Evans
University of Oregon

Mikhail Golosov
University of Chicago

March 2023

Abstract

We develop a perturbational technique to approximate equilibria of a wide class of discrete-time dynamic stochastic general equilibrium heterogeneous-agent models with complex state spaces, including multi-dimensional distributions of endogenous variables. We show that approximating policy functions and stochastic process that governs the distributional state to any order is equivalent to solving small systems of linear equations that characterize values of certain directional derivatives. We analytically derive the coefficients of these linear systems and show that they satisfy simple recursive relations making their numerical implementation quick and efficient. Compared to existing state-of-the-art techniques, our method is faster in constructing first-order approximations and extends to higher orders, capturing the effects of risk that are ignored by many current methods. We apply our method to a broad set of questions such as impacts of first- and second-moment shocks, welfare effect of macroeconomic risk and stabilization policies, endogenous household portfolio formation, and transition dynamics in heterogeneous agent general equilibrium settings.

*We thank Adrien Auclert, participants of SITE conference, and seminar participants at Bank of Canada, Bank of Japan, Bocconi, EIEF, EUI, Minneapolis Fed, University of Chicago, and University of Oregon for helpful comments. Bhandari, Evans, and Golosov thank the NSF for support (grant #36354.00.00.00).

1 Introduction

We develop a numerical method to approximate equilibrium dynamics of a large class of discrete-time heterogeneous agent (HA) models that feature aggregate and idiosyncratic shocks, and occasionally binding borrowing constraints. Our method is based on approximation techniques that scale aggregate shocks and consider Taylor expansions of equilibrium conditions with respect to that scaling parameter. Our method can be used to quickly compute equilibrium effects of higher-moment shocks, welfare effects of risk and macroeconomic stabilization policies, general equilibrium portfolio problems, and transition dynamics to a new steady state in environments with rich heterogeneity.

Our approach intentionally focuses on maintaining the computational speed, flexibility, and ease-of-use of traditional perturbational techniques commonly employed to solve and estimate representative agent dynamic stochastic general equilibrium models, such as those implemented with the DYNARE software package.[1] We use recursive representations of equilibria and approximate equilibrium dynamics around a non-stochastic steady state. In HA models, this steady state includes the invariant distribution of individual state variables, which is a large- (usually, infinite-) dimensional object.

As the first step, we *analytically* derive expressions that fully characterize equilibrium dynamics to an arbitrary order of approximation.[2] These expressions have a simple mathematical structure and proceed inductively. The equilibrium dynamics of a given order of approximation can be obtained by solving a small dimensional linear system of equations. The coefficients in that system depend entirely on objects solved using lower orders of approximation, and are related to each other via recursive relationships. Our analytical expressions show that they depend only on the non-stochastic state-transition kernel and policy functions that characterize the invariant distribution. Existing off-the-shelf numerical methods to solve HA models without aggregate shocks routinely compute such objects. We show that these objects plus the equations describing optimality and market clearing conditions can be used to construct required coefficients quickly and efficiently.

We first describe our approach in a canonical HA economy in the spirit of Krusell and Smith (1998) and show how to apply our method to find its first- and second-order approximations. As a by-product of this analysis, we also show how to use our approach to characterize transition dynamics to a new steady state, and to find equilibrium responses to second moment (volatility) shocks. We then extend our approach to portfolio problems, i.e., general equilibrium models in which agents can invest in more than one asset with different risk characteristics. Extending perturbational techniques to such problems is challenging. Portfolio choices depend on the second-order properties of the model, such as risk premia or covariances of equilibrium variables, but even the first-order equilibrium dynamics are affected by

[1] See Judd (1998) and Schmitt-Grohé and Uribe (2004) for an introduction to such methods.

[2] In the paper, we focus on the first and second orders, but it is straightforward to extend our approach to higher orders.

those choices.[3] This breaks the convenient feature of perturbational techniques whereby the n^{th} order of approximation provides all information necessary to find the $(n+1)^{th}$ approximation order. We show that this difficulty can be sidestepped because it is possible to simultaneously find agents' optimal portfolios, first-order approximation, and the second-order risk premium.

Our method builds on the perturbational techniques developed by Judd (1998) and Schmitt-Grohé and Uribe (2004). When applied to HA models, the key bottleneck is in finding derivatives of policy functions with respect to the state variable. When the number of state variables is small, finding this derivative requires solving a quadratic matrix equation and choosing the stable roots. The theoretical analogue of this object in HA environments is an infinite-dimensional Fréchet derivative that solves a quadratic functional equation that is impossible to compute. A number of papers have tried to overcome this problem by simplifying representation of the aggregate state. Reiter (2009), in his seminal work, uses the histogram method to represent the invariant distribution as a finite number of mass points. Subsequent work, e.g., Ahn et al. (2018); Childers (2018); Winberry (2018); Gornemann et al. (2021); Bayer et al. (2022), further speeds up Reiter's approach by pursuing additional model reduction steps. Despite significant progress, this approach remains fairly slow and, as we explain shortly, does not generally extend beyond the first order.

Our approach is different from this literature as we first derive exact analytical expressions for approximation terms before implementing them numerically. The key insight of those analytical expressions is that one does not need to know the whole Fréchet derivative of policy functions. It is sufficient to find values of those derivative in a sequence of appropriately chosen directions. This distinction is important: while the Fréchet derivative of a policy function for a certain variable is an infinitely dimensional functional, the value of that derivative in a given direction is just a scalar. This dramatically reduces the dimensionality of the system that one needs to solve in order build the approximation. Moreover, those approximations solve linear rather than non-linear equations, do not require finding stable roots or applying pruning algorithms that are necessary for higher order approximations under standard perturbational methods (see, e.g., Kim et al. (2008)). These advantages allow one to quickly find approximations even in settings with very rich heterogeneity.

Deriving analytical characterization is also important to ensure scalability of approximation techniques beyond the first order. We show that applying perturbational techniques to a discretized histogram and its associated state transition matrix would return incorrect approximations starting from the second order. By locally linearizing policy functions, the histogram method misses some of the terms

[3] Recent papers (e.g., Kaplan et al. (2018)) have emphasized the importance of multiple assets in quantitative HA environments. These papers abstract from risk premium and the allocation of savings into different assets in those models is determined by heterogeneity of costs of trading of those assets. Although, heterogeneous trading costs is straightforward to incorporate into our baseline specification, the challenge really is to develop a computational technique that also allows heterogeneity in risk characteristics to affect agents' portfolio allocations.

that must appear in a second- or higher-order expansion. In a calibrated Krusell-Smith economy we show that these missing terms can be quantitatively large and significantly alter economic conclusions drawn from numerical experiments.

Our technique is related to the approximation method developed by Boppart et al. (2018) and Auclert et al. (2021), or ABRS. Those authors consider the sequence-space formulation of equilibria and use impulse responses to one-time, unanticipated "MIT shocks" to recover a first-order approximation. ABRS show that these responses are described by a linear system of equations that can be solved much faster than approximations using perturbational techniques in the spirit of Reiter (2009). There is a certain equivalence between our approach and that of ABRS. In particular, one can show that as the grid size used in ABRS's discretization goes to zero, their system of equations converges to the analytical expressions we derive. Despite this equivalence, our technique has two advantages. Firstly, our recursive approach improves on the computational speed of ABRS's algorithm to compute first-order approximations. Secondly, and more importantly, our approach naturally extends to higher orders and thus expands the scope of analysis beyond MIT shocks.

Our paper is also related to the approximation method developed in Bhandari et al. (2021). Like us, those authors use perturbational methods to derive analytically various orders of approximations of equilibrium in HA economy, and then find those expressions numerically. Their approximation scaled both aggregate and idiosyncratic shocks and it is not applicable to models in which policy functions have kinks, for example due to the occasionally binding borrowing constraints. Our approach instead approximates only with respect to aggregate shocks. This improves the approximation precision, since our approach remains global with respect to idiosyncratic shocks, and allows us to incorporate such kinks. It also makes analytical characterization of approximation terms significantly more challenging. Deriving those analytical expressions to build the approximations when policy functions are not differentiable is one of the key contributions of this paper.

Our approach, like all perturbational methods, is local as it seeks to find equilibrium dynamics when aggregate shocks are small and the economy is near its steady state. Our goal is to preserve key advantages of these methods – computational speed, simplicity, and flexibility – in HA settings. There exists a complementary strand of literature that aims to develop global methods. Such methods can be used to find equilibria without requiring them to be nearby any specific economy but they tend to be slower, harder to use, and often need to be tailored to the specific economic environment.[4]

The class of economies that we consider in this paper is discrete-time infinite horizon HA models with distributional states. There is a parallel literature that studies continuous-time versions of these

[4]Krusell and Smith (1998) solve their economy using a global method that proved difficult to extend to general HA settings. Some recent work extends global solution methods to more complex environments using machine learning techniques. See Maliar et al. (2021), Kahou et al. (2021), Childers et al. (2022), and Han et al. (2021) for details.

economies. See, for instance, Kaplan et al. (2018), Achdou et al. (2020), Ahn et al. (2018) in the context of consumption-savings models; Alvarez and Lippi (2022) and Alvarez et al. (2023a) in the context of price-setting models; and Bigio et al. (2023) for an application to public debt maturity. In related work, Bilal (2023) and Alvarez et al. (2023b) use mean field game techniques to construct approximations with aggregate shocks in these class of models. Their work shares with us the use of linear operators over infinite-dimensional spaces to characterize the exact derivatives analytically.

The rest of the paper is organized as follows. Section 2 presents our baseline environment, Section 3 and 4 describes our approximation techniques in that environment, Section 5 show how these techniques can be extended to models of transition dynamics, stochastic volatility, and portfolio problems. In Section 6 we discuss relationship to different strands of literature in more details. Section 7 provides numerical illustration of our techniques. Section 8 concludes.

2 Environment

In this paper we develop a novel method to obtain numerical approximations of equilibrium in wide class of heterogeneous agent economies with aggregate and idiosyncratic shocks. To motivate our representation of equilibrium conditions, we start with a familiar economy.

2.1 Prototypical Krusell and Smith economy

Consider the setting studied in Krusell and Smith (1998). The economy is populated by a continuum of heterogeneous households and homogeneous firms. Each household supplies inelastically one unit of labor that is subject to idiosyncratic efficiency shocks $\theta_{i,t}$. Households receive wage W_t and save capital $k_{i,t}$ that earns gross return R_t, that includes rental rate net of depreciation δ. Household i chooses stochastic sequences $\{c_{i,t}, k_{i,t}\}_t$ to maximize utility function $\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t U(c_{i,t})$ subject to the budget constraint $c_{i,t} + k_{i,t} \leq R_t k_{i,t-1} + W_t \exp(\theta_{i,t})$ and the borrowing constraint $k_{i,t} \geq 0$. Initial $k_{i,-1}$ and $\theta_{i,0}$ are given and the distribution of $\{k_{i,-1}, \theta_{i,0}\}_i$ over i is denoted by Ω_0. Efficiency $\theta_{i,t}$ follows an exogenous stationary stochastic process normalized so that $\int \exp(\theta_{i,t}) di = 1$.

Households rent capital and supply efficiency-adjusted labor to firms each period. Firms are competitive and produce output using Cobb-Douglas technology with aggregate productivity $\exp(\Theta_t)$ and capital share of α. Wages W_t and rental rates R_t are determined by the market clearing conditions so that supply of labor and capital by consumers is equal to the demand for those factors by firms.

The equilibrium in this economy can be represented by three set of conditions: the optimality conditions of heterogeneous households, the optimality conditions of firms, and the market clearing conditions. It would be helpful to keep conditions characterizing behavior of households separately from the other conditions. Let $\zeta_{i,t}$ be the Lagrange multiplier on the borrowing constraint of household

i in period t. We can write the optimality conditions of heterogeneous households as

$$R_t U_c(c_{i,t}) - \lambda_{i,t} = 0, \quad U_c(c_{i,t}) + \zeta_{i,t} - \beta \mathbb{E}_t \lambda_{i,t+1} = 0 \text{ for all } i, t, \tag{1}$$

and

$$c_{i,t} + k_{i,t} - R_t k_{i,t-1} - W_t \exp(\theta_{i,t}) = 0, \quad k_{i,t}\zeta_{i,t} = 0 \text{ for all } i, t. \tag{2}$$

Here, equation (1) is households' Euler equations written in a form that will be particularly convenient for numerical analysis, equation (2) has the budget constraint, and the the last equation is the complementary slackness.

The optimality conditions of firms and market clearing conditions are

$$W_t - (1-\alpha)\exp(\Theta_t) K_{t-1}^{\alpha} = 0, \quad K_t - \int k_{i,t} di = 0 \text{ for all } t, \tag{3}$$

$$R_t + \delta - 1 - \alpha \exp(\Theta_t) K_{t-1}^{\alpha-1} = 0 \text{ for all } t. \tag{4}$$

Given initial conditions $\{k_{i,-1}, \theta_{i,0}\}_i$ and Θ_0, equations (1), (2), (3) and (4) fully summarize the equilibrium dynamics of this economy.

2.2 The sequence-space representation

Motivated by this example, we now present a general representation of equilibrium conditions of a broad class of HA economies. Let θ and Θ be vectors of idiosyncratic and aggregate shocks. Let x be the vector of endogenous variables chosen by the agents subject to idiosyncratic shocks and X be the vector of all other endogenous variables. We refer to x and X as idiosyncratic and aggregate endogenous variables. Let $a_{i,t-1} \in x_{i,t-1}$ and $A_{t-1} \in X_{t-1}$ be vectors of individual and aggregate endogenous variables that enter into the the time t equilibrium conditions. We will write them explicitly as

$$a = \mathsf{p}x, \quad A = \mathsf{P}X$$

for some selection matrices p and P. Let Y_t be the stacked vector of all relevant aggregates, $Y_t := [\Theta_t, A_{t-1}, X_t, \mathbb{E}_t X_{t+1}]^{\mathrm{T}}$.[5]

The optimality conditions of heterogeneous agents are represented as

$$F(a_{i,t-1}, \theta_{i,t}, x_{i,t}, \mathbb{E}_{i,t} x_{i,t+1}, Y_t) = 0 \text{ for all } i, t \tag{5}$$

which initial conditions $(a_{i,-1}, \theta_{i,0})$. Let Ω_0 the the (cumulative) distribution of $(a_{i,-1}, \theta_{i,0})$. The remaining equilibrium conditions, that include optimality conditions of agents not subject to idiosyncratic shocks, market clearing conditions, budget constraints for the government, etc, are represented as

$$G\left(\int x_i di, Y_t\right) = 0 \text{ for all } t \tag{6}$$

[5]Keeping track of variables such as A_{t-1} and $\mathbb{E}_t X_{t+1}$ in Y_t allows us to represent equilibria of economies with slow adjustment of aggregates such as settings with capital adjustment costs, and various types of real and nominal rigidities.

with some initial Θ_0 and A_{-1}.

It is easy to see how our example of the Krusell and Smith economy fits into this representation. In that example, we have $x_{i,t} = [k_{i,t}, c_{i,t}, \lambda_{i,t}, \zeta_{i,t}]^{\mathrm{T}}$, $X_t = [K_t, W_t, R_t]^{\mathrm{T}}$, $a_{i,t} = k_{i,t}$, $A_t = K_t$. With these definitions, mapping F captures optimality conditions (1) and (2), while mapping G captures conditions (3) and (4).

We postulate that individual optimality conditions F have an argument $\mathbb{E}_{i,t}x_{i,t+1}$ as opposed to a seemingly more general argument $\mathbb{E}_{i,t}f(x_{i,t+1}, \theta_{i,t+1}, Y_{t+1})$ for some nonlinear vector-valued function f. This choice is without loss of generality, as we can always define a new vector $x'_{i,t} := f(x_{i,t}, \theta_{i,t}, Y_t)$ and incorporate it into $x_{i,t}$ so that the optimality conditions take form (5). For instance, in the Krusell and Smith economy, we used the variable $\lambda_{i,t}$ to represent household's Euler equation $U_c(c_{i,t}) + \zeta_{i,t} = \beta\mathbb{E}_{i,t}U_c(c_{i,t+1})R_{t+1}$ in this form (see equation (1)). The linearity embedded in $\mathbb{E}_{i,t}x_{i,t+1}$ is convenient numerically and increases the computation speed of the method that we develop.

In order to streamline our exposition, for now we assume that $\theta_{i,t}$, Θ_t and $a_{i,t}$ are scalars; we discuss the case when they are finite vectors in the appendix. We assume that $\theta_{i,t}$ and Θ_t follow AR(1) processes

$$\Theta_t = \rho_\Theta \Theta_{t-1} + \mathcal{E}_t, \tag{7}$$

$$\theta_{i,t} = \rho_\theta \theta_{i,t-1} + \varepsilon_{i,t}. \tag{8}$$

Here, $\mathcal{E}_t$ is a mean-zero random variable drawn independently across time from a distribution with bounded support, $\varepsilon_{i,t}$ is a mean zero random variable drawn independently across time and agents with some probability distribution μ, coefficients ρ_Θ and ρ_θ satisfy $|\rho_\Theta|, |\rho_\theta| < 1$.

An *equilibrium* consists of stochastic processes $\{X_t(\mathcal{E}^t)\}_{t,\mathcal{E}^t}$ and $\{x_{i,t}(\mathcal{E}^t, \varepsilon_i^t)\}_{i,t,\mathcal{E}^t,\varepsilon^t}$ that satisfy (5) − (8) given some initial conditions $Z_0 = [\Theta_0, A_{-1}, \Omega_0]$. Our main focus is on characterizing the equilibrium stochastic process $\{X_t\}_t$, which is relevant for most macroeconomic applications. As a by product, we also describe a procedure to recover the stochastic processes $\{x_{i,t}\}_{i,t}$. We refer to equations (5) and (6) as the *sequence-space representation* of equilibrium.

2.3 The state-space representation

Significant gains in both computational speed and flexibility of our approach can be attained by using a recursive representation of the equilibrium. The aggregate state of the system (5) and (6) consists of Θ_t, A_{t-1} and the joint distribution Ω_t over $\{(a_{i,t-1}, \theta_{i,t})\}_i$. We use $\Omega_t\langle a, \theta\rangle$ to denote the mass of agents with $\theta_{i,t} \leq \theta$ and $a_{i,t-1} \leq a$. Let $Z_t = [\Theta_t, A_{t-1}, \Omega_t]^{\mathrm{T}}$ be the aggregate state.[6] We use tildes to denote policy functions in the recursive representation. Thus, $\widetilde{X}(Z)$ are policy functions for aggregate

[6]It is possible for some of the variables in Z_t to be redundant. For example, in the Krusell and Smith economy $A_{t-1} = \int a d\Omega_t$

variables and $\widetilde{x}(a,\theta,Z)$ are policy functions for individual variables. The recursive representation of equilibrium conditions is given by

$$F\left(a,\theta,\widetilde{x},\mathbb{E}_{\varepsilon,\mathcal{E}}\widetilde{x},\widetilde{Y}\right)=0 \text{ for all } (a,\theta,Z), \tag{9}$$

and

$$G\left(\int \widetilde{x}d\Omega,\widetilde{Y}\right)=0 \text{ for all } Z, \tag{10}$$

as well as the *Law of Motion* (LoM) for the aggregate distribution $\widetilde{\Omega}(Z)$ defined as

$$\widetilde{\Omega}(Z)\langle a',\theta'\rangle=\int\int \iota\left(\widetilde{a}(a,\theta,Z)\le a'\right)\iota(\rho_\theta\theta+\varepsilon\le\theta')\mu(\varepsilon)\,d\varepsilon d\Omega\langle a,\theta\rangle \text{ for all } Z, \tag{11}$$

and identities

$$\widetilde{a}=\mathsf{p}\widetilde{x},\quad \widetilde{A}=\mathsf{P}\widetilde{X},\quad \widetilde{Y}=\left[\Theta,A,\widetilde{X},\mathbb{E}_{\mathcal{E}}\widetilde{X}\right]^{\mathrm{T}}. \tag{12}$$

We use $\mathbb{E}_{\varepsilon,\mathcal{E}}$ and $\mathbb{E}_{\mathcal{E}}$ to denote conditional expectation of future policies with respect to $(\varepsilon,\mathcal{E})$ and $\mathcal{E}$, respectively, that is,

$$\mathbb{E}_{\varepsilon,\mathcal{E}}\widetilde{x}=\int \widetilde{x}\left(\tilde{a}(a,\theta,Z),\rho_\theta\theta+\varepsilon,\rho_\Theta\Theta+\mathcal{E},\widetilde{A}(Z),\widetilde{\Omega}(Z)\right)d\mu(\varepsilon)\,d\Pr(\mathcal{E}),$$

$$\mathbb{E}_{\mathcal{E}}\widetilde{X}=\int \widetilde{X}\left(\rho_\Theta\Theta+\mathcal{E},\widetilde{A}(Z),\widetilde{\Omega}(Z)\right)d\Pr(\mathcal{E}).$$

This recursive system is initialized by some initial conditions Z_0.

3 The perturbational approach

We approximate equilibria using a "small-noise" perturbation approach which is a popular in both representative agent (e.g., Schmitt-Grohé and Uribe (2004) or Judd (1998)) and heterogeneous agent (e.g., Reiter (2009)) macroeconomic settings. Under this approach, one considers a sequence of economies parameterized by a scalar parameter $\sigma\ge 0$ that scales the exogenous aggregate shocks, that is,

$$\Theta_t=\rho_\Theta\Theta_{t-1}+\sigma\mathcal{E}_t. \tag{13}$$

The state-space representation must hold for all σ. To obtain approximations of the desired stochastic economy, corresponding to case $\sigma=1$, one uses various orders of Taylor expansions of the state-space representations (8)-(13) with respect to σ evaluated at $\sigma=0$. The $\sigma=0$ case corresponds to the equilibrium of the deterministic economy, that is typically easy to find. We refer to this deterministic economy as the zeroth-order approximation. It will be helpful to describe it formally.

3.1 The zeroth-order economy

Let $\widetilde{X}(Z;\sigma)$, $\widetilde{x}(a,\theta,Z;\sigma)$, $\widetilde{\Omega}(Z;\sigma)$ be policy functions in the perturbed economy, i.e., the economy described by equations (8) – (13). The $\sigma=0$ economy plays an important role in our approximations. We use bars to denote policy functions in the zeroth-order economy, so that $\overline{X}(Z)$ stands for $\widetilde{X}(Z;0)$. Let $\overline{\Theta}(Z)=\rho_{\Theta}\Theta$ and $\overline{Z}(Z):=[\overline{\Theta}(Z),\overline{A}(Z),\overline{\Omega}(Z)]^{\mathrm{T}}$. We use $Z^*=[0,A^*,\Omega^*]^{\mathrm{T}}$ to denote its steady state with Ω^* representing the invariant distribution of the HA model without aggregate shocks. When $Z=Z^*$, we drop explicit reference to the aggregate state, so that $\overline{X}$ will be understood to mean $\widetilde{X}(Z^*;0)$ or, equivalently, $\overline{X}(Z^*)$. The same convention will apply to all other policy functions. Let $\overline{\Lambda}(a',\theta',a,\theta)$ be the transition kernel from (a,θ) to (a',θ') under policy function $\overline{a}(a,\theta)$ and the stochastic process (8) for θ.

Throughout this section, we maintain the following assumption.

Assumption 1. *Let* $\overline{Z}_t:=\underbrace{\overline{Z}(\overline{Z}(....\overline{Z}}_{t\ times}(Z_0)))$.

(a) $\widetilde{X}(Z;\sigma)$ *is sufficiently differentiable*[7] *at* $(Z,\sigma)=(Z^*,0)$*;* $\widetilde{x}(a,\theta,Z;\sigma)$ *is continuous and piecewise sufficiently differentiable at* $(Z,\sigma)=(Z^*,0)$ *for all* (a,θ)*. The points of non-differentiability* $(a,\widetilde{\theta}_j^{\vee}(a,Z;\sigma),Z,\sigma)$ *are described by a finite number of functions* $\{\widetilde{\theta}_j^{\vee}(a,Z;\sigma)\}_j$*, where* $\{\widetilde{\theta}_j^{\vee}(a,Z;\sigma)\}_j$ *are sufficiently differentiable and all* $\overline{\theta}_j^{\vee}(\cdot)$ *are invertable given* (Z,σ) *with* $\overline{a}_j^{\vee}$ *denoting the inverse of* $\overline{\theta}_j^{\vee}$*;*

(b) $\lim_{t\to\infty}\overline{Z}_t(Z_0)=Z^*$ *for all* Z_0 *in a neighborhood of* Z^**;*

(c) The marginal of Ω^* *with respect to* a *has a finite number of mass-points* $\{a_n^*\}_n$*, i.e.,* Ω^* *has density* $\mathring{\omega}^*(a,\theta)+\sum_n\xi_n^*(\theta)\delta(a-a_n^*)$*, where* $\mathring{\omega}^*(a,\theta)$ *and* $\xi_n^*(\theta)$ *are continuous and* δ *is a Dirac delta function.*

Parts (a) and (b) are generalizations of differentiability and stability conditions that are required to apply perturbational methods in representative agent settings (Blanchard and Kahn (1980), Schmitt-Grohé and Uribe (2004)).[8] Unlike that approach, which usually requires all policy functions to be differentiable at the steady state, we permit individual policy functions to have kinks. The set of points at which kinks occur is described by $\{\widetilde{\theta}_j^{\vee}(a,Z;\sigma)\}_j$. Thus, the j^{th} kink of $\widetilde{x}(a,\cdot,Z;\sigma)$ occurs at $\theta=\widetilde{\theta}_j^{\vee}(a,Z;\sigma)$. Condition (c) permits the marginal of the invariant distribution Ω^* to have a finite number of mass points but requires it to have a smooth density otherwise. A direct implication of conditions (a) and (c) is that the set of points for which $\widetilde{x}$ is not-differentiable at $(Z^*,0)$ is of Ω^*-measure zero. This also implies that the integral $\int\widetilde{x}d\Omega$ is differentiable at $(Z^*,0)$.

[7] By "sufficiently differentiable" we mean that policy functions are differentiable at least n times when we consider n^{th} order of approximation. In multi-dimensional cases, all derivatives are understood to be Fréchet derivatives.

[8] In RA settings its possible to find conditions on the primitives under which policy functions are sufficiently differentiable (see, e.g., Jin and Judd (2002), who show smoothness under the assumption that equilibrium conditions are described by analytic functions). This is much harder to do in HA models and we simply assume these conditions directly.

As an illustration, consider our example of the Krusell-Smith economy. In that economy, policy function $\overline{k}(k,\theta)$ is continuous, strictly increasing for $\theta \geq \overline{\theta}^{\vee}(k)$ and equal to zero for $\theta \leq \overline{\theta}^{\vee}(k)$, where $\overline{\theta}^{\vee}(k)$ is the level of θ at which the borrowing constraint starts to bind. Since the distribution of idiosyncratic shocks μ is a density, the marginal of Ω^* can have at most one mass point, at $k=0$. Thus, its density takes the form $\mathring{\omega}^*(k,\theta)+\xi^*(\theta)\delta(k)$, where $\xi^*(\theta)$ is the density of agents with productivity θ at $k=0$.

For the rest of the paper, we assume that policy function $\overline{x}(\cdot,\theta)$ has at most one kink, at some $\overline{a}^{\vee}(\theta)$. This done merely to simplify the exposition. The extension of our formulas to accommodate finite number of kinks is immediate.

3.2 Derivatives and generalized functions

Since our perturbational approach involves approximations of complicated objects, it will be useful to develop notation for derivatives systematically and describe their properties. We use $\mathsf{F}_a(a,\theta)$, $\mathsf{F}_x(a,\theta)$, $\mathsf{F}_{x^e}(a,\theta)$, $\mathsf{F}_Y(a,\theta)$ be derivatives of F with respect to a, x, $\mathbb{E}_{\varepsilon,\mathcal{E}}x$ and Y evaluated at individual state (a,θ) and $(Z,\sigma)=(Z^*,0)$. Similarly G_x and G_Y denotes derivatives of G with respect to each of its arguments, evaluated at $(Z,\sigma)=(Z^*,0)$. Their higher-order analogues are denoted by $\mathsf{F}_{xx}(a,\theta)$, G_{xY}, etc.

We use $\overline{x}_a(a,\theta)$ and $\overline{x}_{aa}(a,\theta)$ to represent derivatives of policy functions with respect to a. Since policy functions may have kinks, these derivatives are not defined in the classical sense at those kinks. To handle this, we treat all derivatives of individual policy functions as generalized or distributional derivatives and represent them as generalized functions.[9] Since generalized functions are infinitely differentiable (in the distributional derivative sense), they allows us present a uniform treatment of approximations at any order. When we want to distinguish between generalized and classical functions, we use symbol $^\circ$ to denote the latter. To see the relationship between the two, let $\overline{x}^{\Delta}(\theta)=\lim_{a\downarrow a^{\vee}}\overline{x}(a,\theta)-\lim_{a\uparrow a^{\vee}}\overline{x}(a,\theta)$ and $\overline{x}_a^{\Delta}(\theta)$ be defined analogously $\overline{x}_a(\cdot,\theta)$. Obviously, $\overline{x}^{\Delta}(\theta)=0$ by continuity of $\overline{x}(\cdot,\theta)$ but $\overline{x}_a^{\Delta}(\theta)\neq 0$ in the presence of kinks. The relationship between generalized and classical derivatives is given by

$$\overline{x}_a(a,\theta)=\mathring{\overline{x}}_a(a,\theta)+\underbrace{\delta\left(a-\overline{a}^{\vee}(\theta)\right)\overline{x}^{\Delta}(\theta)}_{=0},$$
$$\overline{x}_{aa}(a,\theta)=\mathring{\overline{x}}_{aa}(a,\theta)+\delta\left(a-\overline{a}^{\vee}(\theta)\right)\overline{x}_a^{\Delta}(\theta).$$

These relationships imply that integrals of $\overline{x}_a$ and $\mathring{\overline{x}}_a$ always agree, but integrals of $\overline{x}_{aa}$ differ from $\mathring{\overline{x}}_{aa}$

[9]A generalized function is a linear functional over some space of functions. For instance, δ is a generalized function defined by the operation $\delta[\phi]=\int\phi(x)\delta(x)dx=\phi(0)$ for some function ϕ. There is a large mathematical literature on generalized functions (also referred to as distributions) and distributional derivatives, see Kanwal (1998) for an introduction to this subject.

by terms involving jumps at the kinks, e.g.,

$$\int \overline{x}_{aa} d\Omega^* = \int \mathring{\overline{x}}_{aa} d\Omega^* + \int \overline{x}_a^{\Delta}(\theta)\, \omega^*\left(\overline{a}^{\vee}(\theta), \theta\right) d\theta,$$

where $\omega^*(a,\theta) := \mathring{\omega}^*(a,\theta) + \sum_n \xi_n^*(\theta)\delta(a - a_n^*)$. Note that ω^* is a generalized function as well and notations such as $\int \overline{x} d\Omega^*$ and $\int \overline{x}\omega^* d\theta da$ are equivalent under this convention.

We use notation such as $\overline{X}_Z$, $\overline{x}_Z(a,\theta)$ and $\overline{\Omega}_Z$ to denote (the Fréchet) derivatives of zeroth-order policy functions with respect to Z, evaluated at $Z = Z^*$.[10] The derivative of $\overline{Z}$ takes the form $\overline{Z}_Z = \left[\overline{\Theta}_Z, \overline{A}_Z, \overline{\Omega}_Z\right]^{\mathrm{T}}$. Derivatives $\overline{X}_Z$ are complicated linear operators. We use $\overline{X}_Z \cdot \hat{Z}$ to denote the value of derivative $\overline{X}_Z$ evaluated in direction $\hat{Z}$. Similar notation applies to higher orders. For example, $\overline{X}_{ZZ}$ denotes the second-order Fréchet derivative (a bilinear map) and $\overline{X}_{ZZ} \cdot \left(\hat{Z}', \hat{Z}''\right)$ denotes its value in directions $\hat{Z}', \hat{Z}''$.

It is useful to keep in mind the dimensionality of different objects. The dimensionality of $\overline{X}_Z$ is equal to the dimensionality of X times dimensionality of Z, and the dimensionality of $\overline{X}_{ZZ}$ is the dimensionality of X times the square of that of Z. In HA economies Z is a large (theoretically, infinite) dimensional object which makes computing and storing such derivatives very costly. On the other hand, dimensionalities of $\overline{X}_Z \cdot \hat{Z}$, $\overline{X}_{ZZ} \cdot \left(\hat{Z}', \hat{Z}''\right)$ and their higher-order generalizations are equal to dimensionality of $\overline{X}$, i.e., to the number of aggregate variables in the model. These are small-dimensional objects in most HA models that are easy to compute and store. Our perturbational approach builds on the idea that one can represent equilibrium approximations using the Fréchet derivatives of $\overline{X}$ evaluated in an appropriate set of directions.

Derivatives of individual policy functions such as $\overline{x}_Z \cdot \hat{Z}$ or $\overline{x}_{ZZ} \cdot \left(\hat{Z}', \hat{Z}''\right)$ and cross-partials $\overline{x}_{aZ} \cdot \hat{Z}$ will be understood to be represented by generalized functions. We show in the appendix that they satisfy $\overline{x}_Z(a,\theta) \cdot \hat{Z} = \mathring{\overline{x}}_Z(a,\theta) \cdot \hat{Z}$ and

$$\overline{x}_{aZ}(a,\theta) \cdot \hat{Z} = \mathring{\overline{x}}_{aZ}(a,\theta) \cdot \hat{Z} + \delta(a - \overline{a}^{\vee}(\theta))\overline{x}_Z^{\Delta}(\theta) \cdot \hat{Z},$$

$$\overline{x}_{ZZ}(a,\theta) \cdot \left(\hat{Z}', \hat{Z}''\right) = \mathring{\overline{x}}_{ZZ}(a,\theta) \cdot \left(\hat{Z}', \hat{Z}''\right) + \delta(a - \overline{a}^{\vee}(\theta))\overline{x}_a^{\Delta}(\theta)\left(\overline{a}_Z^{\vee}(\theta) \cdot \hat{Z}'\right)\left(\overline{a}_Z^{\vee}(\theta) \cdot \hat{Z}''\right).$$

Here $\overline{a}_Z^{\vee}(\theta) \cdot \hat{Z}$ represents how the kink moves when the aggregate state is changed in direction $\hat{Z}$. Finally, we use notation such as $\overline{X}_\sigma$, $\overline{X}_{\sigma\sigma}$ to denote derivatives policy functions such as $\widetilde{X}(Z;\sigma)$ with respect to σ, evaluated at $(Z,\sigma) = (Z^*, 0)$. We refer to $_{\sigma\sigma}$ derivatives as precautionary motives.

[10] The Fréchet derivative $\overline{X}_Z$ is a linear operator that satisfies $\lim_{\|\hat{Z}\| \to 0} \left\|\overline{X}\left(Z^* + \hat{Z}\right) - \overline{X}(Z^*) - \overline{X}_Z \cdot \hat{Z}\right\| / \left\|\hat{Z}\right\|$ for all $\hat{Z}$. Its value coincides with that of the directional (Gateaux) derivative, $\overline{X}_Z \cdot \hat{Z} = \lim_{\alpha \to 0}\left(\overline{X}\left(Z^* + \alpha\hat{Z}\right) - \overline{X}(Z^*)\right)/\alpha$. Intuitively, $\overline{X}_Z$ is just the gradient of $\overline{X}$, i.e., the vector of partial derivatives of $\overline{X}$ with respect to each of the dimensions of Z. We refer to $\overline{X}_Z \cdot \hat{Z}$ as the "value of that derivative in direction $\hat{Z}$". This terminology is slightly different from the one used by Luenberger (1997) and is meant to highlight the economic meaning of these objects. Luenberger (Chapter 7) would refer to $\overline{X}_Z \cdot \hat{Z}$ as the "Fréchet differential of $\overline{X}$ (at Z^*) with increment $\hat{Z}$".

3.3 Numerical solution of the zeroth-order economy

Existing off-the-shelf numerical techniques can be used to find the zeroth-order economy and we treat the equilibrium of that economy as known for the purposes of our approximation. Moreover, many standard algorithms already return the output that is necessary to construct many of the derivatives that we discussed above. We illustrate how this can be done when policy functions are approximated using a finite set of splines as basis functions.[11] As we discuss below, using splines is convenient not only for interpolating policy functions but also for evaluating their derivatives and conditional expectations.

Let N_X, N_Y and N_x be dimensions of vectors $\overline{X}$, $\overline{Y}$ and $\overline{x}\,(a,\theta)$.[12] Discretize the space of (a,θ) into N_Ω points, where N_Ω is large enough to be practical.[13] Let $(a,\theta)_{[i]}$ be the value of the i^{th} point on the grid. We use arrows to denote discretized analogues of theoretical objects, for example $\overrightarrow{x}$ denotes the values of function $\overline{x}$ on the grid on N_Ω points. We keep track of kinks in policy functions by a set of grid points $\aleph$, where $i \in \aleph$ denotes the point $(\theta,a)_{[i]}$ just below the kink while $(\theta,a)_{[i+1]}$ is just above the kink.

Standard algorithms return a transition probability matrix, an invariant distribution, and values of policy functions on a grid. The transition probability matrix $\overrightarrow{\Lambda}$ is a sparse $N_\Omega \times N_\Omega$ matrix with $\overrightarrow{\Lambda}[i',i]$ being the probability that an agent who has state $(a,\theta)_{[i]}$ in the current period end up in state $(a,\theta)_{[i']}$ in the next period. The invariant distribution $\overrightarrow{d\Omega}^*$ is the N_Ω dimensional vector that satisfies $\overrightarrow{d\Omega}^* = \overrightarrow{\Lambda}\overrightarrow{d\Omega}^*$. Finally, individual policy functions are stored as a vector of coefficients on a set of common basis functions. Let $\left\{\phi^j(\cdot,\cdot)\right\}_{j=1}^{N_{sp}}$ be a collection of basis functions, where each ϕ^j is differentiable and maps from (a,θ) into $\mathbb{R}$ and $N_{sp} << N_\Omega$. Individual policy functions are stored in terms of $N_{sp} \times N_\Omega$ matrix Φ with elements $\Phi[j,i] = \phi^j\left((a,\theta)_{[i]}\right)$ and $N_x \times N_{sp}$ matrix $\overline{x}^{\#}$ so $\overline{x}(a,\theta) \approx \sum_j \overline{x}_j^{\#}\phi^j(a,\theta)$. Using these two objects, the $N_x \times N_\Omega$ matrix of values of policy functions $\overrightarrow{x}$ can be recovered by

$$\overrightarrow{x} = \overline{x}^{\#}\Phi. \tag{14}$$

Similarly, it is easy to compute derivatives of policy functions with respect to a. Let ϕ_a^j be the derivative of ϕ^j with respect to a and construct $N_{sp} \times N_\Omega$ matrix Φ_a with elements $\Phi_a[j,i] = \phi_a^j((a,\theta)_{[i]})$. The computational analogues of $\overline{x}_a$ is recovered as in (14), i.e., $\overrightarrow{x}_a = \overline{x}^{\#}\Phi_a$.[14] Since splines approximate kinked functions with differentiable ones, the theoretical distinction between generalized

[11]In our application we use quadratic splines, though the same algorithm works with linear splines as those derivatives are defined almost everywhere.

[12]Our definition of Y includes vectors X, Θ and $\mathbb{E}_{\mathcal{E}}X$. We did that to show the general framework that encompasses many HA environments. Of course, there is no need to include into $\overline{Y}$ those variables that do not affect equilibrium conditions. For example, in the Krusell and Smith economy, we would set $Y = [K,\Theta,W,R]^{\mathrm{T}}$ and so $N_Y = 4$ in that economy, while $N_X = 3$ and $n_x = 4$.

[13]For Krusell and Smith type applications, one generally uses 1000-5000 points per realization of the idiosyncratic shocks. See Reiter (2009) or discussions in the handbook chapter of Algan et al. (2014) for this convention.

[14]In addition it is possible to compute $\overline{x}_a^{\#}$ directly by differentiating the F function with respect to a. The computational analogue of $\overline{x}_{aa}$ can then be recovered as $\overline{x}_a^{\#}\Phi_a$.

and classical derivatives disappears and integrals such as $\int \overline{x}_{aa} d\Omega^*$ can be computed as a matrix multiplication $\overrightarrow{x}_{aa}\overrightarrow{d\Omega}$. The same approach extends to compute conditional expectations, such as $\mathbb{E}_\varepsilon [\overline{x}|a,\theta] = \int \overline{x}(\overline{a}(a,\theta), \rho_\theta\theta + \varepsilon) d\mu(\varepsilon)$. Discretizing the distribution ε into $\varepsilon_1, ..., \varepsilon_K$ with probabilities $\mu_1, ..., \mu_K$, we construct matrix Φ^e with elements $\Phi^e[j,i] = \sum_{k=1}^K \phi^j \left(\overline{a}(a,\theta)_{[i]}, \rho\theta_{[i]} + \varepsilon_k\right) \mu_k$ on the spline grid $i = 1, ..., N_\Omega$ and then recover $\mathbb{E}[\overline{x}|a,\theta]$ as $\overline{x}^\# \Phi^e$. Substantial speed ups can be obtained by pre-computing all the basis matrices and storing them efficiently as sparse matrices. Pre-computing basis matrices allows us to reduce all necessary calculations to matrix algebra without any further nonlinear function calls.

It is easy to recover many of the derivatives discussed in Section 3.2 directly from the zeroth-order objects. Automatic differentiation of (9) and (10) yields various derivatives of mappings F and G. Evaluating those derivatives using the steady-state policy functions $\overline{x}$ gives $\mathsf{F}_a(a,\theta)$, G_{xY}, etc. We now turn to describing our approach of recovering the values of derivatives such as $\overline{X}_Z \cdot \hat{Z}$ or $\overline{X}_{ZZ} \cdot \left(\hat{Z}', \hat{Z}''\right)$ that are necessary to characterize equilibrium responses to aggregate shocks to various orders of approximations.

4 Approximation of equilibrium responses to aggregate shocks

In this section, we describe our approximation of equilibrium responses to aggregate shocks. Up until this point, our approach closely mirrored traditional applications of perturbational techniques (e.g., Schmitt-Grohé and Uribe (2004) or Reiter (2009)) in that we found the zeroth-order economy and some of derivatives of policy functions. The usual next step is to differentiate the equilibrium conditions around the deterministic steady-state to find $\left(\overline{X}_Z, \overline{Z}_Z\right)$ from a stable solution of a quadratic matrix equation. However, the computational cost of doing that scales with the dimension of Z and exponentially so for higher-orders.

We proceed differently. We start from the observation that in order to construct the equilibrium approximation, one does not need to know $\overline{X}_Z$, but its value in a certain sequences of directions, $\{\overline{X}_Z \cdot \hat{Z}_t\}_t$, where sequence of directions $\{\hat{Z}_t\}_t$ can be characterized analytically. Unlike $\left(\overline{X}_Z, \overline{Z}_Z\right)$, which are large dimensional objects, $\overline{X}_Z \cdot \hat{Z}_t$ is a small vector. We derive the expressions for coefficients of a linear system of equations that determines $\{\overline{X}_Z \cdot \hat{Z}_t\}_t$. These formulas are easy to implement numerically as they require simple operations with objects described in Section 3.3. This allows us to compute the first-order approximation much faster and more efficiently relative to traditional applications of perturbational techniques.

Our approach has several additional advantages. The notion of representing equilibrium approximations using directional derivatives extends to second- and higher-orders. In fact, we show the formulas characterizing second-order approximations closely mirror those of the first order, and their implementation recycles many of the same objects that were constructed to compute the first-order approximation.

Finally, we recover equilibrium approximations of any order by solving only linear equations. It obviates the need to use additional refinements for standard perturbational techniques, such as finding stable roots of quadratic matrix equations to characterize the first order, or to introduce "pruning" to find the second-order approximations, see Kim et al. (2008). All this further simplifies our approach.

4.1 First-order approximations

Our starting point is that equilibrium is the stochastic process $X_t(\mathcal{E}^t)$ that can be constructed from policy functions as follows

$$X_t(\mathcal{E}^t) = \tilde{X}\left(Z_t\left(\mathcal{E}^t;\sigma\right);\sigma\right)\Big|_{\sigma=1}, \tag{15}$$

where $Z_t(\mathcal{E}^t,\sigma)$ is defined recursively as $Z_0 = [\sigma\mathcal{E}_0, \overline{A}, \Omega^*]$ and

$$Z_t\left(\mathcal{E}^t;\sigma\right) = \left[\rho_\Theta\Theta_{t-1}\left(\mathcal{E}^{t-1};\sigma\right) + \sigma\mathcal{E}_t, \tilde{A}\left(Z_{t-1}\left(\mathcal{E}^{t-1};\sigma\right);\sigma\right), \tilde{\Omega}\left(Z_{t-1}\left(\mathcal{E}^{t-1};\sigma\right);\sigma\right)\right]. \tag{16}$$

We can take the first order Taylor expansion with respect to σ of the right hand side of (15) to obtain the first-order approximation. Using well-known results that precautionary motives are absent to the first-order (see Schmitt-Grohé and Uribe (2004)), $\overline{X}_\sigma = 0$, we obtain the following lemma.

Lemma 1^{FO}. *To the first-order approximation, X_t satisfies*

$$X_t\left(\mathcal{E}^t\right) = \overline{X} + \sum_{s=0}^{t}\overline{X}_{Z,t-s}\mathcal{E}_s + O\left(\|\mathcal{E}\|^2\right),$$

where $\overline{X}_{Z,t} := \overline{X}_Z \cdot \hat{Z}_t$ and the sequence of directions $\left\{\hat{Z}_t\right\}_t$ satisfies recursion $\hat{Z}_0 = [1, 0, \mathbf{0}]^{\mathrm{T}}$ and $\hat{Z}_t := \overline{Z}_Z \cdot \hat{Z}_{t-1}$.

We now explain the economic intuition for the characterization in Lemma 1^{FO}. The sequence of directions $\left\{\hat{Z}_t\right\}_t$ traces out the response of the aggregate state to a one-time unit shock to Θ_0. Since A_{-1} and Ω_0 are pre-determined in period 0, the aggregate state changes by $\hat{Z}_0 = [1, 0, \mathbf{0}]^{\mathrm{T}}$. In the next period, policy functions respond to this change in state by $\hat{Z}_1 = \left[\rho_\Theta, \overline{A}_Z \cdot \hat{Z}_0, \overline{\Omega}_Z \cdot \hat{Z}_0\right]^{\mathrm{T}}$, which combines both the decay ρ_Θ of the exogenous component of the aggregate state and endogenous policy responses. Iterating forward, we obtain $\left\{\hat{Z}_t\right\}_t$. The sequence of values $\left\{\overline{X}_{Z,t}\right\}_t$ capture the response of policy functions to this change in state. Sequence $\{\overline{X}_{Z,t}\}_t$ represents what is often referred to colloquially in the literature as an impulse response to a "MIT shock". Analogously, $\overline{x}_{Z,t}\left(a,\theta\right) := \overline{x}_Z\left(a,\theta\right)\cdot\hat{Z}_t$ captures the effect of the MIT shock in period 0 on policy function of individual who has states (a,θ) at date t.

Lemma 1^{FO} re-formulates in the state-space representation the insight in Boppart et al. (2018) that one can construct first-order approximations to stochastic economies using impulse responses to MIT shocks. The main take away for our purposes is that finding the first-order equilibrium approximation is equivalent to finding values of the Fréchet derivative $\overline{X}_Z$ in a sequence of directions $\left\{\hat{Z}_t\right\}_t$. In order to characterize $\left\{\overline{X}_{Z,t}\right\}$, we start by taking a Fréchet derivative of (10) and evaluate it in direction $\hat{Z}_t$:

$$\mathsf{G}_x \left(\int \overline{x} d\Omega\right)_{Z,t} + \mathsf{G}_Y \overline{Y}_{Z,t} = 0, \tag{17}$$

where

$$\overline{Y}_{Z,t} = \left[\rho_\Theta^t, \mathsf{P}\overline{X}_{Z,t-1}, \overline{X}_{Z,t}, \overline{X}_{Z,t+1}\right]^{\mathrm{T}}. \tag{18}$$

Recall that G_x and G_Y are already known from the zeroth-order economy. $\overline{Y}_{Z,t}$ consists of the first-order equilibrium responses $\left\{\overline{X}_{Z,t}\right\}$ that we want to find as well as the rate of decay for exogenous aggregate variables.

To proceed further, we simplify $\left(\int \overline{x} d\Omega\right)_{Z,t}$. Using the definition of directional derivatives and continuity of policy functions at the kinks, it is easy to show that

$$\left(\int \overline{x} d\Omega\right)_{Z,t} = \int \overline{x}_{Z,t} d\Omega^* + \int \overline{x} d\hat{\Omega}_t. \tag{19}$$

The first-order change in the aggregation $\int \overline{x} d\Omega$ consists of two components: the effect of the shock on individual policy functions, $\int \overline{x}_{Z,t} d\Omega^*$, and the effect of the shock on the aggregate distribution, $\int \overline{x} d\hat{\Omega}_t$. In order to characterize these two components, we use two intermediate results. First, we show that there is a tight relationship between responses of individual and aggregate endogenous variables to aggregate shocks. We obtain it by applying the implicit function theorem to mapping F defined in equation (9) and evaluating those expressions at the deterministic steady-state.

Lemma 2^{FO}. *For any t,*

$$\overline{x}_{Z,t}(a,\theta) = \sum_{s=0}^{\infty} \mathsf{x}_s(a,\theta) \overline{Y}_{Z,t+s}, \tag{20}$$

where matrices $\mathsf{x}_s(a,\theta)$ are given by

$$\mathsf{x}_0(a,\theta) = -\left(\mathsf{F}_x(a,\theta) + \mathsf{F}_{x^e}(a,\theta)\mathbb{E}_\varepsilon\left[\overline{x}_a|a,\theta\right]\mathsf{p}\right)^{-1}\mathsf{F}_Y(a,\theta), \tag{21}$$

$$\mathsf{x}_{s+1}(a,\theta) = -\left(\mathsf{F}_x(a,\theta) + \mathsf{F}_{x^e}(a,\theta)\mathbb{E}_\varepsilon\left[\overline{x}_a|a,\theta\right]\mathsf{p}\right)^{-1}\mathsf{F}_{x^e}(a,\theta)\mathbb{E}_\varepsilon\left[\mathsf{x}_s|a,\theta\right] \tag{22}$$

away from the kinks and $\mathsf{x}_s(\overline{a}^\vee(\theta),\theta) = 0$ at the kinks.

Equation (20) shows that change in individual policy functions $\overline{x}_{Z,t}$ is equal to the future changes in aggregates $\left\{\overline{Y}_{Z,t+s}\right\}_s$ weighted with matrices $\{\mathsf{x}_s\}_s$. Matrix x_s has a natural economic interpretation. It captures how much individual change their policy functions today if they expect the aggregates to change s periods in the future, $\partial x_t/\partial Y_{t+s}$. The most important part of Lemma 2^{FO} is that it provides explicit formulas for $\{\mathsf{x}_s\}_s$. The right hand side of (21) is known from Section 3.3 and so x_0 can be constructed from the zeroth-order economy using linear algebra operations. This allows the construction of $\{\mathsf{x}_s\}_{s>0}$ sequentially using (22).

The second intermediate result describes the Law of Motion for $\hat{\Omega}_t$, that helps simplifying the second term on the right hand side of (19). It is helpful to define three operators, $\mathcal{M}$, $\mathcal{L}^{(a)}$, and $\mathcal{I}^{(a)}$ on a space of generalized functions. For a generalized function y,

$$(\mathcal{M}\cdot y)\langle a',\theta'\rangle := \int \overline{\Lambda}(a',\theta',a,\theta)\omega^*(a,\theta)\,y(a,\theta)\,dad\theta,$$
$$\left(\mathcal{L}^{(a)}\cdot y\right)\langle a',\theta'\rangle := \int \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_a(a,\theta)\,y(a,\theta)\,dad\theta,$$
$$\mathcal{I}^{(a)}\cdot y := \int \overline{x}_a(\theta,a)y(\theta,a)\,dad\theta.$$

Operators $\mathcal{M}$ and $\mathcal{L}^{(a)}$ take a function $y(a,\theta)$, multiply it with ω^* and $\overline{a}_a$, respectively, and integrate it over $\overline{\Lambda}(a',\theta',\cdot,\cdot)$ for various values of (a',θ'). Operator $\mathcal{I}^{(a)}$ integrates y weighted with $\overline{x}_a$. To explain the economic forces captured by these operators, we start with the following lemma, which we obtain by explicitly taking derivatives of the LoM of $\overline{\Omega}$ defined in equation (11).

Lemma 3^{FO}. *For any t, $\frac{d}{d\theta}\hat{\Omega}_t$ satisfies a recursion $\frac{d}{d\theta}\hat{\Omega}_0=\mathbf{0}$ and*

$$\frac{d}{d\theta}\hat{\Omega}_{t+1}=\mathcal{L}\cdot\frac{d}{d\theta}\hat{\Omega}_t-\mathcal{M}\cdot\overline{a}_{Z,t}. \tag{23}$$

Equation (23) describes how the aggregate distribution Ω_t is affected by aggregate shocks. On impact of the shock in period 0, the distribution is pre-determined and thus $\frac{d}{d\theta}\hat{\Omega}_0=\mathbf{0}$. Individuals change their choices in period 0. In particular, individual (a,θ) changes her savings behavior by $\overline{a}_{Z,0}(a,\theta)$. Operator $\mathcal{M}$ aggregates these individual-level changes by weighting them with the invariant density ω^* and returns the change in the distribution in period 1, $\frac{d}{d\theta}\hat{\Omega}_1=-\mathcal{M}\cdot\overline{a}_{Z,0}$. Thus, $\mathcal{M}$ captures the first-order effect of changes in individual policy functions on the aggregate distribution next period. For all $t>0$, the distribution Ω_t is affected by two forces. One is mechanical: if the distribution Ω_{t-1} changed in the previous period, Ω_t would also change even if individual policy functions did not change. This mechanical effect is captured by the operator $\mathcal{L}^{(a)}$. The aggregate distribution in period $t+1$ is also affected by the response of individuals in period t, and this behavioral effect is captured by $\mathcal{M}\cdot\overline{a}_{Z,t}$.

To make recursion (23) operational, note that $\int\overline{x}d\hat{\Omega}_t=-\mathcal{I}\cdot\frac{d}{d\theta}\hat{\Omega}_t$ using integration by parts. This leads to the following corollary that characterizes the derivative of the aggregation equation (19).

Corollary 1^{FO}. *For any t,*

$$\left(\overline{\int xd\Omega}\right)_{Z,t}=\sum_{s=0}^{\infty}\mathsf{J}_{t,s}\overline{Y}_{Z,s}$$

where $\{\mathsf{J}_{t,s}\}_{t,s}$ is characterized by the following linear recursive system

$$\mathsf{J}_{t,s}=\int\mathsf{x}_{s-t}d\Omega^*+\mathcal{I}^{(a)}\cdot\mathsf{A}_{t,s}, \tag{24}$$

$$\mathsf{A}_{t,s} = \mathcal{L}^{(a)} \cdot \mathsf{A}_{t-1,s} + \mathcal{M} \cdot \mathsf{a}_{s-t-1}, \tag{25}$$

with convention that $\mathsf{a}_s = \mathsf{p}\mathsf{x}_s$ *and* $\mathsf{a}_t = 0$ *for* $t < 0$ *and* $\mathsf{A}_{0,s} = 0$ *for all* s.

This corollary shows that derivative $\left(\overline{\int x d\Omega}\right)_{Z,t}$ can be expressed purely in terms of $\left\{\overline{X}_{Z,s}\right\}_s$ weighted with matrices $\{\mathsf{J}_{t,s}\}_{t,s}$ that is described by a linear recursive system of equations (24) and (25) with all all objects known from the zeroth-order economy. As we explain in Section 4.1.1, numerically implementing (24) and (25) from the zeroth order solution is quick and efficient. In fact, these equations can be simplified to

$$\mathsf{J}_{t,s} = \mathsf{J}_{t-1,s-1} + \mathcal{I}^{(a)} \cdot \left(\mathcal{L}^{(a)}\right)^{t-1} \cdot \mathcal{M} \cdot \mathsf{a}_s, \tag{26}$$

with initial conditions $\mathsf{J}_{0,t} = \int \mathsf{x}_s d\Omega^*$, $\mathsf{J}_{t,0} = 0$.

Combine Corollary 1^{FO} and equation (17) to obtain the main result of this section:

Proposition 1^{FO}. $\left\{\overline{X}_{Z,t}\right\}_t$ *is the solution to a linear system*

$$\mathsf{G}_x \sum_{s=0}^{\infty} \mathsf{J}_{t,s} \overline{Y}_{Z,s} + \mathsf{G}_Y \overline{Y}_{Z,t} = 0, \tag{27}$$

with $\overline{Y}_{Z,t}$ *given by equation (18) and* $\mathsf{P}\overline{X}_{Z,-1} = 0$.

This proposition provides analytical expressions for all the coefficients in the linear system that determine $\{\overline{X}_{Z,t}\}_t$. Stability condition in Assumption 1 imply that $\overline{X}_{Z,t} \to 0$ as $t \to 0$, so this linear system can be solve by truncating at some sufficiently large T, imposing $\overline{X}_{Z,t} = 0$ for all $t > T$.

4.1.1 Numerical implementation of the first-order approximation

To solve the truncated linear system (27), we only need to construct $\{\mathsf{J}_{t,s}\}_{t,s}$ since G_x and G_Y were already constructed in Section 3.3. From (26), constructing $\{\mathsf{J}_{t,s}\}_{t,s}$ requires the sequence $\{\mathsf{x}_t\}_t$ and operators $\left\{\mathcal{I}^{(a)}, \mathcal{L}^{(a)}, \mathcal{M}\right\}$.

Sequence $\{\mathsf{x}_t\}_t$ is recovered from the recursion (21) and (22). All terms that appear in that formula have already been constructed in Section 3.3. Using the same basis functions as we used for the policy functions, the sequence of functions $\{\mathsf{x}_t\}_t$ is stored as a set of coefficients $\{\mathsf{x}_t^{\#}\}_{t=1}^{T}$ on the course grid of N_{sp} points, where each $\mathsf{x}_t^{\#}$ is an $N_x \times N_Y \times N_{sp}$ array so that $\overrightarrow{\mathsf{x}}_t = \mathsf{x}_t^{\#}\Phi$. The coefficient vectors $\left\{\mathsf{x}_t^{\#}\right\}$are given by[15]

$$\mathsf{x}^{\#}{}_0\Phi[:,j] = -\left(\mathsf{F}_x[j] + \mathsf{F}_{x^e}[j]\left(\mathsf{p}\overline{x}_a^{\#}\right)\Phi^e[:,j]\right)^{-1}\mathsf{F}_Y[j],$$
$$\mathsf{x}^{\#}{}_{s+1}\Phi[:,j] = -\left(\mathsf{F}_x[j] + \mathsf{F}_{x^e}[j]\left(\mathsf{p}\overline{x}_a^{\#}\right)\Phi^e[:,j]\right)^{-1}\mathsf{F}_{x^e}[j]\mathsf{x}^{\#}{}_{s+1}\Phi^e[:,j].$$

The numerical analogue of operators $\mathcal{I}^{(a)}$, $\mathcal{L}^{(a)}$, $\mathcal{M}$ are matrices $\overrightarrow{\mathcal{I}}^{(a)}$, $\overrightarrow{\mathcal{L}}^{(a)}$, and $\overrightarrow{\mathcal{M}}$ with elements $\overrightarrow{\mathcal{I}}^{(a)}[:,i] = \overrightarrow{x}_a[i]$, $\overrightarrow{\mathcal{L}}^{(a)}[i',i] = \overrightarrow{\Lambda}[i',i]\overrightarrow{d}_a[i]$ and $\overrightarrow{\mathcal{M}}[i',j] = \sum_i \overrightarrow{\Lambda}[i',i]\overrightarrow{d\Omega}^*[i]\Phi[i,j]$[16] that are trivial to

[15]Here $\overrightarrow{\mathsf{F}}_x[j]$ represents the matrices $\mathsf{F}_x(a,\theta)$ evaluated at the j^{th} element of the coarse grid points used to approximate splines. For a given $\mathsf{x}_t^{\#}$, $\mathsf{x}_t^{\#}\Phi[:,j]$ represents splines approximation evaluated at the j^{th} element of the coarse grid points.

[16]Note that matrix $\overrightarrow{\mathcal{M}}$ is constructed to operate on the vector of spline coefficients.

construct from the objects built in Section 3.3. $\overrightarrow{\mathcal{L}}^{(a)}$ and $\overrightarrow{\mathcal{M}}$ are large but sparse matrices, and so computing $\{\mathsf{J}_{t,s}\}_{t,s}$ using

$$\mathsf{J}_{t,s} = \mathsf{J}_{t-1,s-1} + \overrightarrow{\mathcal{I}}^{(a)} \left(\overrightarrow{\mathcal{L}}^{(a)}\right)^{t-1} \overrightarrow{\mathcal{M}} \mathsf{a}_s^{\#}$$

can be done very quickly, where $\mathsf{a}_s^{\#} = \mathsf{p}\mathsf{x}_s^{\#}$. Once $\{\mathsf{J}_{t,s}\}_{t,s}$ are constructed, the truncated system of equations (27) is solved for $\{\overline{X}_{Z,t}\}_{t=0}^{T}$ via a $TN_X \times TN_X$ matrix inversion.

Once $\{\overline{X}_{Z,t}\}_{t=0}^{T}$ are computed it is also straightforward to compute how both individual policies and the distribution respond to shocks. Using equation (20), $\overline{x}_{Z,t}^{\#}$ can be computed as $\sum_{s=0}^{T-t} \mathsf{x}_s^{\#} \overline{Y}_{Z,t+s}$. Similarly, the numerical analogue of $\left\{\frac{d}{d\theta}\hat{\Omega}_t\right\}_t$ is a sequence of N_Ω dimensional vectors $\left\{\overrightarrow{\frac{d}{d\theta}\Omega}\right\}_t$ that is constructed recursively by $\overrightarrow{\frac{d}{d\theta}\Omega_0} = \mathbf{0}$ and $\overrightarrow{\frac{d}{d\theta}\Omega_t} = \overrightarrow{\mathcal{L}}^{(a)} \overrightarrow{\frac{d}{d\theta}\Omega_{t-1}} - \overrightarrow{\mathcal{M}} \overline{a}_{Z,t}^{\#}$. Finally, once $\overline{x}_{Z,t}^{\#}$ is known, it is possible to construct $\overrightarrow{a}_{Z,t}^{\vee}[j]$ for all $j \in \aleph$ as it depends on $\overline{x}_{Z,t}^{\Delta}(\theta)$. We provide a formula for it in the appendix. Although $\left\{\overrightarrow{a}_{Z,t}^{\vee}\right\}$ are not necessary for the first-order responses (see the discussion in Section 3.2), they will appear in the expressions for higher-order approximations.

4.2 Second-order approximations

Our approach naturally extends to approximations beyond the first order. In this section, we show that by differentiating twice equations (9), (10) and (11) we obtain linear systems of equations characterizing the second-order equilibrium responses. Moreover, those systems use several of the operators that were introduced in the first-order approximation.

Before we start, it is helpful to illustrate the general structure of second-order approximations using a simple example. Suppose we have a compounded function $f(g(\xi))$, where ξ is a scalar and f and g are smooth, scalar-valued functions with their first two derivatives denoted by f_g, f_{gg} and $g_\xi, g_{\xi\xi}$. The second-order derivatives of f with respect to ξ takes the form

$$\frac{\partial^2}{\partial \xi^2} f = f_g g_{\xi\xi} + f_{gg} g_\xi g_\xi. \tag{28}$$

Thus, the second order effect is equal a sum of two terms: the first-order response of f to the second-order change of g, $f_g g_{\xi\xi}$, and the second-order response of f to the first-order change in g, $f_{gg} g_\xi g_\xi$. With this intuition in place, we proceed with the second-order analogue of Lemma 1^{FO}. We differentiate equation (15) twice with respect to σ and evaluate at $\sigma = 0$.

Lemma 1^{SO}. *To the second-order approximation, X_t satisfies*

$$X_t\left(\mathcal{E}^t\right) = ... + \frac{1}{2}\left(\sum_{s=0}^{t}\sum_{m=0}^{t} \overline{X}_{ZZ,t-s,t-m}\mathcal{E}_s\mathcal{E}_m + \overline{X}_{\sigma\sigma,t}\right) + O\left(\|\mathcal{E}\|^3\right), \tag{29}$$

where ... are the first-order terms and $\left\{\overline{X}_{ZZ,t,s}\right\}_{t,s}$ and $\left\{\overline{X}_{\sigma\sigma,t}\right\}_t$ defined by

$$\overline{X}_{ZZ,t,s} := \overline{X}_Z \cdot \hat{Z}_{t,s} + \overline{X}_{ZZ} \cdot \left(\hat{Z}_t, \hat{Z}_s\right) \text{ for } \hat{Z}_{t,s} = \overline{Z}_Z \cdot \hat{Z}_{t-1,s-1} + \overline{Z}_{ZZ} \cdot \left(\hat{Z}_{t-1}, \hat{Z}_{s-1}\right), \tag{30}$$

$$\overline{X}_{\sigma\sigma,t} := \overline{X}_Z \cdot \hat{Z}_{\sigma\sigma,t} + \overline{X}_{\sigma\sigma} \text{ for } \hat{Z}_{\sigma\sigma,t} = \overline{Z}_Z \cdot \hat{Z}_{\sigma\sigma,t-1} + \left[0, \mathrm{P}\overline{X}_{\sigma\sigma}, \overline{\Omega}_{\sigma\sigma}\right]^{\mathrm{T}}. \tag{31}$$

with $\hat{Z}_{0,s} = \hat{Z}_{t,0} = \hat{Z}_{\sigma\sigma,0} = \mathbf{0}$.

This lemma shows that the second-order equilibrium approximation involves two types of terms: those like $\overline{X}_{ZZ,t,s}$ that capture the interaction effects on current period endogenous variables from shocks that occurred t and s period ago (it includes nonlinear responses when $t = s$) and precautionary motive terms like $\overline{X}_{\sigma\sigma,t}$. Inspection of equation (30) reveals that both direction $\hat{Z}_{t,s}$ and $\overline{X}_{ZZ,t,s}$ has the same mathematical structure as our example in equation (28). Equation for precautionary motive terms (31) is much simpler: since precautionary motive is zero to the first order, all the interaction terms (the analogue of $f_{gg}g_{\xi}g_{\xi}$ in equation (28)) disappear. Given this, we start with the description of the precautionary terms.

4.2.1 Characterization of precautionary terms

We proceed in the same way as in Section 4.1 by differentiating equation (10). We have already differentiated it with respect to Z in the previous section. Now we evaluate that derivative in direction $\hat{Z}_{\sigma\sigma,t}$ and add to it the second derivative G with respect to σ. This allows us to express $\overline{X}_{\sigma\sigma,t}$ using definition (31) as

$$\mathsf{G}_x \left(\int \overline{x} d\Omega\right)_{\sigma\sigma,t} + \mathsf{G}_Y \overline{Y}_{\sigma\sigma,t} = 0, \tag{32}$$

where

$$\overline{Y}_{\sigma\sigma,t} = \left[0, \mathsf{P}\overline{X}_{\sigma\sigma,t-1}, \overline{X}_{\sigma\sigma,t}, \overline{X}_{\sigma\sigma,t+1} + \overline{X}_{ZZ,0,0} var\left(\mathcal{E}\right)\right]^{\mathrm{T}}. \tag{33}$$

These equations have a similar structure as their first-order analogue, equations (17) and (18), with two differences. First, the terms corresponding to exogenous shocks are zero, since exogenous shocks are linear in σ (see equation (13)) and thus their second-order derivatives are zero. Second, precautionary motives also depend on $\overline{X}_{ZZ,0,0} var\left(\mathcal{E}\right)$. This has natural economic interpretation. Precautionary motives depend not on realization of the shock but uncertainty about them, captured by $var\left(\mathcal{E}\right)$, adjusted by the non-linearity in policy functions, captured by $\overline{X}_{ZZ,0,0}$. The same insight carries over to the two key intermediate results – Lemmas 2 and 3 – that were central in our characterization of the first-order approximations.

Lemma 2a^{SO}. *For any* t, k

$$\overline{x}_{\sigma\sigma,t}\left(a,\theta\right) = \sum_{s=0}^{\infty} \mathsf{x}_s\left(a,\theta\right) \overline{Y}_{\sigma\sigma,t+s} + \mathsf{x}_{\sigma\sigma}\left(a,\theta\right), \tag{34}$$

where $\mathsf{x}_{\sigma\sigma}\left(a,\theta\right) = 0$ *at the kinks* $a = \overline{a}^{\vee}(\theta)$ *and solves, for all other* (a,θ),

$$0 = \mathsf{F}_x\left(a,\theta\right)\mathsf{x}_{\sigma\sigma}\left(a,\theta\right) + \mathsf{F}_{x^e}\left(a,\theta\right)\left(\mathbb{E}_{\varepsilon}\left[\overline{x}_{ZZ,0,0}|a,\theta\right] var\left(\mathcal{E}\right) + \mathbb{E}_{\varepsilon}\left[\mathsf{x}_{\sigma\sigma}|a,\theta\right] + \mathbb{E}_{\varepsilon}\left[\overline{x}_a|a,\theta\right]\mathsf{p}\mathsf{x}_{\sigma\sigma}\left(a,\theta\right)\right). \tag{35}$$

Equation (34) shows that the relationship between $\overline{x}_{\sigma\sigma,t}$ and $\overline{Y}_{\sigma\sigma,t}$ is almost the same as between $\overline{x}_{Z,t}$ and $\overline{Y}_{Z,t}$ with an exception of an additional term $\mathsf{x}_{\sigma\sigma}$ which captures how agents react to risk holding aggregates fixed. The LoM for $\hat{\Omega}_{\sigma\sigma,t}$ is still given by equation (23) with $\mathsf{p}\overline{x}_{Z,t}$ replaced with $\mathsf{p}\overline{x}_{\sigma\sigma,t}$, which makes it easy to describe the derivative $\left(\int \overline{x} d\Omega\right)_{\sigma\sigma,t}$ as in Corollary 1^{FO} and find the linear system of equations that characterizes $\left\{\overline{X}_{\sigma\sigma,t}\right\}_t$. We summarize it in the following proposition.

Proposition 1a^{SO}. *$\left\{\overline{X}_{\sigma\sigma,t}\right\}_t$ is the solution to a linear system*

$$\mathsf{G}_x \sum_{s=0}^{\infty} \mathsf{J}_{t,s} \overline{Y}_{\sigma\sigma,s} + \mathsf{G}_Y \overline{Y}_{\sigma\sigma,t} + \mathsf{G}_x \mathsf{H}_{\sigma\sigma,t} = 0, \tag{36}$$

with $\left\{\overline{Y}_{\sigma\sigma,t}\right\}$ given by equation (33) and $\mathsf{P}\overline{X}_{\sigma\sigma,-1} = 0$, where $\mathsf{H}_{\sigma\sigma,0} = \int \mathsf{x}_{\sigma\sigma} d\Omega^*$, $\mathsf{H}_{\sigma\sigma,t} = \mathsf{H}_{\sigma\sigma,t-1} + \mathcal{I}^{(a)} \cdot \left(\mathcal{L}^{(a)}\right)^{t-1} \cdot \mathcal{M} \cdot \mathsf{a}_{\sigma\sigma}$, and $\mathsf{a}_{\sigma\sigma} = \mathsf{p}\mathsf{x}_{\sigma\sigma}$.

The system of equations (36) is almost the same as (27) except for additional terms $\left\{\mathsf{H}_{\sigma\sigma,t}\right\}_t$. These terms satisfy a simplified version of equation (26), which was used to construct $\mathsf{J}_{t,s}$ for the first-order solution. But differently from its first-order analogues, the precautionary motives do not need to die down as $t \to 0$ and this system of equation can be solve by truncating it at T and imposing a terminal condition that $\overline{X}_{\sigma\sigma,t} = \overline{X}_{\sigma\sigma,T}$ for all $t > T$.[17]

4.2.2 Characterization of interaction terms

To find the interaction terms, we proceed in the same way as in Section 4.2.1 by using various derivatives of G to construct $\overline{X}_{ZZ,t,k}$. This derivative can be written as

$$\mathsf{G}_x \left(\int \overline{x} d\Omega\right)_{ZZ,t,k} + G_Y \overline{Y}_{ZZ,t,k} + \mathsf{G}_{\Theta\Theta,t,k} = 0, \tag{37}$$

where

$$\overline{Y}_{ZZ,t,k} = [0, \mathsf{P}\overline{X}_{ZZ,t-1,k-1}, \overline{X}_{ZZ,t,k}, \overline{X}_{ZZ,t+1,k+1}]^{\mathrm{T}}, \tag{38}$$

and $\mathsf{G}_{\Theta\Theta,t,k}$ combines all the interaction terms known from the first order.[18] As $\mathsf{G}_{\Theta\Theta,t,k}$ is relatively straightforward to construct from the first-order solution, we focus on the other terms. As always, the key step is finding the derivative

$$\left(\int \overline{x} d\Omega\right)_{ZZ,t,k} = \int \overline{x}_{ZZ,t,k} d\Omega^* + \int \overline{x} d\hat{\Omega}_{t,k} + \int \overline{x}_{Z,k} d\hat{\Omega}_t + \int \overline{x}_{Z,t} d\hat{\Omega}_k. \tag{39}$$

While equality (39) might seem obvious at first sight, showing it requires some care because of the kinks in policy functions. It can be written in this simple form because $\overline{x}_{ZZ,t,k}$ is a generalized derivative

[17]Note that finding $\left\{\overline{X}_{\sigma\sigma,t}\right\}_t$ requires knowledge of $\overline{X}_{ZZ,0,0}$ and $\overline{x}_{ZZ,0,0}$ and so in practice one needs to solve for $\left\{\overline{X}_{ZZ,t,s}\right\}_{t,s}$ before finding $\left\{\overline{X}_{\sigma\sigma,t}\right\}_t$.

[18] In particular, for each dimension i of G the derivative $\mathsf{G}^i_{\Theta\Theta,t,k} = \left(\int \overline{x} d\Omega\right)^{\mathrm{T}}_{Z,t} \mathsf{G}^i_{xx} \left(\int \overline{x} d\Omega\right)_{Z,k} + \overline{Y}^{\mathrm{T}}_{Z,t} \mathsf{G}^i_{YY} \overline{Y}_{Z,k} + \left(\int \overline{x} d\Omega\right)^{\mathrm{T}}_{Z,t} \mathsf{G}^i_{xY} \overline{Y}_{Z,k} + \left(\int \overline{x} d\Omega\right)^{\mathrm{T}}_{Z,k} \mathsf{G}^i_{xY} \overline{Y}_{Z,t}$.

which we characterized in Section 3.2. The two intermediary results – Lemmas 2 and 3 – are central to characterizing this equation.

We start with Lemma 2 which shows the relationship between $\{\overline{x}_{ZZ,t,k}\}_{t,k}$ and $\{\overline{X}_{ZZ,t,k}\}_{t,k}$

Lemma 2b^{SO}. *For any t, k*

$$\overline{x}_{ZZ,t,k}(a,\theta) = \sum_{s=0}^{\infty} \mathsf{x}_s(a,\theta)\,\overline{Y}_{ZZ,t+s,k+s} + \mathsf{x}_{t,k}(a,\theta), \tag{40}$$

where $\mathsf{x}_{t,k}(a,\theta) = \mathring{\mathsf{x}}_{t,k}(a,\theta) + \overline{x}_a^{\Delta}(\theta)\overline{a}_{Z,t}^{\vee}(\theta)\overline{a}_{Z,k}^{\vee}(\theta)\delta(a - \overline{a}^{\vee}(\theta))$ with $\mathring{\mathsf{x}}_{t,k}$ solving a recursion

$$\mathring{\mathsf{x}}_{t,k}(a,\theta) = \left(\mathsf{F}_x(a,\theta) + \mathsf{F}_{x^e}(a,\theta)\,\mathbb{E}_\varepsilon\left[\overline{x}_a|a,\theta\right]\mathsf{p}\right)^{-1}\left(\mathsf{F}_{t,k}(a,\theta) + \mathsf{F}_{x^e}(a,\theta)\,\mathbb{E}_\varepsilon\left[\mathring{\mathsf{x}}_{t+1,k+1}|a,\theta\right]\right), \tag{41}$$

and $\mathsf{F}_{t,k}(a,\theta)$ combines known first-order interaction terms given explicitly in the appendix.

As with our simple example (28), the second-order change in individual policy functions $\overline{x}_{ZZ,t,k}$ consists of two terms: the first-order response to the second-order changes in the aggregates, captured by the infinite sum, and second-order responses to the first-order interactions of shock, captured by $\mathsf{x}_{t,k}$. Importantly, this lemma also provides an explicit formula for $\{\mathsf{x}_{t,k}\}$ exclusively in terms of objects from the first-order solution. The generalized function $\mathsf{x}_{t,k}$ consists of two parts: the classical derivative $\mathring{\mathsf{x}}$ that has form very similar to equation (22) in Lemma 2^{FO} and kink adjustments captured by the delta function. We discuss in Section 4.2.3 how to construct the coefficients $\{\mathsf{x}_{t,k}\}$ numerically and how to appropriately handle the delta function component.

The second intermediate result describes the LoM for $\hat{\Omega}_{t,k}$. In Section 4.1 we show that operators $\mathcal{M}$, $\mathcal{L}^{(a)}$, and $\mathcal{I}^{(a)}$ were central to describing the LoM of $\hat{\Omega}_t$. Modifications of the same three operators characterize $\hat{\Omega}_{t,k}$. Let $\mathcal{L}_{Z,t}^{(a)}$ be the derivatives of $\mathcal{L}^{(a)}$ with respect to Z evaluated in direction $\hat{Z}_t$. Mathematically, it takes the same form as $\mathcal{L}^{(a)}$ except $\overline{a}_a$ in the definition is replaced with $\overline{a}_{aZ,t}$. Similarly, we use notations $\mathcal{L}^{(aa)}$, $\mathcal{L}^{(a,a)}$, etc to denote modifications of these operators where $\overline{a}_a$ is replaced with $\overline{a}_{aa}$ and $\overline{a}_a\overline{a}_a$ respectively. Analogous convention applies to $\mathcal{I}^{(a)}$. Finally, we use notation $y' \odot y''$ for two generalized functions y', y'' to denote their point-wise product.

Lemma 3b^{SO}. *For all t, k*

$$\frac{d}{d\theta}\hat{\Omega}_{t+1,k+1} = \mathcal{L}^{(a)} \cdot \frac{d}{d\theta}\hat{\Omega}_{t,k} - \mathcal{M} \cdot \overline{a}_{ZZ,t,k} + \frac{d}{da}\mathsf{c}_{t,k} - \mathsf{b}_{t,k}, \tag{42}$$

where $\mathsf{b}_{t,k}$ and $\mathsf{c}_{t,k}$ satisfy

$$\mathsf{b}_{t,k} = -\mathcal{L}_{Z,t}^{(a)} \cdot \frac{d}{d\theta}\hat{\Omega}_k - \mathcal{L}_{Z,k}^{(a)} \cdot \frac{d}{d\theta}\hat{\Omega}_t,$$

$$\mathsf{c}_{t,k} = \mathcal{M} \cdot (\overline{a}_{Z,t} \odot \overline{a}_{Z,k}) - \mathcal{L}^{(a)} \cdot \left(\frac{d}{d\theta}\hat{\Omega}_t \odot \overline{a}_{Z,k}\right) - \mathcal{L}^{(a)} \cdot \left(\frac{d}{d\theta}\hat{\Omega}_k \odot \overline{a}_{Z,t}\right).$$

This lemma shows that the LoM for $\hat{\Omega}_{t+1,k+1}$ consists of two types of terms: the first order response of the LoM $\overline{\Omega}$ to the second-order changes in policy functions and distribution $\overline{a}_{ZZ,t,k}$, $\hat{\Omega}_{t,k}$, and the second-order response of $\overline{\Omega}$ to the first-order changes, captured by $\mathsf{c}_{t,k}$, $\mathsf{b}_{t,k}$. Inspection of their form reveals that just like $\mathsf{x}_{t,k}$ both $\mathsf{c}_{t,k}$ and $\mathsf{b}_{t,k}$ depend only on first-order terms and thus can be constructed explicitly from the solution in Section 4.1.

Lemmas 2b^{SO} and 3b^{SO} allow us to characterize the second order derivative of the aggregation equations.

Corollary 1b^{SO}. *For all t, k*

$$\left(\overline{\int x d\Omega}\right)_{ZZ,t,k} = \sum_{s=0}^{\infty} \mathsf{J}_{t,s} \overline{Y}_{ZZ,t,k} + \mathsf{H}_{t,k},$$

where $\{\mathsf{H}_{t,k}\}_{t,k}$ *is characterized by the following linear recursive system*

$$\mathsf{H}_{t,k} = \int \mathsf{x}_{t,k} d\Omega^* + \mathcal{I}^{(a)} \cdot \mathsf{B}_{t,k} + \mathcal{I}^{(aa)} \cdot \mathsf{C}_{t,k} - \mathcal{I}^{(a)}_{Z,k} \cdot \frac{d}{d\theta}\hat{\Omega}_t - \mathcal{I}^{(a)}_{Z,t} \cdot \frac{d}{d\theta}\hat{\Omega}_k,$$

$$\mathsf{C}_{t+1,k+1} = \mathcal{M} \cdot (\overline{a}_{Z,t} \odot \overline{a}_{Z,k}) - \mathcal{L}^{(a)} \cdot \left(\frac{d}{d\theta}\hat{\Omega}_t \odot \overline{a}_{Z,k}\right) - \mathcal{L}^{(a)} \cdot \left(\frac{d}{d\theta}\hat{\Omega}_k \odot \overline{a}_{Z,t}\right) + \mathcal{L}^{(a,a)} \cdot \mathsf{C}_{t,k},$$

$$\mathsf{B}_{t+1,k+1} = \mathcal{M} \cdot \mathsf{a}_{t,k} - \mathcal{L}^{(a)}_{Z,t} \cdot \frac{d}{d\theta}\hat{\Omega}_k - \mathcal{L}^{(a)}_{Z,k} \cdot \frac{d}{d\theta}\hat{\Omega}_t + \mathcal{L}^{(a)} \cdot \mathsf{B}_{t,k} + \mathcal{L}^{(aa)} \cdot \mathsf{C}_{t,k},$$

where $\mathsf{a}_{t,k} = \mathsf{p}\mathsf{x}_{t,k}$.

Comparison of Corollaries 1b^{SO} and 1^{FO} reveals several insights. The second-order relationship between the aggregating equation $\int \overline{x} d\Omega$ and aggregate variables $\overline{Y}$ is almost the same as the first-order one except for an additional term $\mathsf{H}_{t,k}$. This term, that captures various first-order interaction effects, has mathematical structure similar to that of $\mathsf{J}_{t,k}$ that we constructed in the first-order approximations. While formulas for $\mathsf{H}_{t,k}$ are lengthier, they recycle several of the same operators that we needed to construct the first-order solution. As we shall see in Section 4.2.3, this makes extending the first order numerical implementation to the second order quite easy.

Equipped with Corollaries 1b^{SO}, it is now easy to go back to equation (37) to find the system of equations that characterizes the second order interaction terms.

Proposition 1b^{SO}. $\left\{\overline{X}_{ZZ,t,k}\right\}_{t,k}$ *is the solution to linear system of equations*

$$\mathsf{G}_x \sum_{s=0}^{\infty} \mathsf{J}_{t,s} \overline{Y}_{ZZ,s,k-t+s} + \mathsf{G}_Y \overline{Y}_{ZZ,t,k} + \mathsf{G}_x \mathsf{H}_{t,k} + \mathsf{G}_{\Theta\Theta,t,k} = 0, \tag{43}$$

with $\left\{\overline{Y}_{ZZ,t,k}\right\}$*given by equation (38) and* $\mathsf{P}\overline{X}_{ZZ,-1,k-t-1} = 0$.

These systems of equations are solved analogously to (27), by imposing the terminal conditions $\overline{X}_{ZZ,s,k+s} = 0$ for all $s + k \geq T$. There are two insights to be gleaned from this proposition. Firstly, our solution is automatically stable. This contrast with traditional perturbation methods that feature unstable paths and require "pruning" to select stable paths (see, e.g., Kim et al. (2008)). By restricting our attention to the second-order derivatives in directions necessary for the second-order expansion, we automatically select the stable path much in the same manner as we avoid the unstable roots in the first-order approximation. Second, Lemma 1^{SO} can be used to find the ergodic mean of the stochastic steady state.[19] Taking expectation of equation (29) and then the limit as $t \to \infty$ finds the long run average level of X to be

$$\mathbb{E}[X] = \overline{X} + \sum_{s=0}^{\infty} \overline{X}_{ZZ,s,s} var(\mathcal{E}) + \lim_{t\to\infty} \overline{X}_{\sigma\sigma,t} + O\left(\|\mathcal{E}\|^3\right). \tag{44}$$

If aggregate welfare is included in X then equation (44) can be quickly used to compute ergodic welfare and evaluate policies that vary over the business cycle. The observation that the difference $\mathbb{E}[X] - \overline{X}$ is of the second order, implies that the approximation error would not improve if one were to approximate the equilibrium dynamics around the mean of the ergodic distribution under aggregate shocks rather than Ω^*.

4.2.3 Numerical implementation of the second-order approximation

To find the interaction terms $\left\{\overline{X}_{ZZ,t,k}\right\}_{t,k}$, we need to construct the new objects in Proposition 1b^{SO}: $\{\mathsf{G}_{\Theta\Theta,t,k}, \mathsf{H}_{t,k}\}$. The terms $\mathsf{G}_{\Theta\Theta,t,k}$ are simply constructed from the first-order objects (see footnote 18), while $\{\mathsf{H}_{t,k}\}_{t,k}$ is given in Corollary 1b^{SO} and requires constructing various modifications of operators from Section 4.1 and computing of $\{\mathsf{x}_{t,k}\}$ from Lemma 2b^{SO}.

Numerical analogues of $\mathcal{I}^{(aa)}$ and $\mathcal{I}^{(a)}_{Z,t}$ are in parallel with $\overrightarrow{\mathcal{I}}^{(a)}$, i.e., $\overrightarrow{\mathcal{I}}^{(aa)}[:,i] = \overrightarrow{x}_{aa}[i]$ and $\overrightarrow{\mathcal{I}}^{(a)}_{Z,t}[:,i] = \overrightarrow{x}_{aZ,t}[i]$.[20] Similarly, the numerical analogues of $\mathcal{L}^{(aa)}$ and $\mathcal{L}^{(a,a)}$ are constructed in parallel with $\overrightarrow{\mathcal{L}}^{(a)}$: $\overrightarrow{\mathcal{L}}^{(aa)}[i',i] = \overrightarrow{\Lambda}[i',i]\overrightarrow{a}_{aa}[i]$ and $\overrightarrow{\mathcal{L}}^{(a,a)}[i',i] = \overrightarrow{\Lambda}[i',i]\overrightarrow{a}_{a}[i]\overrightarrow{a}_{a}[i]$. Finally, construct terms like $\mathcal{L}^{(a)}_{Z,t} \cdot \frac{d}{d\theta}\hat{\Omega}_k$ as $\overrightarrow{\Lambda}\left(\overrightarrow{a}_{aZ,t} \odot \overrightarrow{\frac{d}{d\theta}\Omega_k}\right)$ where $\odot$ is simply point-wise matrix multiplication. Since we approximate policy functions with splines, no additional adjustments is necessary for kinks.

For $\mathsf{x}_{t,k}$, we construct the approximation to the classical and the generalized part separately. For the classical component, we start by setting $\mathring{\mathsf{x}}_{t+1,k+1} = 0$ for $s + k \geq T$ in equation (41) and using backward induction to compute $\mathring{\mathsf{x}}^{\#}_{t,k}$. To adjust for the delta function part, rewrite the expression in Lemma 2b^{SO}

$$\mathsf{x}_{t.k}(z,\theta) = \mathring{\mathsf{x}}_{t+1,k+1} + \frac{d}{da}\left(\underbrace{\iota\left(a \geq \overline{a}^{\vee}(\theta)\right)\overline{x}^{\Delta}(\theta)\overline{a}^{\vee}_{Z,t}(\theta)\overline{a}^{\vee}_{Z,k}(\theta)}_{\equiv \mathsf{x}^{\delta}_{t,k}(a,\theta)}\right).$$

[19] This stochastic steady state should not be confused with the risky steady state of Coeurdacier et al. (2011). The risky steady state relies on a linear approximation of policy rules and, thus, captures only the $\lim_{t\to\infty}\overline{X}_{\sigma\sigma,t}$ term of (44).

[20] $\overrightarrow{x}_{aZ,t}$ is constructed simply as $\overline{x}^{\#}_{Z,t}\Phi_a$.

The function $\mathsf{x}^{\delta}_{t,k}(z,\theta)$ is a step function (that depends on $\overrightarrow{a}^{\vee}_{Z,t}$) that we can approximate with a spline with coefficients $\mathsf{x}^{\delta\#}_{t,k}$. We then recover $\overrightarrow{\mathsf{x}}_{t,k}$ as simply $\mathring{\mathsf{x}}^{\#}_{t,k}\Phi + \mathsf{x}^{\delta,\#}_{t,k}\Phi_a$. The rest of the approach is the same as for other terms.

Once we solved for $\left\{\overline{X}_{ZZ,t,k}\right\}_{t,k}$, we obtain $\overline{X}_{ZZ,0,0}$ and $\overline{x}_{ZZ,0,0}$ that are needed to find $\left\{\overline{X}_{\sigma\sigma,t}\right\}_t$. The only term that still needs to be found in order to solve the system of equations in Proposition 1a^{SO} is $\mathsf{x}_{\sigma\sigma}$. $\mathsf{x}_{\sigma\sigma}$ solves the linear system (35) and thus we find $\mathsf{x}^{\#}_{\sigma\sigma}$ by evaluating (35) at each element of the course grid used for the spline approximation.

$$
\begin{aligned}
0 = \overrightarrow{\mathsf{F}}_x[j]\mathsf{x}^{\#}_{\sigma\sigma}\Phi[:,j] &+ \overrightarrow{\mathsf{F}}_{x^e}[j]\left(\overline{x}^{\#}_{ZZ,0,0}\Phi^e[:,j] + \mathsf{x}^{\delta,\#}_{0,0}\Phi^e_a[:,j]\right) var\left(\mathcal{E}\right) \\
&+ \overrightarrow{\mathsf{F}}_{x^e}[j]\overline{x}^{\#}_{\sigma\sigma}\Phi^e[:,j] + \overrightarrow{\mathsf{F}}_{x^e}[j]\overline{x}^{\#}\Phi^e_a[:,j]\left(\mathsf{p}\mathsf{x}^{\#}_{\sigma\sigma}\right)\Phi[:,j].
\end{aligned}
$$

The conditional expectations are determined using the sparse matrices Φ^e are precomputed from the zeroth-order approximation. This equation is linear in $\mathsf{x}^{\#}_{\sigma\sigma}$ and thus can be solved with a single linear operation.

5 Extensions

We now discuss how our approach can be extended to three classes of problems: models with transition dynamics from some initial distribution to its long run steady state, models with stochastic volatility, and portfolio problems. The first class of problems emerges when one considers permanent shocks or policy changes that induce transition to a new steady state, the second class of problems occurs frequently in studies of asset prices, and the third one emerges whenever agents can invest in more than one asset with different risk characteristics.

5.1 Transition dynamics

Consider an economy as in Section 3 but suppose that the initial condition is given by $(0,\Omega_0)$, where Ω_0 is some distribution that does not necessarily coincides with Ω^*. The equilibrium in this economy now features deterministic transition dynamics as state converges to its steady-state value. It is easy to adapt techniques that we developed in Section 4.1 to compute this transition path.

Let $\hat{\Omega}_0 := \Omega_0 - \Omega^*$ and consider a sequence of directions $\{\hat{Z}'_t\}_t$ defined recursively by $\hat{Z}'_0 = [0,\hat{\Omega}_0]^{\mathrm{T}}$ and $\hat{Z}'_t = \overline{Z}_Z \cdot \hat{Z}'_{t-1}$. Similarly, define $\{\overline{X}'_{Z,t}\}_t$ as $\overline{X}'_{Z,t} := \overline{X}_Z \cdot \hat{Z}'_t$. This sequence characterizes transition dynamics to the first order.

Lemma 1^{TD}. *To the first-order approximation, X_t satisfies*

$$
\mathbb{E}_0 X_t = \overline{X} + \overline{X}'_{Z,t} + O\left(\left\|\mathcal{E},\hat{\Omega}_0\right\|^2\right).
$$

Finding sequence $\{\overline{X}'_{Z,t}\}_t$ is simple. The LoM for the aggregate distribution must still satisfy the recursion shown in Lemma 3^{FO}, except that it is initialized by the difference between the initial and the steady state distributions, $\frac{d}{d\theta}\hat{\Omega}_0$. The rest of the analysis goes unchanged, giving us a proposition that characterizes $\{\overline{X}'_{Z,t}\}_t$:

Proposition 1^{TD}. *$\{\overline{X}'_{Z,t}\}_t$ is the solution to*

$$\mathsf{G}_x \sum_{s=0}^{\infty} \mathsf{J}_{t,s} \overline{Y}'_{Z,s} + \mathsf{G}_X \overline{Y}'_{Z,t} + \mathsf{G}_x \mathsf{J}'_t = 0, \tag{45}$$

$\overline{Y}'_{Z,t} = \left[0, \mathsf{P}\overline{X}'_{Z,t-1}, \overline{X}'_{Z,t}, \overline{X}'_{Z,t+1}\right]^{\mathrm{T}}$, *and* $\mathsf{P}\overline{X}'_{Z,-1} = 0$, *where* $\mathsf{J}'_t = \mathcal{I}^{(a)} \cdot \left(\mathcal{L}^{(a)}\right)^t \cdot \left(-\frac{d}{d\theta}\hat{\Omega}_0\right)$.

The main difference between Proposition 1^{TD} and Proposition 1^{FO} is the last term in equation (45). This term generalizes equation (27) to account for the fact that Ω_0 may differ from Ω^*.

5.2 Stochastic volatility

Many applications that study financial markets or effects of government policies require the volatility of exogenous aggregate variables to be time-varying. A standard way to approximate such models is to consider third-order expansions (see, e.g., discussion in Fernández-Villaverde et al. (2011)). While it is possible to use a third-order extension of our techniques to model stochastic volatility, in this section we show a much simpler second-order approximation that attains the same goal.

Suppose that stochastic process for Θ_t is given by (7) but $\mathcal{E}_t$ is not homoskedastic but rather follows the process

$$\mathcal{E}_t = \sqrt{1 + \Upsilon_{t-1}} \mathcal{E}_{\Theta,t}, \tag{46}$$

$$\Upsilon_t = \rho_\Upsilon \Upsilon_{t-1} + \mathcal{E}_{\Upsilon,t}, \tag{47}$$

where $|\rho_\Upsilon| < 1$ and $\mathcal{E}_{\Theta,t}$ and $\mathcal{E}_{\Upsilon,t}$ are mean-zero i.i.d. variables with support of $\mathcal{E}_{\Upsilon,t}$ bounded so that Υ_t always remains greater than -1. The conditional volatility of aggregate innovations is stochastic and satisfies $var_{t-1}(\mathcal{E}_t) = (1 + \Upsilon_{t-1}) var(\mathcal{E}_{\Theta,t})$. This model collapses to that of Section 2 when Υ_t is a degenerate stochastic process, $\Upsilon_t \equiv 0$.

The state in the recursive representation now consists of a triplet $(\Upsilon, \Theta, \Omega)$. One way to approximate this economy is to scale both shocks $\mathcal{E}_{\Theta,t}$ and $\mathcal{E}_{\Upsilon,t}$ with σ and approximate the equilibrium around the deterministic point $(0, 0, \Omega^*)$. In order to capture time-varying volatility, this approach would indeed require using third-order approximations. Instead, a much faster and simpler method is to proceed as in Section 2 and scale only the combined shock $\mathcal{E}_t$ with σ, just as we did in equation (13). Since shocks $\mathcal{E}_{\Upsilon,t}$ and $\mathcal{E}_{\Theta,t}$ are not scaled with σ, Υ_t still satisfies (47) in the zeroth-order economy. Thus, our approximations are around $(\Upsilon, 0, \Omega^*)$ where Υ is stochastic.

Observe that realizations of Υ have no effect on equilibrium variables in the zeroth-order economy. Therefore, the invariant distribution Ω^* is independent of Υ and coincides with the invariant distribution we discussed in Section 3. Similarly, the derivatives of policy functions $\widetilde{X}(\Upsilon, Z; 0)$ and $\widetilde{x}(z, \theta, \Upsilon, Z; 0)$ with respect to $Z = (\Theta, \Omega)$ are also independent of Υ and coincide with those in Section 3. Derivatives $\overline{X}_\sigma$, $\overline{x}_\sigma$ are equal to zero for any Υ but precautionary motive terms $\overline{X}_{\sigma\sigma}(\Upsilon)$, $\overline{x}_{\sigma\sigma}(\Upsilon)$ are generally non-zero and depend on Υ. This dependence captures equilibrium response to volatility shocks. We now show how our methods can be used to find these coefficients.

Define a sequence of directions $\{\hat{Z}_{\sigma\sigma,t}(\mathcal{E}^t_\Upsilon)\}_{t,\mathcal{E}^t_\Upsilon}$ recursively with $\hat{Z}_{\sigma\sigma,-1} = \mathbf{0}$ and

$$\hat{Z}_{\sigma\sigma,t}\left(\mathcal{E}^t_\Upsilon\right) = \left[0, \overline{\Omega}_{\sigma\sigma}(\Upsilon_{t-1})\right]^{\mathrm{T}} + \overline{Z}_Z \cdot \hat{Z}_{\sigma\sigma,t-1}\left(\mathcal{E}^{t-1}_\Upsilon\right),$$

and the corresponding $\{\overline{X}_{\sigma\sigma,t}(\mathcal{E}^t_\Upsilon)\}_{t,\mathcal{E}^t_\Upsilon}$ as

$$\overline{X}_{\sigma\sigma,t}\left(\mathcal{E}^t_\Upsilon\right) := \overline{X}_{\sigma\sigma}(\Upsilon_t) + \overline{X}_Z \cdot \hat{Z}_{\sigma\sigma,t}\left(\mathcal{E}^t_\Upsilon\right)$$

and

$$\overline{Y}_{\sigma\sigma,t}\left(\mathcal{E}^t_\Upsilon\right) = \left[0, \mathsf{P}\overline{X}_{\sigma\sigma,t-1}\left(\mathcal{E}^{t-1}_\Upsilon\right), \overline{X}_{\sigma\sigma,t}\left(\mathcal{E}^t_\Upsilon\right), \mathbb{E}\left[\overline{X}_{\sigma\sigma,t+1}\middle|\mathcal{E}^t_\Upsilon\right] + \Upsilon_t \overline{X}_{ZZ,0,0} var\left(\mathcal{E}\right)\right]^{\mathrm{T}}.$$

These definitions naturally generalize those of $\{\hat{Z}_{\sigma\sigma,t}\}_t$ and $\{\overline{X}_{\sigma\sigma,t}\}_t$ in Section 4.2 and lead to the extension of Lemma 1^{SO} to settings with stochastic volatility.

Lemma 1^{SV}. *To the second-order approximation, X_t satisfies*

$$X_t\left(\boldsymbol{\mathcal{E}}^t\right) = \overline{X} + \sum_{s=0}^{t} \overline{X}_{Z,t-s}\mathcal{E}_s + \frac{1}{2}\left(\sum_{s=0}^{t}\sum_{m=0}^{t} \overline{X}_{ZZ,t-s,t-m}\mathcal{E}_s\mathcal{E}_m + \overline{X}_{\sigma\sigma,t}\left(\mathcal{E}^t_\Upsilon\right)\right) + O\left(\|\mathcal{E}\|^3\right), \tag{48}$$

where sequences $\left\{\overline{X}_{Z,t}\right\}_t$, $\left\{\overline{X}_{ZZ,t,k}\right\}_{t,k}$ are the same as in Sections 4.1 and 4.2.

Finding $\{\overline{X}_{\sigma\sigma,t}(\mathcal{E}^t_\Upsilon)\}_{t,\mathcal{E}^t_\Upsilon}$ would be complicated if $\overline{X}_{\sigma\sigma,t}(\cdot)$ were an arbitrary non-linear function. Fortunately, this is not the case. Stochastic process (46) simplifies our problem as it implies that $\overline{X}_{\sigma\sigma,t}(\cdot)$ is a linear function of $\mathcal{E}^t_\Upsilon$ that takes a particular simple form. Characterizing this linearity is as easy as finding the first-order impulse responses in Proposition 1^{FO}. The key step to establishing this result is the following proposition. Let $\{\overline{x}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_\Upsilon)\}_{t,\mathcal{E}^t_\Upsilon}$ be defined analogously to $\{\overline{X}_{\sigma\sigma,t}(\mathcal{E}^t_\Upsilon)\}_{t,\mathcal{E}^t_\Upsilon}$. We have the following relationship between these two objects:

Lemma 3^{SV}. *For any t,*

$$\overline{x}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_\Upsilon) = \sum_{s=0}^{\infty} \mathsf{x}_s(a,\theta)\mathbb{E}\left[\overline{Y}_{\sigma\sigma,t+s}|\mathcal{E}^t_\Upsilon\right] + \mathsf{x}_{\sigma\sigma}(a,\theta) + \mathsf{x}^\Upsilon_{\sigma\sigma}(a,\theta)\Upsilon_t, \tag{49}$$

where x_s, $\mathsf{x}_{\sigma\sigma}$ are the same as in Lemma $2b^{SO}$ and where $\mathsf{x}^\Upsilon_{\sigma\sigma}(a,\theta) = 0$ at the kinks $a = \overline{a}^\vee(\theta)$ and solves, for all other (a,θ),

$$0 = \mathsf{F}_x(a,\theta)\,\mathsf{x}^\Upsilon_{\sigma\sigma}(a,\theta) + \mathsf{F}_{x^e}(a,\theta)\left(\mathbb{E}\left[\overline{x}_{ZZ,0,0}|a,\theta\right] var\left(\mathcal{E}\right) + \rho_\Upsilon \mathbb{E}\left[\mathsf{x}^\Upsilon_{\sigma\sigma}|a,\theta\right] + \mathbb{E}\left[\overline{x}_a|a,\theta\right]\mathsf{px}^\Upsilon_{\sigma\sigma}(a,\theta)\right). \tag{50}$$

The key insight of this lemma is that the direct effect of volatility shocks on individual policy functions is linear in Υ_t and is captured by a coefficient $\mathsf{x}_{\sigma\sigma}^{\Upsilon}$. Moreover, this coefficient is closely related to the $\mathsf{x}_{\sigma\sigma}$ coefficient defined by (35), with the only difference being the ρ_{Υ} term which captures the transient nature of the volatility shock: $\mathsf{x}_{\sigma\sigma}^{\Upsilon} \to \mathsf{x}_{\sigma\sigma}$ as $\rho_{\Upsilon} \to 1$. A direct result of (49) is that all policy functions are linear in volatility shocks as well. Moreover, since $\mathsf{x}_{\sigma\sigma}^{\Upsilon}$ is known explicitly, approximation coefficients on $\mathcal{E}_{\Upsilon}^{t}$ can be found much in the same way as the coefficients on $\mathcal{E}^t$ in Proposition 1^{FO}:

Proposition 1^{SV}. *The stochastic process $\overline{X}_{\sigma\sigma,t}(\mathcal{E}_{\Upsilon}^{t})$ satisfies*

$$\overline{X}_{\sigma\sigma,t}\left(\mathcal{E}_{\Upsilon}^{t}\right) = \overline{X}_{\sigma\sigma,t} + \sum_{s=0}^{t} \overline{X}_{\sigma\sigma,t-s}^{\Upsilon} \mathcal{E}_{\Upsilon,s},$$

where $\{\overline{X}_{\sigma\sigma,t}\}_t$ is the same as in Section 4.2 and $\{\overline{X}_{\sigma\sigma,t}^{\Upsilon}\}_t$ satisfies

$$\mathsf{G}_x \sum_{s=0}^{\infty} \mathsf{J}_{t,s} \overline{Y}_{\sigma\sigma,s}^{\Upsilon} + \mathsf{G}_X \overline{Y}_{\sigma\sigma,t}^{\Upsilon} + \mathsf{G}_x \mathsf{H}_{\sigma\sigma,k}^{\Upsilon} = 0, \tag{51}$$

$\overline{Y}_{\sigma\sigma,t}^{\Upsilon} = \left[0, \mathsf{P}\overline{X}_{\sigma\sigma,t-1}^{\Upsilon}, \overline{X}_{\sigma\sigma,t}^{\Upsilon}, \overline{X}_{\sigma\sigma,t+1}^{\Upsilon} + \rho_{\Upsilon}^{t} \overline{X}_{ZZ,0,0} var\left(\mathcal{E}\right)\right]^{\mathrm{T}}$, *and* $\overline{X}_{\sigma\sigma,-1}^{\Upsilon} = 0$ *with* $\mathsf{H}_{\sigma\sigma,0}^{\Upsilon} = \int \mathsf{x}_{\sigma\sigma}^{\Upsilon} d\Omega^{*}$ *and* $\mathsf{H}_{\sigma\sigma,t}^{\Upsilon} = \rho_{\Upsilon} \mathsf{H}_{\sigma\sigma,t-1}^{\Upsilon} + \mathcal{I}^{(a)} \cdot \left(\mathcal{L}^{(a)}\right)^{t-1} \cdot \mathcal{M} \cdot \mathsf{a}_{\sigma\sigma}^{\Upsilon}$

Equation (51) consists of mostly the same objects that were already created to solve equations (36) and (27), so that $\{\overline{X}_{\sigma\sigma,t}^{\Upsilon}\}_t$ can be found very quickly. Comparing Proposition $1\mathrm{b}^{SO}$ to Proposition 1^{SV} with the knowledge that $\mathsf{x}_{\sigma\sigma}^{\Upsilon} \to \mathsf{x}_{\sigma\sigma}$ as $\rho_{\Upsilon} \to 1$, we can easily see that $\overline{X}_{\sigma\sigma,t}^{\Upsilon} \to \overline{X}_{\sigma\sigma,t}$ as $\rho_{\Upsilon} \to 1$ which allows us to interpret $\overline{X}_{\sigma\sigma,t}$ as the response to a permanent change in risk.

5.3 Portfolio problems

Portfolio choice problems–in which agents can allocate their wealth in more than one asset with different risk characteristics–are commonplace, yet incorporating portfolio choice in HA settings presents a unique set of challenges. Optimal portfolios depend on second-order moments, such as risk premia, and having correct portfolios matters even for the first-order responses of equilibrium quantities. This breaks the convenient inductive feature of many approximation methods that proceed sequentially from lower order approximations to higher order approximations. Moreover, in the zeroth-order economy, all assets are risk-free, and in absence of security specific trading costs, optimal portfolios are undetermined. This violates our Assumption 1 and requires adjustments to the techniques that we developed in Section 3.

Portfolio choice problems have a specific mathematical structure that can be exploited to extend our approach to such problems with minimal changes. To illustrate this structure, we start with a simple two asset version of the Krusell and Smith economy before extending our representation to a general class of portfolio problems.

Consider the same economy as in Section 2.1 except suppose that agents can also trade a one-period risk-free bond that is available in the zero net supply. Let R_t^f be the interest rate on this bond between periods $t-1$ and t, and $R_t^x = R_t - R_t^f$ be the excess return to capital. We use $a_{i,t}$ to denote the total wealth of agent i in period t and $k_{i,t}$ as the holdings in capital. Bond holdings are given by $a_{i,t} - k_{i,t}$. Assuming for concreteness that the borrowing constraint is on total assets holdings, agents' optimization problem can be written as the choice of stochastic sequences $\{c_{i,t}, a_{i,t}, k_{i,t}\}_t$ to maximize their utility subject to the borrowing constraint $a_{i,t} \geq 0$ and the budget constraint

$$c_{i,t} + a_{i,t} - W_t \exp(\theta_{i,t}) - R_t^f a_{i,t-1} - R_t^x k_{i,t-1} = 0. \tag{52}$$

Agents' optimality conditions are represented by stochastic sequences $\{a_{i,t}, c_{i,t}, k_{i,t}, \zeta_{i,t}, \lambda_{i,t}\}_{i,t}$ that satisfy (52) and

$$R_t^f U_c(c_{i,t}) - \lambda_{i,t} = 0, \quad U_c(c_{i,t}) + \zeta_{i,t} - \beta \mathbb{E}_t \lambda_{i,t+1} = 0, \quad a_{i,t}\zeta_{i,t} = 0, \tag{53}$$

$$\mathbb{E}_{t-1}\left[\lambda_{i,t} R_t^x\right] = 0. \tag{54}$$

Market clearing conditions for aggregate variables $\left\{K_t, W_t, R_t^f, R_t^x\right\}_t$ are given by (3) and

$$R_t^f + R_t^x + \delta - \alpha \exp\left(\Theta_t\right) K_{t-1}^{\alpha-1} - 1 = 0, \quad K_{t-1} - \int k_{i,t-1} di = 0, \quad K_t - \int a_{i,t} di, \tag{55}$$

with the last equation being equivalent to imposing that the the bond market clears: $\int (a_{i,t} - k_{i,t}) di = 0$.

This example shows the portfolio allocations affect agents' optimality condition via term $R_t^x k_{i,t-1}$, where R_t^x satisfies equation (54). Motivated by this example, consider now a general class of portfolio problems that preserves this feature. Individual optimality conditions are given by

$$F\left(a_{i,t-1}, x_{i,t}, \mathbb{E}_{i,t} x_{i,t+1}, Y_t, R_t^x k_{i,t-1}\right) = 0 \text{ for all } i, t, \tag{56}$$

and

$$\mathbb{E}_{t-1}\left[m_{i,t} R_t^x\right] = 0 \text{ for all } i, t, \tag{57}$$

where $m_{i,t}$ is an some element of $x_{i,t}$ and R_t^x is some element of X_t and $Y_t = [\Theta_t, \mathsf{P}X_{t-1}, X_t, \mathbb{E}_t X_{t+`1}]$. We can write these as $m_{i,t} = \mathsf{S} x_{i,t}$ and $R_t^x = \mathsf{R} Y_t$ for some selection matrices S and R. Aggregate feasibility conditions are given by

$$G\left(\int x_{i,t} di, Y_t\right) = 0 \text{ for all } t \tag{58}$$

and

$$\int k_{i,t-1} di - \mathsf{K} Y_t = 0 \text{ for all } t. \tag{59}$$

Equations (56) and (57) preserve the features of the portfolio problems that we highlighted with the Krusell and Smith example. Equation (59) could be folded into (58) but, as it should get clear shortly, it is useful to keep this equation separate.

This representation naturally nests our example of the Krusell and Smith economy, where assets $a_{i,t-1}$ serve as the individual endogenous state variable $a_{i,t-1}$, and vector $x_{i,t}$ includes all individual choices and appropriate Lagrange multipliers except for $k_{i,t}$. In the appendix, we show how other types of portfolio problems, such as small open economy models and models with different types of risky technologies, map into the same representation. We treat $k_{i,t}$ as a scalar in the text (so that portfolio choice is between two assets) solely for simplicity of exposition, in the appendix, we show that our approach extends naturally to the case with $k_{i,t}$ is an arbitrary-sized vector.

To apply perturbational techniques to portfolio problems, it is important to choose the state-space representation judiciously. In particular, we choose a representation in which the total wealth of an agent is a part of her state, but the allocation of that wealth into different assets is not. Under this convention, policy functions need to depend both on the current and previous period realization of shocks. We use θ_- and Θ_- to denote previous period shocks and portfolio decision from the previous period will depend only on those shocks, while realized return on the portfolios, agents' consumption and savings for next period will additionally depend on current period shocks θ and Θ. Thus, the individual state is a triple (a, θ, θ_-), or $(a, \boldsymbol{\theta})$ for short, the distribution Ω over $(a, \boldsymbol{\theta})$, and the aggregate state $Z = [\Theta, \Theta_-, A, \Omega]^{\mathrm{T}}$ were A represents the pre-determined aggregate variables. Additional measurability restrictions need to be imposed to ensure that a subset of policy functions do not depend on (θ, Θ).

The recursive representation consists of policy functions $\widetilde{x}(a, \boldsymbol{\theta}, Z)$, $\widetilde{k}(z, \boldsymbol{\theta}, Z)$, $\widetilde{X}(Z)$ that satisfy

$$F\left(a, \theta, \widetilde{x}, \mathbb{E}_{\varepsilon,\mathcal{E}}\widetilde{x}, \widetilde{Y}, \widetilde{R}^x\widetilde{k}\right) = 0 \text{ for all } (a, \boldsymbol{\theta}, Z), \tag{60}$$

$$\mathbb{E}_{\varepsilon,\mathcal{E}}\left[\mathsf{S}\widetilde{x}\widetilde{R}^x|\theta_-, \Theta_-, A, \Omega\right] = 0 \text{ for all } (a, \theta_-, \Theta_-, A, \Omega), \tag{61}$$

$$G\left(\int \widetilde{x}d\Omega, \widetilde{Y}\right) = 0 \text{ for all } Z, \tag{62}$$

$$\widetilde{R}^x(Z) = \mathsf{R}\widetilde{Y}(Z), \qquad \int \widetilde{k}d\Omega = \mathsf{K}\widetilde{Y} \text{ for all } Z, \tag{63}$$

$$\mathsf{T}\widetilde{Y}(Z) \text{ and } \widetilde{k}(a, \boldsymbol{\theta}, Z) \text{ are independent of } \Theta \text{ and } (\theta, \Theta) \text{ for all } Z, \tag{64}$$

as well as the LoM for the distribution,

$$\widetilde{\Omega}(Z)\left\langle a', \boldsymbol{\theta}'\right\rangle = \int\int \iota\left(\widetilde{a}(a, \boldsymbol{\theta}, Z) \leq z'\right)\iota(\rho_\theta\theta + \varepsilon \leq \theta')\iota\left(\theta \leq \theta'_-\right)\mu(\varepsilon)\, d\varepsilon d\Omega\left\langle z, \boldsymbol{\theta}\right\rangle \text{ for all } Z. \tag{65}$$

and identities

$$\tilde{a} = \mathsf{p}\tilde{x}, \qquad \tilde{A} = \mathsf{P}\tilde{X}, \qquad \tilde{Y} = \left[\Theta, A, \tilde{X}, \mathbb{E}_{\mathcal{E}}\tilde{X}\right]$$

where $\mathbb{E}_{\varepsilon,\mathcal{E}}\widetilde{x}$ and $\mathbb{E}_{\mathcal{E}}\tilde{X}$ are as defined in section 2.3. The conditional expectation in (61) represents expectations over current shocks, i.e.,

$$\mathbb{E}_{\varepsilon,\mathcal{E}}\left[\widetilde{x}|\theta_-, \Theta_-, A, \Omega\right] = \int \tilde{x}(a, \rho_\theta\theta_- + \varepsilon, \theta_-, \rho_\Theta\Theta_- + \mathcal{E}, \Theta_-, A, \Omega)d\mu(\varepsilon)d\Pr(\mathcal{E}).$$

In this representation, R is the selection matrix that picks out the policy function corresponding to the excess return from vector $\widetilde{Y}$, K is a selection matrix that picks out the aggregate quantity of the risky asset from $\widetilde{Y}$, and T picks out variables only measurable with respect previous period shocks. This last component is necessary due our choice of the state space. Aggregate wealth is a pre-determined variable but prices that clear asset markets that are necessary for the portfolio choice, e.g., R_t^f, are not. These variables will have measurability restrictions that must be imposed via T. Finally, portfolio choices $\widetilde{k}$ are not included in vector $\widetilde{x}$ because they are undetermined in the zeroth-order economy and their limit as $\sigma \to 0$ needs to be found separately.

The key reason for choosing this representation is that it simplifies finding optimal portfolios in the $\sigma \to 0$ limit. In our representation individual wealth and hence the invariant distribution Ω^* are determined in the zeroth-order economy. The limiting portfolios are found together by characterizing the first-order dynamics, and this extra step requires only modest changes to the approach we developed in Section 3. In contrast, other representations that include total wealth and its allocation as parts of the individual state need to confront that the aggregate state is not pinned down in the zeroth-order economy. Under this representation, one would need to find a way to select the appropriate point of approximation Ω^* out of continuum of invariant distributions, which would require additional steps.

We now develop a method to compute the first-order approximation to the equilibrium of this economy. Our analysis closely follows Section 4.1. Directions $\{\hat{Z}_t\}_t$ and $\{\overline{X}_{Z,t}\}_t$ are defined as in that section, except the initial direction is $\hat{Z}_0 = [1, 0, 0, \mathbf{0}]^{\mathrm{T}}$. With this adjustment, statement of Lemmas 1^{FO} as well as equations (17) and (18) carry over without changes to portfolio settings. The other steps of the analysis require some adjustments.

Let $\overline{R}_{Z,0} = \mathsf{R}\overline{Y}_{Z,0}$ be the contemporaneous effect of the aggregate shock on risky returns and let $\mathfrak{S}(\overline{R}_{Z,0})$ be a mapping defined as

$$\mathfrak{S}\left(\overline{R}_{Z,0}\right) := \frac{1}{(\overline{R}_{Z,0})^2 var\left(\mathcal{E}\right)}.$$

Similarly, let $\overline{X}_{\sigma\sigma,0}$ and $\overline{Y}_{\sigma\sigma,0}$ be defined as in Section 4.2 and let $\overline{R}_{\sigma\sigma,0} = \mathsf{R}\overline{Y}_{\sigma\sigma,0}$ be the risk premium of the asset. This is the only second-order term that we need to find in order to characterize agents' portfolio problems.[21]

Lemma 2^{PF}. $\overline{x}_{Z,0}\left(a, \boldsymbol{\theta}\right)$ *satisfies*

$$\overline{x}_{Z,0}\left(a, \boldsymbol{\theta}\right) = \sum_{s=0}^{\infty} \mathsf{x}_s\left(a, \theta\right)\overline{X}_{Z,s} + \mathsf{r}\left(a, \theta\right)\overline{R}_{Z,0}\overline{k}\left(a, \theta_{\text{-}}\right), \tag{66}$$

[21] The observation that only a small number of second-order moments is needed to find the optimal portfolio exploits the structure of the portfolio problems embedded in equations (60) and (61). We build on the insights of Devereux and Sutherland (2011).

where

$$\overline{k}(a,\theta_{-}) = \mathfrak{S}\left(\overline{R}_{Z,0}\right)\left[\mathsf{k}_{\sigma\sigma}(a,\theta_{-})\overline{R}_{\sigma\sigma,0} + \sum_{s=0}^{\infty}\mathsf{k}_s(a,\theta_{-})\overline{X}_{Z,s}\overline{R}_{Z,0}var(\mathcal{E})\right] \tag{67}$$

and explicit expressions for r, $\mathsf{k}_{\sigma\sigma}$ *and* k_s *are given in the appendix.* $\overline{x}_{Z,t}\left(a,\boldsymbol{\theta}\right)$ *is independent of* θ_{-} *and satisfies (20) and* $\overline{X}_{Z,t}$ *satisfies* $\mathsf{R}\overline{X}_{Z,t} = 0$ *for all* $t > 0$.

This lemma describes the first-order relationship between individual and aggregate policy functions. Portfolio problems introduce one additional term, $\mathsf{r}\overline{R}_{Z,0}\overline{k}$, to the equation that describes $\overline{x}_{Z,0}$; expressions that describe $\overline{x}_{Z,t}$ for all other t remain the same as in Lemma 2^{FO}.

We now explain the intuition for these results. $\overline{k}(a,\theta_{-})$ is the optimal portfolio for agent (a,θ_{-}) in the limit as $\sigma \to 0$. The expression for this portfolio is given in equation (67) and takes a form that is familiar from the classical portfolio theory (see, e.g., Viceira (2001)). $\overline{R}_{\sigma\sigma,0}$ captures assets' expected excess returns, $\overline{Y}_{Z,s}\overline{R}_{Z,0}var(\mathcal{E})$ is the covariance between returns and the s-period ahead aggregate variable, and $\mathfrak{S}\left(\overline{R}_{Z,0}\right)$ is the inverse of the covariance matrix of returns that captures risk adjustments. Thus, the optimal portfolio depends on assets' risk-adjusted returns $\mathfrak{S}\left(\overline{R}_{Z,0}\right)\overline{R}_{\sigma\sigma,0}$ and their risk-adjusted hedging characteristics $\{\mathfrak{S}\left(\overline{R}_{Z,0}\right)\overline{Y}_{Z,s}\overline{R}_{Z,0}var(\mathcal{E})\}_s$. Coefficients $\mathsf{k}_{\sigma\sigma}$ and $\{\mathsf{k}_s\}_s$ describe how agents weight these objectives. These coefficients depend on agents' risk aversion as well as on how aggregate shocks affect agents' non-tradable labor income.

The intuition for the remaining parts of Lemma 2^{PF} is straightforward. $\overline{R}_{Z,0}\overline{k}$ is the realized return on the risky portfolio and r captures how this return affects individual policy functions at the time of the shock, $\overline{x}_{Z,0}$. There are no analogues of this term in the expressions for other $\overline{x}_{Z,t}$ since those analogues would depend on the expected excess returns, which are zero to the first order (this effect is captured by the result that $\mathsf{R}\overline{X}_{Z,t} = 0$ for $t > 0$ stated in the lemma).

The observation that portfolio problems introduce an additional adjustment term in the initial period policy rules carries over to another key result from Section 4.1, the recursive characterization of the LoM for the aggregate distribution. If we extend the definitions of operators $\mathcal{L}^{(a)}$, $\mathcal{M}$, and $\mathcal{I}^{(a)}$ so that they integrate over (θ_{-},θ), and use $\frac{d}{d\boldsymbol{\theta}}\hat{\Omega}_t$ to denote $\frac{d}{d\theta}\frac{d}{d\theta_{-}}\hat{\Omega}_t$ the the statement of Lemma 3^{FO} carries over with $\frac{d}{d\theta}\hat{\Omega}_t$ replaced with $\frac{d}{d\boldsymbol{\theta}}\hat{\Omega}_t$. As a result, we can extend Corollary 1^{FO} by including the additional terms in $\overline{a}_{Z,0}$ resulting from the portfolio choice and propagating them to future periods via the LOM.

Corollary 1^{PF}. *For any* t,

$$\left(\overline{\int xd\Omega}\right)_{Z,t} = \sum_{s=0}^{\infty}\mathsf{J}_{t,s}\overline{Y}_{Z,s} + \left[\sum_{s=0}^{\infty}\left(\mathsf{J}_{t,s}^{PF}\overline{Y}_{Z,s}\right)var(\mathcal{E})\overline{R}_{Z,0}\mathfrak{S}\left(\overline{R}_{Z,0}\right) + \mathsf{J}_{\sigma\sigma,t}^{PF}\overline{R}_{\sigma\sigma,0}\mathfrak{S}\left(\overline{R}_{Z,0}\right)\right]\overline{R}_{Z,0}$$

where $\mathsf{J}_{t,s}$ *is the same as in Corollary* 1^{FO}, $\mathsf{J}_{\sigma\sigma,0}^{PF} = \int \mathsf{r}(a,\theta)\mathsf{k}_{\sigma\sigma}(a,\theta_{-})d\Omega^{*}$ *and* $\mathsf{J}_{0,s}^{PF} = \int \mathsf{r}(a,\theta)\mathsf{k}_s(a,\theta_{-})d\Omega^{*}$

and, for $t \geq 1$,

$$\mathsf{J}^{PF}_{\sigma\sigma,t} = \mathcal{I}^{(a)} \cdot \left(\mathcal{L}^{(a)}\right)^{t-1} \cdot \mathcal{M} \cdot \mathsf{a}^{PF}_{\sigma\sigma} \tag{68}$$

$$\mathsf{J}^{PF}_{t,s} = \mathcal{I}^{(a)} \cdot \left(\mathcal{L}^{(a)}\right)^{t-1} \cdot \mathcal{M} \cdot \mathsf{a}^{PF}_{s} \tag{69}$$

with $\mathsf{a}^{PF}_{\sigma\sigma}(a,\theta,\theta_{\text{-}}) = \mathsf{pr}(a,\theta)\mathsf{k}_{\sigma\sigma}(a,\theta,\theta_{\text{-}})$ *and* $\mathsf{a}^{PF}_{s}(a,\theta,\theta_{\text{-}}) = \mathsf{pr}(a,\theta)\mathsf{k}_{s}(a,\theta,\theta_{\text{-}})$

The difference between Corollary 1^{FO} and Corollary 1^{PF} captures how the realized portfolio returns alter the aggregated decisions of the agents. As the portfolio choice depends on the assets' risk-adjusted returns $\mathfrak{S}\left(\overline{R}_{Z,0}\right)\overline{R}_{\sigma\sigma,0}$ and their risk-adjusted hedging characteristics $\{\mathfrak{S}\left(\overline{R}_{Z,0}\right)\overline{Y}_{Z,s}\overline{R}_{Z,0}var(\mathcal{E})\}_s$, there are corresponding weighting matrices, e.g., $\mathsf{J}^{PF}_{\sigma\sigma,t}$ and $\mathsf{J}^{PF}_{t,s}$, which summarize their affects on $\left(\overline{\int x d\Omega}\right)_{Z,t}$. While these portfolio returns only directly affect decisions at $t=0$, they alter $\left(\overline{\int x d\Omega}\right)_{Z,t}$ through the LoM of the distribution as described by equations (68) and (69). As they are built using the same operators, the $\{\mathsf{J}^{PF}_{t,s}\}_{t,s}$ and $\{\mathsf{J}^{PF}_{\sigma\sigma,t}\}_t$ are constructed similarly to $\{\mathsf{J}_{t,s}\}_{t,s}$ and $\{\mathsf{J}_{\sigma\sigma,t}\}_t$ as we described in Section 3.

Combining Lemma 2^{PF} and Corollary 1^{PF} with equations (17) and (18) yields the main result of this section

Proposition 1^{PF}. $\left\{\overline{X}_{Z,t}\right\}_t$ *and* $\overline{R}_{\sigma\sigma,0}$ *are the solution to*

$$\begin{aligned}0 =& \mathsf{G}_x \sum_{s=0}^{\infty} \mathsf{J}_{t,s}\overline{Y}_{Z,s} + \mathsf{G}_X \overline{Y}_{Z,t} \\ &+ \mathsf{G}_x \left[\sum_{s=0}^{\infty}\left(\mathsf{J}^{PF}_{t,s}\overline{Y}_{Z,s}\right) var(\mathcal{E})\overline{R}_{Z,0}\mathfrak{S}\left(\overline{R}_{Z,0}\right) + \mathsf{J}^{PF}_{\sigma\sigma,t}\overline{R}_{\sigma\sigma,0}\mathfrak{S}\left(\overline{R}_{Z,0}\right)\right]\overline{R}_{Z,0},\end{aligned} \tag{70}$$

as well as

$$\mathsf{K}_{\sigma\sigma}\overline{R}_{\sigma\sigma,0}\mathfrak{S}\left(\overline{R}_{Z,0}\right) + \sum_{s=0}^{\infty}\mathsf{K}_s\overline{X}_{Z,s}var\left(\mathcal{E}\right)\overline{R}_{Z,0}\mathfrak{S}\left(\overline{R}_{Z,0}\right) = \mathsf{K}\overline{Y}, \tag{71}$$

$$\mathsf{T}\overline{Y}_{Z,0} = 0, \qquad \mathsf{R}\overline{Y}_{Z,t} = 0 \text{ for } t \geq 1, \tag{72}$$

$$\overline{R}_{Z,0} = \mathsf{R}\overline{Y}_{Z,0}, \tag{73}$$

(18) *and* $\mathsf{P}\overline{X}_{Z,-1} = 0$ *where* $\mathsf{K}_{\sigma\sigma} = \int \mathsf{k}_{\sigma\sigma} d\Omega^*$, $\mathsf{K}_s = \int \mathsf{k}_s d\Omega^*$.

Equation (70) shows that portfolio problems add one extra term, given in the second line, to the characterization of the first-order equilibrium dynamics. This additional term has by now familiar structure and interpretation.

Equation (70) depends on one second-order term, $\overline{R}_{\sigma\sigma,0}$. Fortunately, this term can be found from equation (71), circumventing the need to do the full second-order expansion. This both simplifies the first-order analysis and preserves the convenient structure of the perturbation approach that allows to find the rest of higher order terms using the first-order solution of Proposition 1^{PF}.

Equation (71) has a simple interpretation. Its left hand side is the demand for the risky asset, which simply integrates individual portfolio problems (67). The right hand side of (71) is asset supply. Thus, equation (71) determines the risk premium that clears the asset market in general equilibrium.

The system of equations (70), (71), (72) and (73) is non-linear in $\{\overline{X}_{Z,t}\}_t$ and $\overline{R}_{\sigma\sigma,0}$ due to the nonlinear operator $\mathfrak{S}(\overline{R}_{Z,0})$. Despite this nonlinearity, it can be solved quite easily. Observe that for a fixed $\overline{R}_{Z,0}$ equations (70), (71) and (72) form a linear system that determine $\{\overline{X}_{Z,t}\}_t$ and $\overline{R}_{\sigma\sigma,0}$. This observation provides a natural algorithm for solving the system of equations in Proposition 1^{PF}: guess $\overline{R}_{Z,0}$ and solve the linear system (70) – (72) for $\{\overline{X}_{Z,t}\}_t$ and $\overline{R}_{\sigma\sigma,0}$; verify if the initial guess satisfies equation (73); if necessary, adjust the guess for $\overline{R}_{Z,0}$ and iterate until a solution is found.

6 Comparison to literature

Our approach builds on the perturbational techniques in the spirit of Judd (1998) and Schmitt-Grohé and Uribe (2004) originally developed to study dynamic representative agent models. The key difficulty in extending them to HA environments lies in the fact that derivatives of policy functions with respect to the aggregate state (captured by $\overline{X}_Z$, $\overline{X}_{ZZ}$, etc in our notation) are intractably large objects. The seminal paper by Reiter (2009) takes a step overcome this hurdle by discretizing the state space and the transition probability matrix (using the so-called "histogram method", see also Young (2010)), using finite-dimensional histogram as the representation of Ω, and then applying standard perturbational techniques to this finite dimensional state space. To obtain first-order approximations, this method requires solving large quadratic matrix equations and proved to be too slow and imprecise in many standard HA environments.[22]

One strand of literature, originally proposed by Boppart et al. (2018) and then significantly developed by Auclert et al. (2021), abandons the state-space representation used in Reiter (2009) and variants and works with the sequence-space formulation of the problem. The key observation for that approach is that the first-order impulse responses of the stochastic economy can be fully constructed from deterministic responses to MIT shocks, and that these response can recovered numerically fairly easily from the sequence problem. Auclert et al. (2021) show that this can be done very fast as those impulse responses

[22]The key issues are both the time taken to compute and space needed to store those derivatives. This is most clear from Reiter's implementation of the Krusell and Smith model that is solved by discretizing the θ process and using a histogram to store the distribution Ω. If we were to follow the standard convention of using between 1000-5000 points per θ for the histogram, and use 10 points for the shocks, the size of the histogram $N_\Omega \sim 10^4$. This means that $\overline{\Omega}_Z \sim 10^8$ and $\overline{\Omega}_{ZZ} \sim 10^{16}$ entries. Assuming that 4 bytes (float) are required to store an entry, this would mean that one needs 450 megabytes of RAM to store the first derivative and 4 terabyte of RAM to store the second-order derivative, which is clearly outside the scope of the current computing architectures. The general argument applies to variants of Reiter (2009) such as Bayer et al. (2022); Ahn et al. (2018); Childers (2018); Winberry (2018); Gornemann et al. (2021) in different degrees depending on the method and the application studied.

solve a linear system of equations which coefficients can be constructed using linear recursive equations.

Our approach combines insights from both strands of the literature but also approaches approximations differently from Reiter (2009) and the literature that followed him. This allows us to improve on the computational speed of Auclert et al. (2021) method, and to have our approach scalable to second- and higher-orders of approximation, which is one of the key features of classical perturbational techniques a-la Judd (1998) and Schmitt-Grohé and Uribe (2004) but not of Reiter (2009) and papers building on it.

The key distinction of our approach is that we start with the theoretical distribution and its LoM and derive exact analytical expressions for approximations of various orders; numerical values of those expressions are then computed using appropriate discretization. This contrast with papers following the Reiter (2009) tradition that start with an approximate (i.e., already discretized) distribution and a transition probability matrix before further approximating with respect to aggregate shocks. There is no guarantee that this latter approach correctly would correctly recover aggregate responses beyond the first order. In particular, we show in the appendix D that the second-order approximation of the transition probability obtained under histogram method generically misses some of the second-order order terms and that does not converge to the exact second-order expressions as the grid size goes to zero. The intuition for this results is that the histogram method locally linearizes the LoM for the aggregate distribution, which misses terms capturing second-order responses of the LoM to the first-order changes in policy functions.[23]

Our description of approximations as a sequence of values of derivatives such as $\{\overline{X}_{Z,t}\}_t$, $\{\overline{X}_{ZZ,t,s}\}_{t,s}$, etc is related to the "MIT shock" strand of literature. One can show that to the first order, our approach is equivalent to that of Auclert et al. (2021) in the sense that as the grid size of their approximations goes to zero, the linear system of equations they use to describe approximations converges to our system (27). Despite this equivalence, using state-space representation has advantages even to the first order, as it allows us to derive analytically and then construct recursively coefficients $\mathsf{x}_s = \partial x_0 / \partial X_t$ in Lemma 2^{FO}. In contrast, the sequence-space approach finds $\{\partial x_0 / \partial X_t\}_t$ using numerical differentiation of the (truncated) infinite system of equations (5). This process is both slower and less stable numerically. State space representation also significantly simplifies and speeds up computation of first-order transition dynamics, as in Section 5.1.

The bigger advantage of our use of the state-space representation is that it enables recursive representations of the necessary directional derivatives at all orders. For instance, by explicitly specifying directions that characterize the effects of persistent risk $\left\{\hat{Z}_{\sigma\sigma,t}\right\}$, our approach can incorporate risk and go beyond MIT shocks. This is imperative for questions such as finding first-order impulse responses

[23] In models set in continuous time, several authors used the exact LoM of the aggregate distribution (see, e.g., Alvarez and Lippi (2022); Alvarez et al. (2023b), Bilal (2022)) to characterize analytically equilibrim properties.

in models with portfolio choice, understanding effect of risk or welfare costs of aggregate shocks, and studying trade-offs involved in designing macroeconomic stabilization policies.

7 Numerical results

In this section, we apply our algorithm to calibrated versions of the Krusell and Smith (1998) model. First, we use the calibrated model to report diagnostics such as sped and accuracy and compare them to alternative methods. Second, we use extensions of the Krusell and Smith model to study several applications that illustrate the usefulness of our methods over and above what can be achieved with existing approaches. These applications include welfare analysis of fiscal stabilization policies, aggregate and distributional effects of fluctuations in macroeconomic uncertainty, role of household portfolios, and transitions across steady states.

7.1 Baseline Model

Our baseline model extends the Krusell and Smith framework of Section 2 to include capital adjustment costs. This allows the model to generate volatile returns to risky capital that is useful for some of our applications. To enable convenient aggregation, we introduce a competitive mutual funds sector whose shares are owned and traded by households. The household's budget constraint is modified to

$$c_{i,t} + k_{i,t} = w_t e^{\theta_{i,t}} + R_t k_{i,t-1},$$

where $k_{i,t}$ now is the date t wealth of the household. The mutual fund gathers rental income from the corporate sector, owns and invests in physical capital subject to a convex adjustment of the form

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

In the appendix, we show that the equilibrium is given by the same equations (2), (1), (3) with a modified equation for return on savings[24]:

$$R_t = \frac{(1-\alpha)\exp(\Theta_t) K_t^{\alpha} - I_t - \frac{\phi}{2}\left(\frac{I_t}{K_t} - \delta\right)^2 K_t + Q_t K_{t+1}}{Q_{t-1} K_t}, \quad Q_t = 1 + \phi\left(\frac{I_t}{K_t} - \delta\right). \tag{74}$$

Calibration To calibrate our model, we set the period length to one quarter. The parameter α is set to 0.36 to target the capital share of income. We use an isoelastic period utility $U(c) = \frac{c^{1-\gamma}}{1-\gamma}$ and vary the risk aversion parameter γ between 2 and 7. For each choice of the risk aversion parameter, we set adjustment cost parameter ϕ to match a 3% standard deviation of un-leveraged quarterly returns to equity. Unless otherwise specified, the plots in this section are for risk aversion set to 5. For the parameters governing the aggregate and idiosyncratic labor productivity in (7) and (8), we choose values used by Auclert et al. (2021). The calibrated parameters are summarized in Table 1.

[24]The date t aggregate capital K_t used in production satisfies $Q_{t-1}K_t = \int k_{i,t-1}$ and appropriately adjust (1)

Table 1: CALIBRATION OF THE KRUSELL-SMITH ECONOMY

Parameter	Description	Value
α	Capital share	0.36
β	Discount factor	0.983
γ	Risk aversion	$[2, 7]$
δ	Depreciation rate of capital	1.77%
ϕ	Adjustment cost of capital	$[32, 125]$
ρ_ϵ	Idiosyncratic mean reversion	0.966
$\sigma_\epsilon/\sqrt{1-\rho_\epsilon^2}$	Cross-sectional std of log earnings	0.503
ρ_Θ	Aggregate mean reversion	0.80
σ_Θ	Std of Aggregate TFP shocks	0.014
N_ϵ	Points in Markov chain for ϵ	7
N_z	Grid points for the policy rule $\bar{x}^i(z)$	60
I_z	Grid points for the distribution $\bar{\omega}_i$	1000
T	Time horizon (in quarters) for IRF	400

Simulations We use the baseline calibration to simulate policy functions using Lemma 1^{SO}. This involves computing the zeroth-, first- and second-order terms.

To compute the zeroth-order terms, we solve the non-stochastic steady state policy functions using an endogenous grid method after discretizing the productivity with $N_\epsilon = 7$ and asset grid $N_z = 60$. and we use $N_\Omega = 1000 \times 7$ points to store the distribution. To compute the first- and second-order terms, we implement the steps detailed in Section 4.1.1 and 4.2.3. In Table 2, we report total time taken to compute those terms and break up the time by each step stage of the algorithm. The timings for the first-order approximation are reported in the first two columns of the table and the timings for steps to compute the second order are reported rest of the columns.[25]

Table 2: COMPUTATIONAL SPEED: FIRST AND SECOND ORDER

First Order		Second Order		
Step	Time	Step	Time (ZZ)	Time$(\sigma\sigma)$
		Additional First-Order Terms	0.70s	
Lemma 2^{FO} Terms	0.07s	Lemma 2b^{SO} Terms	0.64s	0.05s
Lemma 3^{FO} Terms	0.13s	Lemma 3b^{SO} Terms	0.21s	0.45s
Corollary 1^{FO} Terms	0.17s	Corollary 1b^{SO} Terms	0.07s	0.05s
Proposition 1^{FO} Terms	0.13s	Proposition 1b^{SO} Terms	0.19s	0.28s
Total	0.5s		1.81s	0.83s
ABRS	1.51s			

All told, once the steady state has been computed, our algorithm takes 0.5 seconds to solve for the $\overline{X}_{Z,t}$ terms with roughly equal time spent in all 4 of the main steps. As Lemma 1^{FO} highlights, $\overline{X}_{Z,t}$ are all that is needed to simulate the path of aggregates and to compute ergodic moments from the

[25] All numbers are reported using a 20 core M1 ultra mac studio.

first-order approximation. The other first-order terms, $\bar{x}_{Z,t}$ and $\bar{\Omega}_{Z,t}$, are required for the second-order approximation and take an additional 0.7 seconds to compute.

We compare this to our own implementation of the Sequence Space Jacobian of ABRS which takes approximately 1.5 seconds to compute the equivalent on the $\overline{X}_{Z,t}$. Of that time, approximately 1.35 seconds are spent on the backward and forwards iteration steps which are the equivalent of the terms computed in Lemma 2^{FO} and 3^{FO}. As we detailed in Section 4.1.1, once the steady state is solved our algorithm requires only sparse linear operations which can be done quickly and efficiently independently of how the steady state is solved for. The methodology of ABRS generally relies on numerical differentiation of global transition code, and is therefore limited by the efficiency of that global code. Moreover, very careful attention has to be paid to those numeric derivatives in order to ensure that they are accurate, (See appendix C.1 of Auclert et al. for details). These numerical issues would be amplified with a second-order approximation as calculating second derivatives are more prone to numerical error. By giving explicit expressions for these second derivatives in terms of derivatives of F and G we sidestep these issues.

The addition time to compute the second-order approximation is broken out in the last two columns of Table 2. As highlighted in Section 4.2 there are two additional types of terms in the second-order approximation: the curvature terms, $\bar{X}_{ZZ,t,k}$, and risk correction terms $\bar{X}_{\sigma\sigma,t}$. As they follow the same mathematical structure, we break out the computational time separately for both types. The curvature terms take 1.11 seconds to compute[26] while the risk adjustment terms take 0.83 seconds. The vast majority of the computational time for the curvature terms is spent on Lemma 2b^{SO} and Proposition 1b^{SO} which is a result of a large number of quadratic forms required to compute the $\mathsf{x}_{t,k}(z,\theta)$ and $\mathsf{G}_{\Theta,t,k}$ terms. All combined, computing the second-order approximation requires an additional 2 to 3 seconds relative to the first-order approximation.

Next, we use the terms we computed to construct simulations using alternative methods. In the left panel of Figure 1, we plot simulations for one-time impulse at date $t = 0$, that is, $\{\mathcal{E}_t\} = (1, 0, 0, \ldots)$. The plots shows that the first-order approximation under our method and the approximation of ABRS overlap. This overlap is reassuring that issues related to numerical derivatives and coarseness of asset grid are not quantitatively large. The plots also show that simualted path with first and second-order terms is quite different from ABRS emphasizing the quantitative importance of the second-order terms. In the case shown in Figure 1, the differences are mainly due to the $X_{\sigma,\sigma,t}$ terms from equation (36).

[26] Here we report only the time required to compute that $\bar{X}_{ZZ,t,t}$ terms. We do this for two reasons. Firstly, for most ergodic moments only the $\bar{X}_{ZZ,t,t}$ are required. Secondly, computing the addition $\bar{X}_{ZZ,t,t+i}$ terms are trivially parallelizable for each i so, with enough processors, computing all the $\bar{X}_{ZZ,t,k}$ terms would not require any additional time.

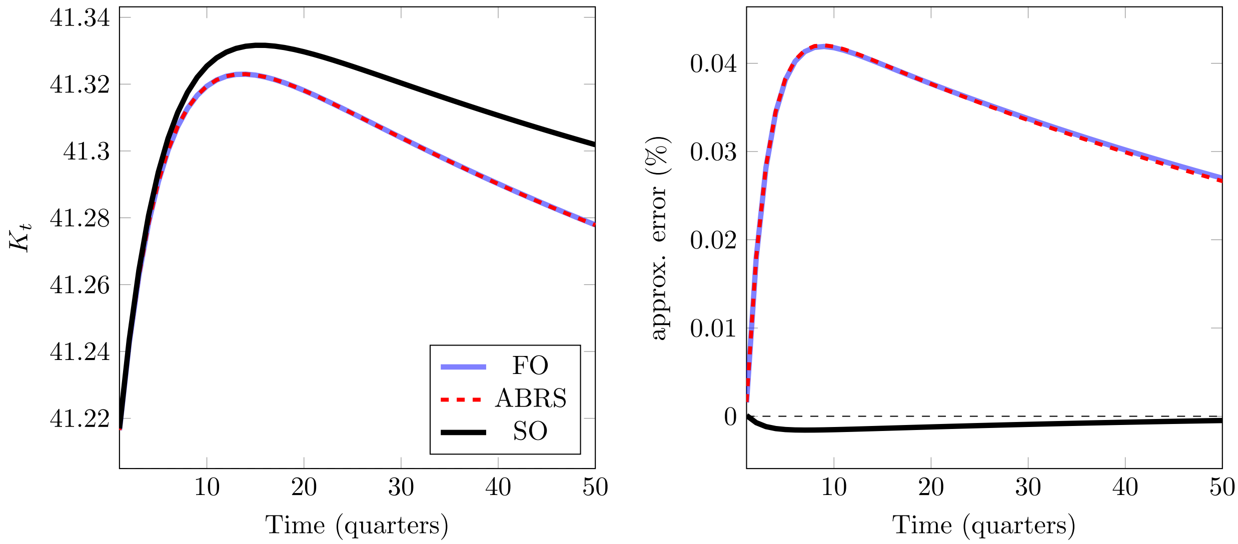

Figure 1: The left panel is the simulated path of aggregate capital $K_t\left(\mathcal{E}^t\right)$ and the right panel is the Nonlinear error $\frac{\hat{K}_t-\tilde{K}_t}{\tilde{K}_t}$ for $\{\mathcal{E}_t\}_t = (1, 0, 0, \ldots)$

Accuracy Measuring the numerical errors made in the simulation of heterogeneous agents models with aggregate shocks is known to be difficult since there are no "reference point" to compare with. As such, we test the accuracy of ours and alternative methods by studying the response to a one-time, one standard deviation positive shock to TFP which can be solved non-linearly and compared to the approximations $\hat{X}_t$ under alternative perturbation methods. We can measure the accuracy by comparing $\hat{X}_t$ with the non-linear counterpart. In the right panel of Figure 1, we plot the % error in the capital stock . For comparison purposes we show errors using our approach described in Section 4.1 and the Sequence Space Jacobian approach of ABRS. As would be anticipated by the right panel of Figure 1 all three approaches have roughly the same error to first order, with the maximal error being on the order of 0.04% of the capital stock. At higher orders, our approach has errors which remain very small over time.

7.2 Applications

In this section we study four applications that highlight the usefulness our method for heterogeneous agent models.

7.2.1 Welfare from stabilization policy

Second-order approximations can be used to evaluate welfare effects of stabilization policies such as fiscal or monetary rules that describe how taxes or interest rates vary over business cycles. Here, we extend the baseline Krusell and Smith model to include a fiscal rule that in form of a time varying labor-tax

$$\tau_t = \tau_\Theta \Theta_t,$$

which is returned lump-sum to the households. Households with labor productivity $\theta_{i,t}$ will receive transfers T_t and $(1-\tau_t)W_t \exp(\theta_{i,t})$ in after-tax labor income in the current period.

A redistributive planner faces a non-trivial tradeoff in transferring resources across agents and insurance across states with the choice of τ_Θ. We are interested in illustrating how to use our method to conduct welfare analysis across different choices of τ_Θ. Welfare comparisons across τ_Θ can be meaningfully answered only with a minimum of second-order expansion. The stabilization coefficient, τ_Θ doesn't affect the non-stochastic steady state and to the first-order of approximation welfare is constant with respect to τ_Θ as certainty equivalence holds.

We follow steps from section 4.2 and equation (44) to approximate welfare. For a given τ_Θ, define utilitarian welfare as $\mathcal{W}(\Omega, \Theta; \tau_\Theta) = \int v(\theta, k, \Theta, \Omega; \tau_\Theta)\, d\Omega$ where v is the value of an individual with who starts with idiosyncratic states (θ, k) when the aggregate state is (Θ, Ω) under policy indexed by τ_Θ. We extend x and X to include v and $\mathcal{W}$, respectively, and add the Bellman equation that solves the value function v to the mapping F and the definition of welfare $\mathcal{W}$ to the mapping G, our framework computes welfare automatically. As mentioned before, it takes only a few seconds to calculate welfare for a given τ_Θ.

In Figure 2 we plot ergodic welfare (computed as consumption equivalent relative to steady state) as a function of the tax parameter τ_Θ. For risk aversion set to 2, we see that relative to a laissez-faire policy, $\tau_\Theta = 0$, making the tax policy more countercyclical initially raises welfare with a distinct maximum (denoted by τ_Θ^*) achieved at $\tau_\Theta = -3.1$ which amounts to raising taxes by 3.1 percentage points for every percentage point decrease in TFP. In fact, roughly 22% of the welfare losses from business cycles in this model can be ameliorated by this tax policy. In Table 3, column τ_Θ^*, we report the optimal cyclicalty for other values of risk aversion. We find that higher the risk aversion, lower the cyclicality. Because gains costs of insurance increase with risk aversion, higher values of risk aversion are associated with lower values for tax cyclicality.

This application also serves as a valuable tool for illustrating the shortcomings associated with employing the histogram technique. In Section 6, we emphasized the consequences of naively extending the histogram approach, which overlook specific second-order terms. These second-order terms become particularly crucial when calculating welfare derived from stabilization policy, which is inherently a second-order object. In columns $\frac{\mathcal{W}^{\text{hist}}(\tau_\Theta^*)}{\mathcal{W}(\tau_\Theta^*)}$ and $\frac{\tau_\Theta^{*,\text{hist}}}{\tau_\Theta^*}$ of Table 3, the ergodic welfare and optimal cyclicality under the histogram method are presented. It is evident that both the magnitude of welfare corresponding to a particular τ_Θ and the gradient of welfare in relation to τ_Θ are inaccurate when utilizing the histogram approach.

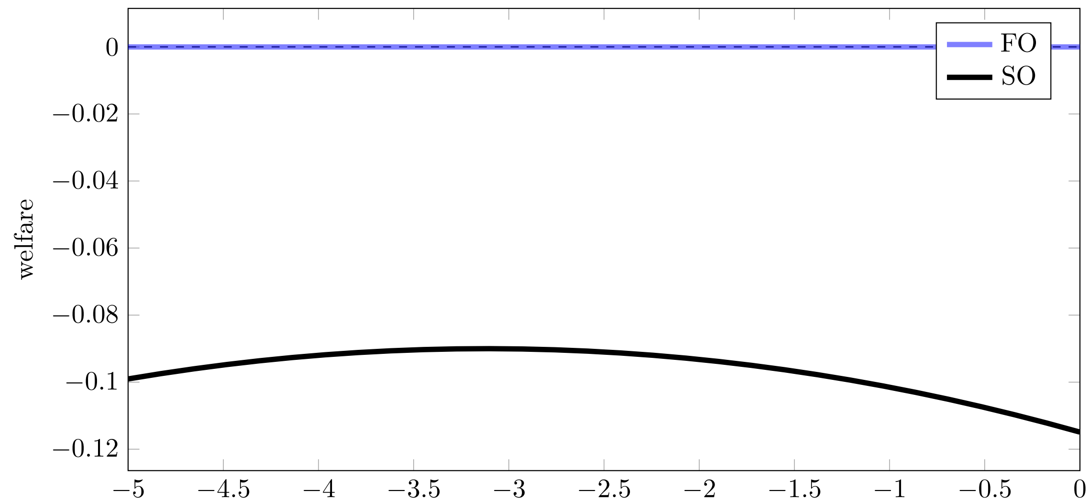

Figure 2: Welfare for various values of τ_Θ when risk aversion is set to 2.

Table 3: STABILIZATION POLICY τ_Θ^*

risk aversion	τ_Θ^*	$\frac{\mathcal{W}^{\text{hist}}(\tau_\Theta^*)}{\mathcal{W}(\tau_\Theta^*)}$	$\frac{\tau_\Theta^{*,\text{hist}}}{\tau_\Theta^*}$
2	-3.10	-348%	161%
3	-1.90	-230%	209%
4	-1.03	-226%	167%
5	-0.69	-217%	125%
7	-0.52	-187%	67%

Notes: Optimal τ_Θ as we vary the risk aversion parameter. The $\mathcal{W}^{\text{hist}}(\tau_\Theta^*)$ uses the histogram method to compute the welfare and $\tau_\Theta^{*,\text{hist}}$ is the optimal policy using $\mathcal{W}^{\text{hist}}(\tau_\Theta)$ as the measure of welfare

7.2.2 Stochastic Volatility

We next use techniques from Section 5.2 to study aggregate and distributional consequences of changes in macroeconomic risk. To do that, we first extend the baseline model to include equations (46)–(47) as the new process for aggregate shocks. We use impulse responses to study the effect of uncertainty of quantities such as aggregate capital and welfare.

Define the impulse response of aggregate X as

$$IRF_k^\Upsilon(\{\mathcal{E}_{\Upsilon,t}\}) = \mathbb{E}_t\left[X_{t+k}\middle|\mathcal{E}_{\Upsilon,t}\right] - \mathbb{E}\left[X_{t+k}|\mathcal{E}_{\Upsilon,t}=0\right].$$

Applying Lemma 1^{SV} and Proposition 1^{SV} we find that

$$IRF_k^\Upsilon(\{\mathcal{E}_{\Upsilon,t}\}) \approx \frac{1}{2}\left(\sum_{j=0}^{k-1}\overline{X}_{ZZ,j,j}\rho_\Upsilon^{k-1-j}\sigma_\Theta^2 + \sum_{j=0}^{k}\bar{X}_{\sigma\sigma,j}^\Upsilon\rho_\Upsilon^{k-j}\right)\mathcal{E}_{\Upsilon,t}. \tag{75}$$

We have already discussed the computations of the terms $\{\overline{X}_{ZZ,j,j}\}$. The sequence $\{\bar{X}_{\sigma\sigma,j}^\Upsilon\}_j$ is computed using the linear system (51), in which most terms are precomputed with the only computationally

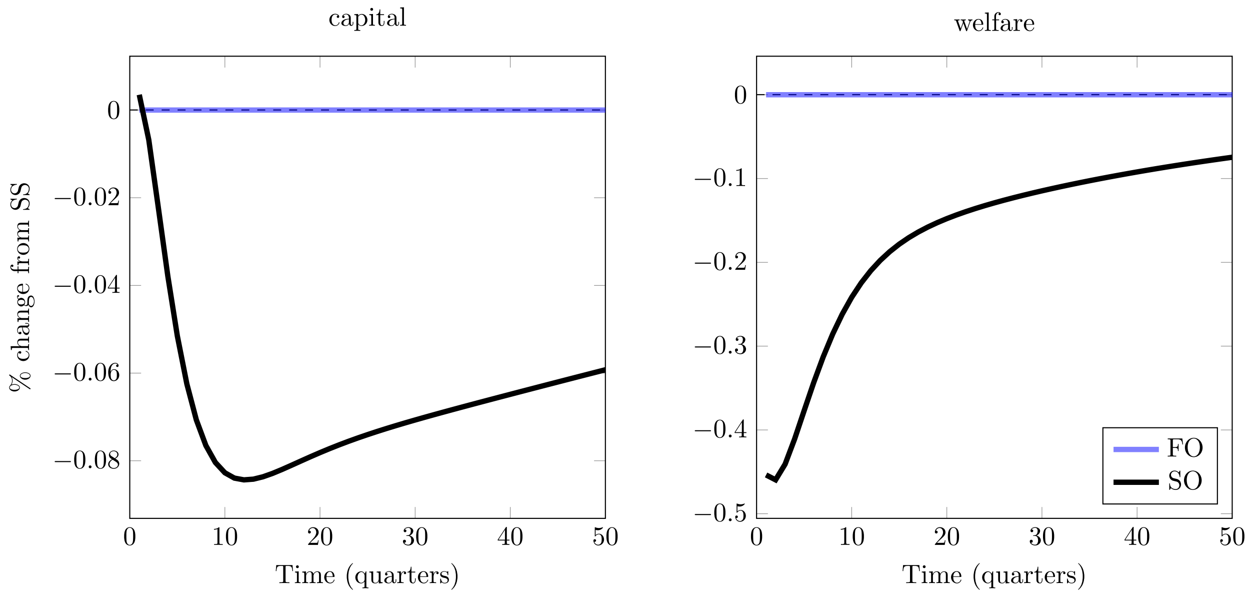

Figure 3: Impact of uncertainty shock

intensive step is obtaining function $\mathsf{x}^{\Upsilon}_{\sigma\sigma}(z,\theta)$ from equation (49). This takes an additional second of computational time. When individual Bellman equations are a part of x, the corresponding component of term $\mathsf{x}^{\Upsilon}_{\sigma\sigma}(z,\theta)$ captures effect on $t=0$ welfare $v_0(z,\theta)$, and can be conveniently used to compute the distributional effects of uncertainty with no additional time.

To illustrate the responses in the context of the Krusell and Smith baseline, we need to pick a value ρ_{Υ} and the impulse $\{\mathcal{E}_{\Upsilon,t}\}$. To do that, we use the CBOE Volatility Index (VIX), which shows large fluctuations, with rapid increases of up to 4-5 times the average (for instance in 2008 and also during the COVID pandemic) that take a couple of years to mean revert. Interpreting the VIX as a measure of uncertainty, we study a impulse response to a one-time, large but transitory shock to the uncertainty of the TFP process, similar to what we saw in the recent crisis. The shock increases the standard deviation by a factor of 5 and mean reverts with a persistence of 0.75. In Figure 3, we show the response of aggregate capital and welfare (convert the magnitudes into certainty equivalents) following the shock. We observe that the shock leads to a decrease in capital accumulation of about 1% and a fall in aggregate welfare of about 0.5%. In addition to the impact on aggregate variables, we investigate the effect of the shock on individual welfare. In Figure 4, we plot the welfare losses by assets, normalized by per capita GDP. The average welfare loss amounts to approximately half a percentage point of per-period consumption, and these losses range from 0.81% to 0.20% across the asset distribution. The most significant welfare losses are experienced by asset-poor agents who are closer to the borrowing constraints.

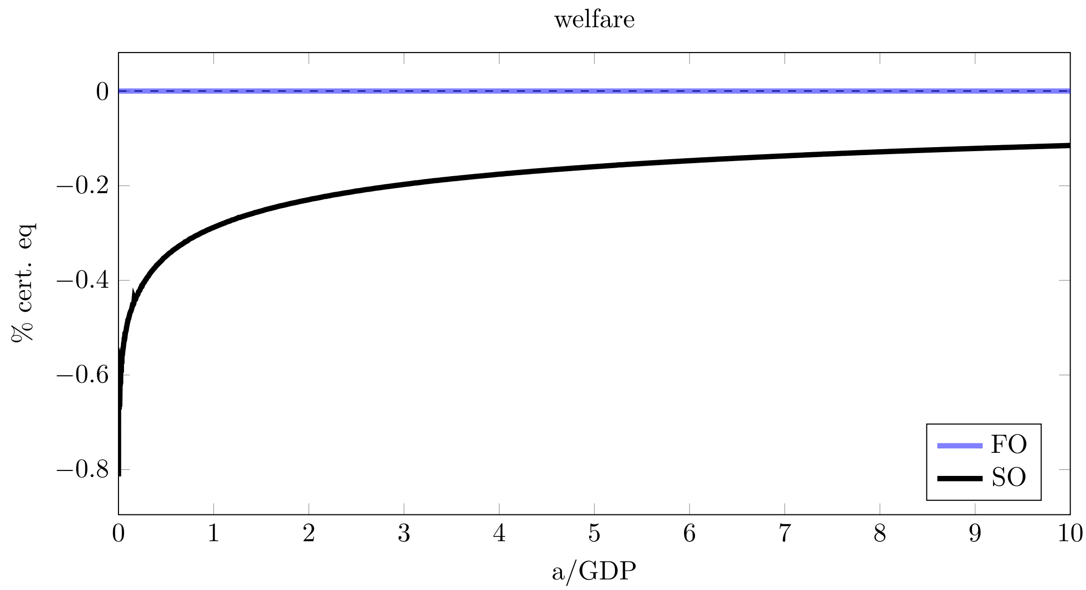

Figure 4: Distribution of per-period certainty equivalent that households forgo to avoid the one-time uncertainty shock

7.2.3 Portfolio Choice

We now illustrate the extension of our algorithm in Section 5.3 to capture portfolio choice. Extend the baseline Krusell and Smith model allowing agents to trade risk-free debt, b, which has a zero net supply, in addition to claims on risky capital whose market value we denote by k. Total wealth is $a = k + b$. We impose a constraint that prevents households from short-selling capital.

The key computational step here is to solve for the zeroth-order holdings of the risky capital $\overline{k}(\theta, a)$ and the first-order response of excess returns on capital to the contemporaneous TFP shock $R_{Z,0}$. Conditional on these two objects, the steps and the time to compute responses of other variables is the similar as in Table 2. We use a nonlinear root solver to implement equations (70) – (73) to solve for a one dimensional unknown $R_{Z,0}$. The time taken to solve for $R_{Z,0}$ naturally depends on the initial guess and tolerance/tuning parameters of the root finding algorithm. Using the initial guess from the equal-portfolio economy, it took about a minute to find the $R_{Z,0}$ using default algorithm of Julia's NLsolve package with a tolerance of $1e-11$. As a by product, we get $\{X_{Z,t}\}$ and $R_{\sigma,\sigma,0}$ from equations (70) and (71) that got used in constructing the non linear equation. We use $\{X_{Z,t}\}$ to construct the first-order responses and then use $R_{\sigma,\sigma,0}$ along with $R_{Z,0}$ to construct the portfolio function using equation (67). Both of these are linear systems and take negligible amount of additional time.

We now explore the predictions of the baseline Krusell and Smith model for the cross-sectional distribution of portfolios as well as the role of portfolios in shaping aggregate responses. In the left panel of Figure 5, we depict the distribution of household portfolios by assets normalized by per capita

GDP. The model qualitatively aligns with the observed pattern (see Yogo and Wachter (2011) who use data from the Survey of Consumer Finances) wherein poorer households hold more bonds and wealthier households hold more stocks. Households closest to the borrowing constraint are most exposed to aggregate shocks, and they optimally reduce their exposure by adjusting their portfolios towards risk-free bonds. In the right panel of Figure 5, we plot the first-order response of aggregate capital with optimal portfolio and compare it to the response if we force households to hold the same portfolios .

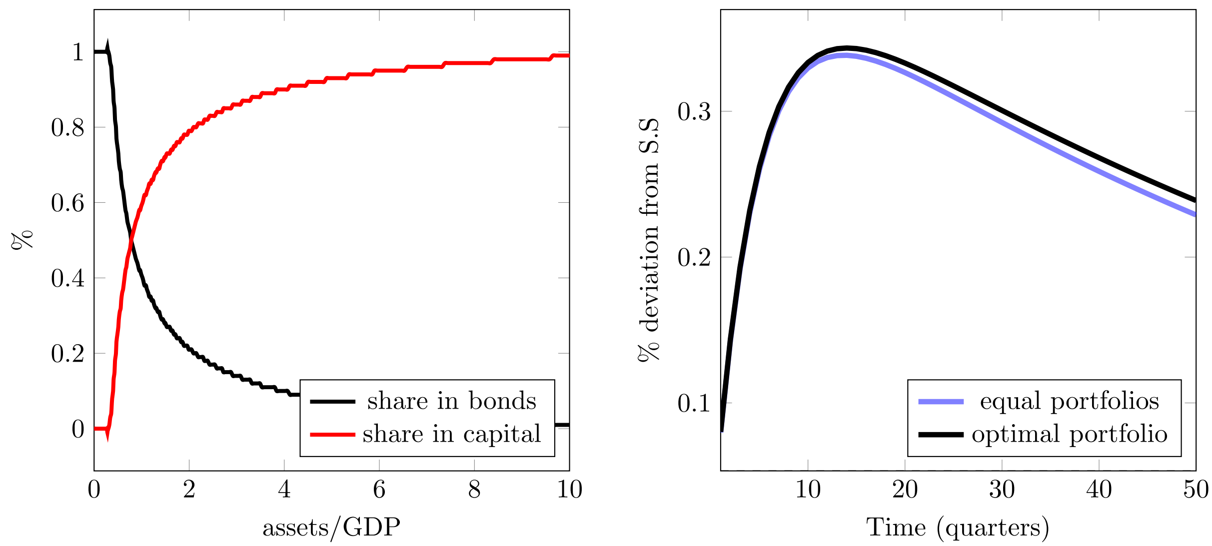

Figure 5: The left panel shows the distribution of household portfolios by assets. The right panel shows the first-order impulse response of aggregate capital.

7.2.4 Transitions

Our final application shows how to apply our method to compute deterministic transitions across two steady states. We modify the aggregate TFP process to have a parameter $\overline{\Theta}$ that controls the mean and consider a one-time permanent change of 5% to $\overline{\Theta}$. In the economy with high TFP, the distribution of savings shifts to the right to accommodate the higher demand of the capital which is now more productive.

To apply the insights from Section 5.1, we need to compute $\hat{\Omega}_0 = \Omega^* - \Omega_0$. We set the asset distribution in the non-stochastic economy with high TFP to be Ω^* and asset distribution in the non-stochastic economy with low TFP to be Ω_0. This allows us to construct the new term in Lemma 1^{TD}, $\mathcal{I} \cdot \mathcal{L}^t \cdot \frac{d}{d\theta}\hat{\Omega}_0$ which takes negligible amount of time given that we have precomputed operators $\mathcal{I}$ and $\{\mathcal{L}^t\}$. We truncate T when the difference between $\overline{X}\left(\Omega^*; \overline{\Theta} = 1\right) + \overline{X}_{Z,0}$ and $\overline{X}\left(\Omega^*; \overline{\Theta} = .95\right)$ is below a threshold.

In Figure 6, we plot the first-order expansions of the mean path of aggregate capital and the distribution of capital between the two steady states. We see capital slowly approaching a higher level and the distribution of wealth shifts rightwards.

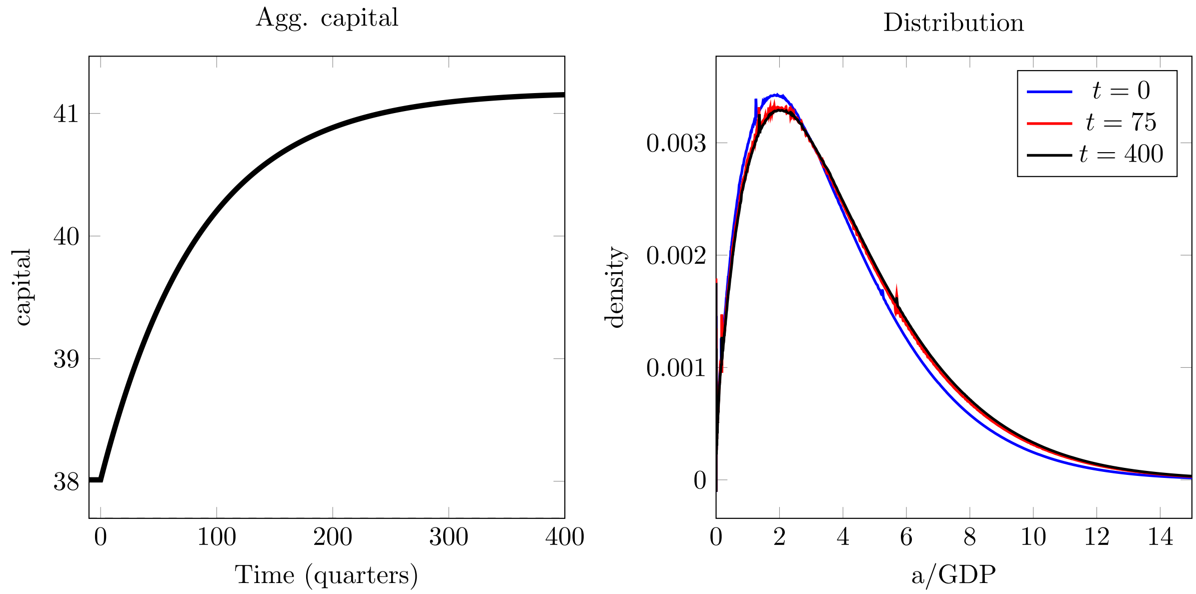

Figure 6: Transition path for capital and the distribution of savings after a 5% permanent increase in agg. TFP

8 Conclusion

In this paper, we propose a novel perturbation technique to approximate a wide variety of stochastic heterogeneous-agent (HA) models. Our methods goes beyond the MIT shock approach prevalent in existing literature by employing higher-order approximations. Utilizing Fréchet derivative techniques, we demonstrate that all-order approximations can be represented using analytically derived coefficients that are straightforward to implement numerically. Our approach broadens the range of research questions that can be addressed within these model classes. We showcase the practicality of our method by applying it to examine welfare implications of stabilization policies, portfolio choice, and time-varying uncertainty in a calibrated economy.

References

Achdou, Yves, Jiequn Han, Jean Michel Lasry, Pierre Louis Lions, and Ben Moll. 2020. "Income and Wealth Distribution in Macroeconomics: a continuous-time approach." *Review of Economic Studies.*

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

Algan, Yann, Olivier Allais, Wouter J Den Haan, and Pontus Rendahl. 2014. "Solving and Simulating Models with Heterogeneous Agents and Aggregate Uncertainty." In *Handbook of Computational Economics*. 3: Elsevier, 277–324.

Alvarez, Fernando, Andrea Ferrara, Erwan Gautier, Herve Le Bihan, and Francesco Lippi. 2023a. "Empirical Investigation of a Sufficient Statistic for Monetary Shocks."

Alvarez, Fernando, and Francesco Lippi. 2022. "The Analytic Theory of a Monetary Shock." *Econometrica*, 90(4): 1653–1682.

Alvarez, Fernando, Francesco Lippi, and Panagiotis Souganidis. 2023b. "Price Setting with Strategic Complementarities as a Mean Field Game."

Auclert, Adrien, Bence Bardóczy, Matthew Rognlie, and Ludwig Straub. 2021. "Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models." *Econometrica*, 89(5): 2375–2408.

Bayer, Christian, Benjamin Born, and Ralph Luetticke. 2022. "Shocks, Frictions, and Inequality in US Business Cycles." CEPR Discussion Papers 14364.

Bhandari, Anmol, David Evans, Mikhail Golosov, and Thomas J. Sargent. 2021. "Inequality, Business Cycles, and Monetary-Fiscal Policy." *Econometrica*, 89(6): 2559–2599.

Bigio, Saki, Galo Nuño, and Juan Passadore. 2023. "Debt-Maturity Management with Liquidity Costs."

Bilal, Adrien. 2022. "Solving Heterogeneous Agent Models with the Master Equation."Technical report.

Bilal, Adrien. 2023. "Solving Heterogeneous Agent Models with the Master Equation."

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

Boppart, Timo, Per Krusell, and Kurt Mitman. 2018. "Exploiting MIT Shocks in Heterogeneous-Agent Economies: The Impulse Response as a Numerical Derivative." *Journal of Economic Dynamics and Control*, 89 68–92.

Childers, David. 2018. "On the Solution and Application of Rational Expectations Models with Function-Valued States."Technical report.

Childers, David, Jesús Fernández-Villaverde, Jesse Perla, Christopher Rackauckas, and Peifan Wu. 2022. "Differentiable State-Space Models and Hamiltonian Monte Carlo Estimation."Technical report, National Bureau of Economic Research.

Coeurdacier, Nicolas, Helene Rey, and Pablo Winant. 2011. "The risky steady state." *American Economic Review*, 101(3): 398–401.

Devereux, Michael B., and Alan Sutherland. 2011. "Country Portfolios in Open Economy Macro-Models." *Journal of the European Economic Association*, 9(2): 337–369.

Fernández-Villaverde, Jesús, Pablo Guerrón-Quintana, Juan F. Rubio-Ramírez, and Martin Uribe. 2011. "Risk Matters: The Real Effects of Volatility Shocks." *American Economic Review*, 101(6): 2530–61.

Gornemann, Nils, Keith Kuester, and Makoto Nakajima. 2021. "Doves for the Rich, Hawks for the Poor? Distributional Consequences of Systematic Monetary Policy."

Han, Jiequn, Yucheng Yang et al. 2021. "Deepham: A global solution method for heterogeneous agent models with aggregate shocks." *arXiv preprint arXiv:2112.14377*.

Jin, He-Hui, and Kenneth L Judd. 2002. "Perturbation Methods for General Dynamic Stochastic Models."Technical report.

Judd, Kenneth L. 1998. *Numerical Methods in Economics*. MIT Press Books: The MIT Press.

Kahou, Mahdi Ebrahimi, Jesús Fernández-Villaverde, Jesse Perla, and Arnav Sood. 2021. "Exploiting symmetry in high-dimensional dynamic programming."Technical report, National Bureau of Economic Research.

Kanwal, Ram P. 1998. *Generalized functions: Theory and technique.*: Springer Science & Business Media.

Kaplan, Greg, Benjamin Moll, and Giovanni L Violante. 2018. "Monetary Policy according to HANK." *American Economic Review*, 108(3): 697–743.

Kim, Jinill, Sunghyun Kim, Ernst Schaumburg, and Christopher A Sims. 2008. "Calculating and using Second-Order Accurate Solutions of discrete-time Dynamic Equilibrium Models." *Journal of Economic Dynamics and Control*, 32(11): 3397–3414.

Krusell, Per, and Anthony A Smith, Jr. 1998. "Income and Wealth Heterogeneity in the Macroeconomy." *Journal of Political Economy*, 106(5): 867–896.

Luenberger, David G. 1997. *Optimization by Vector Space Methods*. New York, NY, USA: John Wiley & Sons.

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

Reiter, Michael. 2009. "Solving Heterogeneous-Agent Models by Projection and Perturbation." *Journal of Economic Dynamics and Control*, 33(3): 649–665.

Schmitt-Grohé, Stephanie, and Martın Uribe. 2004. "Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function." *Journal of Economic Dynamics and Control*, 28(4): 755–775.

Viceira, Luis M. 2001. "Optimal Portfolio Choice for Long Horizon Investors with Nontradable Labor Income." *Journal of Finance*, 56(2): 433–470.

Winberry, Thomas. 2018. "A method for solving and estimating heterogeneous agent macro models." *Journal of Economic Dynamics and Control*, 89 68–92.

Yogo, Motohiro, and Jessica A. Wachter. 2011. "Why Do Household Portfolio Shares Rise in Wealth?" *Review of Financial Studies*, 24(11): 3929–3965.

Young, Eric R. 2010. "Solving the Incomplete Markets Model with Aggregate Uncertainty using the Krusell–Smith Algorithm and Non-Stochastic Simulations." *Journal of Economic Dynamics and Control*, 34(1): 36–41.

A Section 4 Proofs

A.1 Derivatives of Kinks and Generalized Functions

Assumption 1(a) states that the policy rules $\tilde{x}(a,\theta,Z;\sigma)$ are smooth everywhere except for the locations $\tilde{a}^{\vee,j}(\theta,Z;\sigma)$. We let $N^{\vee}$ represent the number of kinks. As such the classical derivatives, e.g. $\mathring{\tilde{x}}_{aa}(a,\theta)$, are not defined at those kinks, and for the purposes of integration, we represent them as generalized functions. We will find it convenient to use the notation $\overline{x}^{\Delta,j}(\theta) \equiv \lim_{a\downarrow\overline{a}^{\vee,j}(\theta)} \overline{x}(a,\theta) - \lim_{a\uparrow\overline{a}^{\vee,j}(\theta)} \overline{x}(a,\theta)$ to represent the size of the discontinuity at the kink. Fore conciseness, we'll define the upper and lower limits w.r.t. a as $\overline{x}^{+}(a,\theta) = \lim_{h\downarrow 0} \overline{x}(a+h,\theta)$ and $\overline{x}^{-}(a,\theta) = \lim_{h\uparrow 0} \overline{x}(a+h,\theta)$ respectively. Continuity of the policy rules implies that $\overline{x}^{\Delta,j}(\theta) = \overline{x}^{+}(\overline{a}^{\vee,j}(\theta),\theta) - \overline{x}^{-}(\overline{a}^{\vee,j}(\theta),\theta) = 0$, but the derivatives themselves are allowed to to be discontinuous at the kinks: $\overline{x}_a^{\Delta,j}(\theta) \neq 0$.

Before formally studying the distributional derivatives, it is necessary to understand how the kinks themselves respond to the shocks. Continuity of the policy rules allows us to get the following relationship the derivative of the kink, $\overline{a}_{Z,t}^{\vee,j}(\theta)$, and the size of the discontinuity of the derivative of the policy rules at that kink, $\overline{x}_{Z,t}^{\Delta,j}(\theta)$.

Lemma 1. *For all t, the derivatives of the kinks satisfy*

$$\overline{x}_a^{\Delta}(\theta)\overline{a}_{j,Z,t}^{\vee}(\theta) = -\overline{x}_{Z,t}^{\Delta,j}(\theta),$$

and, in particular,

$$\overline{a}_{j,Z,t}^{\vee}(\theta) = -\overline{a}_a^{\Delta}(\theta)^{-1}\overline{a}_{Z,t}^{\Delta,j}(\theta). \tag{76}$$

Proof. Continuity implies that

$$\tilde{x}^{+}(\tilde{a}^{\vee,j}(\theta,Z),\theta,Z) = \tilde{x}^{-}(\tilde{a}^{\vee,j}(\theta,Z)+h,\theta,Z).$$

Differentiating with respect to Z in direction $\hat{Z}_t$ at $\sigma = 0$ yields

$$\overline{x}_a^{+}(\overline{a}^{\vee,j}(\theta),\theta)\overline{a}_{Z,t}^{\vee,j}(\theta) + \overline{x}_{Z,t}^{+}(\overline{a}^{\vee,j}(\theta),\theta) = \overline{x}_a^{-}(\overline{a}^{\vee,j}(\theta),\theta)\overline{a}_{Z,t}^{\vee,j}(\theta) + \overline{x}_{Z,t}^{-}(\overline{a}^{\vee,j}(\theta),\theta),$$

which implies that

$$\overline{x}_a^{\Delta}(\theta)\overline{a}_{j,Z,t}^{\vee}(\theta) = -\overline{x}_{Z,t}^{\Delta,j}(\theta).$$

Applying p to both sides and dividing by $\overline{a}_a^{\Delta}(\theta)$ yields (76). □

The distributional derivates themselves are defined by how the operate as linear functionals over a space of smooth test functions, φ, with compact support. We use these definitions to establish the following relationships

Claim 1. For all t, k distributional derivatives of $\tilde{x}$ satisfy

$$\begin{aligned}
\overline{x}_{Z,t}(a,\theta) &= \mathring{\overline{x}}_{Z,t}(a,\theta) \\
\overline{x}_{a}(a,\theta) &= \mathring{\overline{x}}_{a}(a,\theta) \\
\overline{x}_{aa}(a,\theta) &= \mathring{\overline{x}}_{aa}(a,\theta) + \sum_j \overline{x}_a^{\Delta,j}(\theta)\delta(a - \overline{a}^{\vee,j}(\theta)) \\
\overline{x}_{aZ,t}(a,\theta) &= \mathring{\overline{x}}_{aZ,t}(a,\theta) + \sum_j \overline{x}_{Z,t}^{\Delta,j}(\theta)\delta(a - \overline{a}^{\vee,j}(\theta)) \\
\overline{x}_{ZZ,t,k}(a,\theta) &= \mathring{\overline{x}}_{ZZ,t,k}(a,\theta) + \sum_j \overline{x}_a^{\Delta,j}(\theta)\overline{a}_{Z,t}^{\vee,j}(\theta)\overline{a}_{Z,k}^{\vee,j}(\theta)\delta(a - \overline{a}^{\vee,j}(\theta))
\end{aligned}$$

Proof. The distribution derivative $\overline{x}_a(a,\theta)$ is defined by[27]

$$\begin{aligned}
\iint \overline{x}_a(a,\theta)\varphi(a,\theta)dad\theta &= -\iint \overline{x}(a,\theta)\varphi_a(a,\theta)dad\theta \\
&= -\int \sum_{j=0}^{N^\vee} \int_{\overline{a}^{\vee,j}(\theta)}^{\overline{a}^{\vee,j+1}(\theta)} \overline{x}(a,\theta)\varphi_a(a,\theta)dad\theta
\end{aligned}$$

for any test function φ. On each of these intervals the functions are smooth so we can apply integration by parts to get

$$\iint \overline{x}_a(a,\theta)\varphi(a,\theta)dad\theta = \iint \mathring{\overline{x}}_a(a,\theta)\varphi(a,\theta)dad\theta + \int \sum_{j=1}^{N^\vee} \overline{x}^{\Delta,j}(\theta)\varphi(\overline{a}^{\vee,j}(\theta),\theta)d\theta = \iint \mathring{\overline{x}}_a(a,\theta)\phi(a,\theta)dad\theta$$

where the last equality used continuity. This implies $\overline{x}_a(a,\theta) = \mathring{\overline{x}}_a(a,\theta)$.

Next we turn to $\overline{x}_{aa}(a,\theta)$, which is defined by

$$\iint \overline{x}_{aa}(a,\theta)\varphi(a,\theta)dad\theta = -\iint \overline{x}_a(a,\theta)\varphi_a(a,\theta)dad\theta = -\iint \mathring{\overline{x}}_a(a,\theta)\varphi_a(a,\theta)dad\theta$$

Splitting up the integral over a we have

$$\begin{aligned}
\iint \overline{x}_{aa}(a,\theta)\varphi(a,\theta)dad\theta &= -\int \sum_{j=0}^{N^\vee} \int_{\overline{a}^{\vee,j}(\theta)}^{\overline{a}^{\vee,j+1}(\theta)} \mathring{\overline{x}}_a(a,\theta)\varphi_a(a,\theta)dad\theta \\
&= \iint \mathring{\overline{x}}_{aa}(a,\theta)\varphi(a,\theta)dad\theta + \int \sum_{j=1}^{N^\vee} \overline{x}_a^{\Delta,j}(\theta)\varphi(\overline{a}^{\vee,j}(\theta),\theta)d\theta \\
&= \iint \left(\mathring{\overline{x}}_{aa}(a,\theta) + \sum_{j=1}^{N^\vee} \overline{x}_a^{\Delta,j}(\theta)\delta(a - \overline{a}^{\vee,j}(\theta)) \right) \varphi(a,\theta)dad\theta
\end{aligned}$$

where the second equality used integration by parts.

Next we have $\overline{x}_{Z,t}(a,\theta)$ which is defined by

$$\int \overline{x}_{Z,t}(a,\theta)\varphi(a,\theta)dad\theta = \left(\iint \overline{x}(a,\theta,Z)\varphi(a,\theta)dad\theta \right)_{Z,t}.$$

[27] To concisely represent these integrals we use the convention that $\overline{a}^{\vee,0}(\theta) = -\infty$ and $\overline{a}^{\vee,N^\vee+1}(\theta) = \infty$

As

$$\iint \overline{x}(a,\theta,Z)\varphi(a,\theta)dad\theta = \int \sum_{j=0}^{N^\vee} \int_{\overline{a}^{\vee,j}(\theta,Z)}^{\overline{a}^{\vee,j+1`}(\theta,Z)} \overline{x}(a,\theta,Z)\varphi(a,\theta)dad\theta,$$

when we take the derivative we get

$$\begin{aligned}\left(\iint \overline{x}(a,\theta,Z)\varphi(a,\theta)dad\theta\right)_{Z,t} &= \iint \mathring{\overline{x}}_{Z,t}(a,\theta)\varphi(a,\theta)dad\theta - \int \sum_{j=1}^{N^\vee} \overline{x}^{\Delta,j}(\theta)\overline{a}^\vee_{Z,t}(\theta)\varphi(\overline{a}^{\vee,j}(\theta),\theta)d\theta \\ &= \iint \overline{x}_{Z,t}(a,\theta)\varphi(a,\theta)dad\theta\end{aligned}$$

which implies that $\overline{x}_{Z,t}(a,\theta) = \mathring{\overline{x}}_{Z,t}(a,\theta)$.

The distributional derivative $\overline{x}_{aZ,t}(a,\theta)$ is defined by

$$\iint \overline{x}_{aZ,t}(a,\theta)\varphi(a,\theta)dad\theta = -\iint \overline{x}_{Z,t}(a,\theta)\varphi_a(a,\theta)dad\theta$$

for any test function φ. As $\overline{x}_{Z,t}(a,\theta) = \mathring{\overline{x}}_{Z,t}(a,\theta)$ we have

$$\begin{aligned}-\iint \overline{x}_{Z,t}(a,\theta)\varphi_a(a,\theta)dad\theta &= -\int \sum_{j=0}^{N^\vee} \int_{\overline{a}^{\vee,j}(\theta)}^{\overline{a}^{\vee,j+1`}(\theta)} \mathring{\overline{x}}_{Z,t}(a,\theta)\varphi_a(a,\theta)dad\theta \\ &= \iint \mathring{\overline{x}}_{aZ,t}(a,\theta)\varphi(a,\theta)dad\theta + \int \sum_{j=1}^{N^\vee} \overline{x}^{\Delta,j}_{Z,t}(\theta)\varphi(\overline{a}^{\vee,j}(\theta),\theta)d\theta \\ &= \iint \left(\mathring{\overline{x}}_{aZ,t}(a,\theta) + \sum_{j=1}^{N^\vee} \overline{x}^{\Delta,j}_{Z,t}(\theta)\delta(a-\overline{a}^{\vee,j}(\theta))\right)\varphi(a,\theta)dad\theta.\end{aligned}$$

We conclude $\overline{x}_{aZ,t}(a,\theta) = \mathring{\overline{x}}_{aZ,t}(a,\theta) + \sum_j \overline{x}^{\Delta,j}_{Z,t}(\theta)\delta(a-\overline{a}^{\vee,j}(\theta))$. Finally the distributional derivative, $\dot{\overline{x}}_{ZZ,t,k}(a,\theta)$, is defined by

$$\int \dot{\overline{x}}_{ZZ,t,k}(a,\theta)\varphi(a,\theta)dad\theta = \left(\iint \overline{x}(a,\theta,Z)\varphi(a,\theta)dad\theta\right)_{ZZ,t,k}$$

As

$$\begin{aligned}\left(\iint \overline{x}(a,\theta,Z)\varphi(a,\theta)dad\theta\right)_{ZZ,t,k} &= \left(\iint \mathring{\overline{x}}_{Z,t}\left(a,\theta,Z\right)\varphi(a,\theta)dad\theta\right)_{Z,k} + \left(\iint \overline{x}(a,\theta,Z)\varphi(a,\theta)dad\theta\right)_Z \cdot \hat{Z}_{t,k} \\ &= \left(\int \sum_{j=0}^{N^\vee} \int_{\overline{a}^{\vee,j}(\theta,Z)}^{\overline{a}^{\vee,j+1`}(\theta,Z)} \mathring{\overline{x}}_{Z,t}(a,\theta,Z)\varphi(a,\theta)dad\theta\right)_{Z,k} + \iint \mathring{\overline{x}}_Z(a,\theta)\cdot\hat{Z}_{t,k}\varphi(a,\theta)dad\theta \\ &= \iint \mathring{\overline{x}}_{ZZ,t,k}(a,\theta)\varphi(a,\theta)dad\theta - \int \sum_{j=1}^{N^\vee} \overline{x}^{\Delta,j}_{Z,t}\left(\theta\right)\overline{a}^{\vee,j}_{Z,k}(\theta)\varphi\left(\overline{a}^{\vee,j}(\theta),\theta\right)d\theta \\ &= \iint \left(\mathring{\overline{x}}_{ZZ,t,k}(a,\theta) + \sum_{j=1}^{N^\vee} \overline{x}^{\Delta,j}_a(\theta)\overline{a}^{\vee,j}_{Z,t}(\theta)\overline{a}^{\vee,j}_{Z,k}(\theta)\delta\left(a-\overline{a}^{\vee,j}(\theta)\right)\right)\varphi(a,\theta)dad\theta\end{aligned}$$

where the last time used the relationship $-\overline{x}^{\Delta,j}_a(\theta)\overline{a}^{\vee,j}_{Z,t}(\theta) = \overline{x}^{\Delta,j}_{Z,t}(\theta)$ from the above. Thus

$$\overline{x}_{ZZ,t,k}(a,\theta) = \mathring{\overline{x}}_{ZZ,t,k}(a,\theta) + \sum_j \overline{x}^{\Delta,j}_a(\theta)\overline{a}^{\vee,j}_{Z,t}(\theta)\overline{a}^{\vee,j}_{Z,k}(\theta)\delta\left(a-\overline{a}^{\vee,j}(\theta)\right).$$

□

The distributional derivatives in Claim 1 provide a succinct way to summarize how changes in the location of the kink affect derivatives of integrals over individual policies. To keep the the analysis in this appendix as accessible as possible we'll derive all our main results without explicitly the generalized derivatives of $\overline{x}$. Instead, we will explicitly track the limits of integration and only summarize our results at the end using these δ-functions. As the distributional and classical derivatives align to first order we will use $\overline{x}_{Z,t}$ and $\mathring{\overline{x}}_{Z,t}$ interchangeably. We will only explicitly emphasize the classical derivative at second order.

Finally, we want to highlight an important feature of these additional terms that arise from kinked policy functions. Namely, they can always be determined from lower order derivatives. We see this in all of the generalized second derivatives, who's δ-function components depend only on first derivatives. This implies that all of the δ function components in the second order derivates can be determined before the classical second order derivatives, e.g. $\mathring{\overline{x}}_{ZZ,t,k}$, are found.

A.2 Proof of Lemma ??

Taking a first-order derivative of (16) and (15) around the $\sigma = 0$ steady state yields $\overline{Z}_{0,\sigma}(\mathcal{E}^0) = \hat{Z}_0\mathcal{E}_0$ and, for $t \geq 0$,

$$\overline{Z}_{t+1,\sigma}\left(\mathcal{E}^{t+1}\right) = \overline{Z}_Z \cdot \overline{Z}_{t,\sigma}\left(\mathcal{E}^{t-1}\right) + \hat{Z}_0\mathcal{E}_t + \overline{Z}_\sigma \tag{77}$$

$$\overline{X}_{t,\sigma}\left(\mathcal{E}^{t}\right) = \overline{X}_Z \cdot \overline{Z}_{t,\sigma}\left(\mathcal{E}^{t}\right) + \overline{X}_\sigma, \tag{78}$$

with $\hat{Z}_0$ and $\overline{Z}_Z$ being defined in the main text and $\overline{Z}_\sigma := \left[0, \overline{\Omega}_\sigma\right]$. Our first step is to show that $\overline{X}_\sigma$ and $\overline{Z}_\sigma$ are both 0 which we codify in the following claim

Claim 2. The first derivatives with respect to σ, $\left(\overline{X}_\sigma, \overline{\Omega}_\sigma, \overline{x}_\sigma\right)$, are all 0.

Proof. Differentiating the F, G, and LoM mappings with respect to σ yields the following system of equations[28]

$$0 = \mathsf{F}_x(a,\theta)\overline{x}_\sigma(a,\theta) + \mathsf{F}_X(a,\theta)\overline{X}_\sigma + \mathsf{F}_{x'}(a,\theta)\left(\mathbb{E}\left[\overline{x}_a|a,\theta\right]\overline{a}_\sigma(a,\theta) + \mathbb{E}\left[\overline{x}_\sigma|a,\theta\right] + \mathbb{E}\left[\overline{x}_Z \cdot \overline{Z}_\sigma|a,\theta\right]\right)$$

$$0 = \mathsf{G}_x \int \overline{x}_\sigma d\Omega^* + \mathsf{G}_X\overline{X}_\sigma$$

$$\overline{\Omega}_\sigma\langle a', \theta'\rangle = -\iint \delta\left(\overline{a}(a,\theta) - a'\right)\iota(\rho_\theta\theta + \epsilon \leq \theta')\overline{a}_\sigma(a,\theta)\mu(\epsilon)d\epsilon d\Omega^*.$$

This system of equations is homogeneous of degree 1 in $\left(\overline{X}_\sigma, \overline{\Omega}_\sigma, \overline{x}_\sigma\right)$ and, therefore, is solved by setting all terms to zero. □

[28]Here we have exploited the knowledge that $\mathbb{E}\left[\overline{x}_\Theta(a',\theta')\mathcal{E}'\right] = 0$

Next we show the following claim relating $\overline{Z}_{t,\sigma}\left(\mathcal{E}^t\right)$ to the directions $\hat{Z}_t$

Claim 3. For all t

$$\overline{Z}_{t,\sigma}\left(\mathcal{E}^t\right)=\sum_{s=0}^{t}\hat{Z}_{t-s}\mathcal{E}_s. \tag{79}$$

Proof. We proceed via induction as $\overline{Z}_{0,\sigma}(\mathcal{E}^0)=\hat{Z}_0\mathcal{E}_0$ implies equation (79) holds for $t=0$. Assuming (79) holds for $t-1$ we have

$$\begin{aligned}\overline{Z}_{t,\sigma}\left(\mathcal{E}^t\right)&=\overline{Z}_Z\cdot\left(\sum_{s=0}^{t-1}\hat{Z}_{t-1-s}\mathcal{E}_s\right)+\hat{Z}_0\mathcal{E}_0\\&=\left(\sum_{s=0}^{t-1}\hat{Z}_{t-s}\mathcal{E}_s\right)+\hat{Z}_0\mathcal{E}_0\\&=\sum_{s=0}^{t}\hat{Z}_{t-s}\mathcal{E}_s,\end{aligned}$$

where in the second line we used $\hat{Z}_{k+1}\equiv\overline{Z}_Z\cdot\hat{Z}_k$. □

Finally, substituting for $\overline{Z}_{t,\sigma}\left(\mathcal{E}^t\right)$ in (15) yields

$$\overline{X}_{t,\sigma}\left(\mathcal{E}^t\right)=\overline{X}_Z\cdot\left(\sum_{s=0}^{t}\hat{Z}_{t-s}\mathcal{E}_s\right)=\sum_{s=0}^{t}\overline{X}_{Z,t-s}\mathcal{E}_s$$

which completes the proof.

A.3 Derivations of Equations (17), (18), and (19)

Differentiating the G mapping, equation (10), in direction $\hat{Z}_t=[\rho_\Theta^t,\overline{A}_{Z,t-1},\hat{\Omega}_t]^\intercal$ is equivalent to differentiating

$$G\left(\int\overline{x}\left(a,\theta,Z^*+\alpha\hat{Z}_t\right)d\left(\Omega^*+\alpha\hat{\Omega}_t\right),\overline{Y}(Z^*+\alpha\hat{Z}_t)\right)=0$$

w.r.t. α. Doing so yields

$$\mathsf{G}_x\underbrace{\left(\left(\int\overline{x}d\Omega^*\right)_{Z,t}+\int\overline{x}d\hat{\Omega}_t\right)}_{\left(\int\overline{x}d\Omega\right)_{Z,t}}+\mathsf{G}_Y\overline{Y}_{Z,t}=0.$$

Writing out the integration explicitly, we have

$$\begin{aligned}\left(\int\overline{x}d\Omega^*\right)_{Z,t}&=\left(\int\sum_{j=0}^{N^\vee}\int_{\overline{a}^{\vee,j}(\theta,Z)}^{\overline{a}^{\vee,j+1}(\theta,Z)}\overline{x}(a,\theta,Z)\omega^*(a,\theta)dad\theta\right)_{Z,t}\\&=\iint\overline{x}_Zd\Omega^*-\int\sum_{j=1}^{N^\vee}\overline{x}^{\Delta,j}(\theta)\omega^*(\overline{a}^{\vee,j}(\theta),\theta)d\theta\\&=\iint\overline{x}_Zd\Omega^*.\end{aligned}$$

To arrive at the second equality we exploited that the components of ω^* are continuous and that the mass of agents on the kink is zero. Finally, that the mass of agents on the kink is zero implies that $\iint \overline{x}_Z d\Omega^*$ is well defined.

Differentiating the of $\tilde{Y}$ in equation (12) in direction $\hat{Z}_t = [\rho_\Theta^t, \overline{A}_{Z,t-1}, \hat{\Omega}_t]^\intercal$ immediately gives

$$\overline{Y}_{Z,t} = \left[\rho_\Theta^t, \mathsf{P}\overline{X}_{Z,t-1}, \overline{X}_{Z,t}, \overline{X}_{Z,t+1}\right]$$

A.4 Proof of Lemma 2^{FO}

We begin by differentiating the F mapping, equation (9), in direction $\hat{Z}_t$ at a point not on the kinks. Doing so yields

$$\mathsf{F}_x(a,\theta)\overline{x}_Z(a,\theta)\cdot\hat{Z}_t + \mathsf{F}_Y(a,\theta)\overline{Y}_Z\cdot\hat{Z}_t + \mathsf{F}_{x^e}(a,\theta)\left(\mathbb{E}_\varepsilon\left[\overline{x}|a,\theta,Z\right]\right)_Z\cdot\hat{Z}_t = 0$$

where

$$\mathbb{E}_\varepsilon\left[\overline{x}|a,\theta,Z\right] = \int \overline{x}\left(\overline{a}(a,\theta,Z), \rho_\theta\theta+\varepsilon, \overline{Z}(Z)\right)\mu(\epsilon)d\epsilon.$$

Applying the derivative and exploiting continuity of $\overline{x}$

$$\mathbb{E}_\varepsilon\left[\overline{x}_a|a,\theta\right]\mathsf{p}\overline{x}_Z(a,\theta)\cdot\hat{Z}_t + \mathbb{E}_\varepsilon\left[\overline{x}_Z\cdot\overline{Z}_Z\cdot\hat{Z}_t|a,\theta\right]$$

Replacing $\overline{x}_Z\cdot\hat{Z}_t = \overline{x}_{Z,t}$, $\overline{X}_Z\cdot\hat{Z}_t = \overline{X}_{Z,t}$ and $\hat{Z}_{t+1} = \overline{Z}_Z\cdot\hat{Z}_t$ we get the difference equation

$$\mathsf{F}_x(a,\theta)\overline{x}_{Z,t}(a,\theta) + \mathsf{F}_Y(a,\theta)\overline{X}_{Z,t} + \mathsf{F}_{x^e}(a,\theta)\left(\mathbb{E}_\varepsilon\left[\overline{x}_a|a,\theta\right]\mathsf{p}\overline{x}_{Z,t}(a,\theta) + \mathbb{E}_\varepsilon\left[\overline{x}_{Z,t+1}|a,\theta\right]\right) = 0. \quad (80)$$

Our claim is that $\overline{x}_{Z,t} = \sum_{s=0}^\infty \mathsf{x}_s\overline{X}_{Z,t+s}$ solves this equation where x_s are defined via (21) and (22). To see this, note that

$$\begin{aligned}
\mathsf{F}_{x^e}(a,\theta)\mathbb{E}_\varepsilon\left[\overline{x}_{Z,t+1}|a,\theta\right] &= \sum_{s=0}^\infty \mathsf{F}_{x^e}(a,\theta)\mathbb{E}_\varepsilon\left[\mathsf{x}_s|a,\theta\right]\overline{Y}_{Z,t+1+s} \\
&= -\left(\mathsf{F}_x(a,\theta) + \mathsf{F}_{x^e}(a,\theta)\mathbb{E}_\varepsilon\left[\overline{x}_a|a,\theta\right]\mathsf{P}\right)\sum_{s=0}^\infty \mathsf{x}_{s+1}(a,\theta)\overline{Y}_{Z,t+1+s} \\
&= -\left(\mathsf{F}_x(a,\theta) + \mathsf{F}_{x^e}(a,\theta)\mathbb{E}_\varepsilon\left[\overline{x}_a|a,\theta\right]\mathsf{P}\right)\sum_{s=1}^\infty \mathsf{x}_s(a,\theta)\overline{Y}_{Z,t+s}
\end{aligned}$$

where the second line comes from applying equation (22). Combined with equation (21) we have

$$\mathsf{F}_Y(a,\theta) + \mathsf{F}_{x^e}(a,\theta)\mathbb{E}_\varepsilon\left[\overline{x}_{Z,t+1}|a,\theta\right] = -\left(\mathsf{F}_x(a,\theta) + \mathsf{F}_{x^e}(a,\theta)\mathbb{E}_\varepsilon\left[\overline{x}_a|a,\theta\right]\mathsf{P}\right)\overline{x}_{Z,t}(a,\theta)$$

which guarantees (80) and completes the proof.

A.5 Proof of Lemma 3^{FO}

Differentiating the LoM, equation 11, in direction $\hat{Z}_t$ is equivalent to differentiating

$$\overline{\Omega}(Z^*+\alpha\hat{Z}_t)\langle a',\theta'\rangle = \int \sum_{j=0}^{N^\vee} \int_{\overline{a}^{\vee,j}(\theta,Z^*+\alpha\hat{Z}_t)}^{\overline{a}^{\vee,j+1}(\theta,Z^*+\alpha\hat{Z}_t)} \int \iota\left(\bar{a}(a,\theta,Z^*+\alpha\hat{Z}_t)\le a'\right)\iota\left(\rho_\theta\theta+\epsilon\le\theta'\right)\mu(\epsilon)d\epsilon d\left(\Omega^*+\alpha\hat{\Omega}_t\right)\langle a,\theta\rangle$$

with respect to α. This yields

$$\begin{aligned}\overline{\Omega}_Z\cdot\hat{Z}_t\langle a',\theta'\rangle = &-\iint \delta\left(\overline{a}(a,\theta)-a'\right)\iota\left(\rho_\theta\theta+\epsilon\le\theta'\right)\mu(\epsilon)d\epsilon\overline{a}_Z(a,\theta)\cdot\hat{Z}_t d\Omega^*\langle a,\theta\rangle\\ &+\iint \iota\left(\bar{a}(a,\theta)\le a'\right)\iota\left(\rho_\theta\theta+\epsilon\le\theta'\right)\mu(\epsilon)d\epsilon d\hat{\Omega}_t\langle a,\theta\rangle.\\ &+\int\sum_{j=1}^{N^\vee}\int\left(\iota\left(\bar{a}^+(\overline{a}^{\vee,j}(\theta),\theta)\le a'\right)-\iota\left(\bar{a}^-(\overline{a}^{\vee,j}(\theta),\theta)\le a'\right)\right)\iota\left(\rho_\theta\theta+\epsilon\le\theta'\right)\mu(\epsilon)d\epsilon d\Omega^*\langle\overline{a}^\vee(\theta),\theta\rangle\end{aligned}$$

Continuity of $\overline{a}$ implies that this last term is 0. Applying $\frac{d}{d\theta'}$ to both sides and substituting for $\overline{a}_{Z,t}$ yields

$$\begin{aligned}\frac{d}{d\theta'}\hat{\Omega}_{t+1}\langle a',\theta'\rangle &= -\int \overbrace{\delta\left(\overline{a}(a,\theta)-a'\right)\mu(\theta^{'}-\rho_\theta\theta)}^{\overline{\Lambda}(a^{'},\theta^{'},a,\theta)}\,\overline{a}_{Z,t}(a,\theta)d\Omega^*\langle a,\theta\rangle\\ &\quad+\int \iota\left(\bar{a}(a,\theta)\le a'\right)\mu(\theta^{'}-\rho_\theta\theta)d\hat{\Omega}_t\langle a,\theta\rangle.\\ &= -\left(\mathcal{M}\cdot\overline{a}_{Z,t}\right)\langle a',\theta'\rangle\\ &\quad+\int \overbrace{\delta\left(\overline{a}(a,\theta)-a'\right)\mu(\theta^{'}-\rho_\theta\theta)}^{\overline{\Lambda}(a^{'},\theta^{'},a,\theta)}\,\overline{a}_a(a,\theta)\frac{d}{d\theta}\hat{\Omega}_t\langle a,\theta\rangle da d\theta\\ &= -\left(\mathcal{M}\cdot\overline{a}_{Z,t}\right)\langle a',\theta'\rangle+\left(\mathcal{L}^{(a)}\cdot\frac{d}{d\theta}\hat{\Omega}_t\right)\langle a',\theta'\rangle\end{aligned}$$

Where the second equality is achieved via integration by parts. To conclude, we need to show that all the integrals are well defined. We start with the following Claim

Claim 4. If y is a piecewise smooth with kinks at $\overline{a}^{\vee,j}(\theta)$ then $\mathcal{M}\cdot\mathsf{y}$ is a generalized function with a finite number of mass points a_n^*.

Proof. We will show that $\left(\mathcal{M}\cdot\mathsf{y}\right)\langle a',\theta'\rangle$ is of the form

$$\mathsf{m}(a',\theta')+\sum_n \mathsf{m}_n^\delta(\theta')\delta(a-a_n^*).$$

From our definition of $\mathcal{M}$

$$\begin{aligned}\left(\mathcal{M}\cdot\mathsf{y}\right)\langle a',\theta'\rangle &= \iint \overline{\Lambda}(a',\theta',a,\theta)\mathsf{y}(a,\theta)\left(\omega^*(a,\theta)+\sum_n\xi_n^*(\theta)\delta(a-a_n^*)\right)da d\theta\\ &= \iint \overline{\Lambda}(a',\theta',a,\theta)\mathsf{y}(a,\theta)\mathring{\omega}^*(a,\theta)da d\theta\\ &\quad+\sum_n\int\overline{\Lambda}(a',\theta',a_n^*,\theta)\mathsf{y}(a_n^*,\theta)\xi_n^*(\theta)d\theta\end{aligned}$$

For points $a^{'} \neq a_n^*$ we have

$$\mathring{\omega}^*(a',\theta') = \iint \overline{\Lambda}(a',\theta',a,\theta)\mathring{\omega}^*(a,\theta)dad\theta + \sum_n \int \overline{\Lambda}(a',\theta',a_n^*,\theta)\xi_n^*(\theta)d\theta,$$

and since y is piecewise smooth with discontinuities that don't align with the mass-points a_n^* we conclude that

$$\begin{aligned}\mathsf{m}(a',\theta') = &\iint \overline{\Lambda}(a',\theta',a,\theta)\mathsf{y}(a,\theta)\mathring{\omega}^*(a,\theta)dad\theta \\ &+ \sum_n \int \overline{\Lambda}(a',\theta',a_n^*,\theta)\mathsf{y}(a_n^*,\theta)\xi_n^*(\theta)d\theta\end{aligned}$$

exists and is continuous for all $a' \neq a_n^*$. At the mass points we have

$$\iint \overline{\Lambda}(a_n^*,\theta',a,\theta)\mathring{\omega}^*(a,\theta)dad\theta + \sum_m \int \overline{\Lambda}(a_n^*,\theta',a_m^*,\theta)\xi_n^*(\theta)d\theta = \xi_n^*(\theta')\delta(a-a_n^*)$$

where

$$\xi_n^*(\theta') = \int \int_{\overline{\theta}(a,a_n^*)} \mu\left(\theta' - \rho_\theta\theta\right)\mathring{\omega}^*(a,\theta)d\theta da + \sum_m \int_{\overline{\theta}(a_m^*,a_n^*)} \mu(\theta' - \rho_\theta\theta)\xi_n^*(\theta)d\theta$$

with $\overline{\theta}(a,a') = \{\theta : \overline{a}(a,\theta) = a'\}$. As y is piecewise smooth with discontinuities that don't align with the mass-points a_n^*

$$\begin{aligned}\mathsf{m}_{\delta,n}(\theta') = &\int \int_{\overline{\theta}(a,a_n^*)} \mu\left(\theta' - \rho_\theta\theta\right)\mathring{\omega}^*(a,\theta)\mathsf{y}\left(a,\theta\right)d\theta da \\ &+ \sum_m \int_{\overline{\theta}(a_m^*,a_n^*)} \mu(\theta' - \rho_\theta\theta)\xi_n^*(\theta)\mathsf{y}\left(a_m^*,\theta\right)d\theta\end{aligned}$$

is well defined. □

This claim implies that $\mathcal{M}\cdot\overline{a}_{Z,t}$ is a generalized function with a finite number of mass-points at a_n^*. As $\hat{\Omega}_0 = 0$ we conclude that

$$\frac{d}{d\theta}\hat{\Omega}_1 = -\mathcal{M}\cdot\overline{a}_{Z,0}$$

is a generalized function with a finite number of mass-points at a_n^*. Our next claim allows us to extend this to all $\frac{d}{d\theta}\hat{\Omega}_t$ via induction

Claim 5. If $\frac{d}{d\theta}\hat{\Omega}$ is a generalized function with a finite number of mass-points at a_n^*, then $\mathcal{L}^{(a)}\cdot\frac{\mathsf{d}}{\mathsf{d}\theta}\hat{\Omega}$ is a generalized function with a finite number of mass-points at a_n^*.

Proof. Repeat the steps of the Claim 4 replacing y with $\overline{a}_a$ and

$$\mathring{\omega}^*(a,\theta) + \sum_n \xi_n^*(\theta)\delta(a-a_n^*)$$

with $\frac{d}{d\theta}\hat{\Omega}$. □

Induction then implies that $\frac{d}{d\theta}\hat{\Omega}_t$ is a density with a finite number of mass points $\{a_n^*\}$ for the remainder of the proof we will write

$$\frac{d}{d\theta}\hat{\Omega}_t = \hat{\omega}_t(a,\theta) + \sum_n \hat{\xi}_{t,n}(\theta)\delta(a - a_n^*)$$

A.6 Proof of Corollary 1^{FO}

We start with our first claim

Claim 6. $\frac{d}{d\theta}\hat{\Omega}_t$ is given by

$$\frac{d}{d\theta}\hat{\Omega}_t = -\sum_{s=0} \mathsf{A}_{t,s}\overline{Y}_{Z,s}$$

where $\mathsf{A}_{t,s}$ is as defined in Corollary 1^{FO}.

Proof. We proceed by induction. It's trivially true from $t = 0$ as $\mathsf{A}_{0,s} = 0$ and $\frac{d}{d\theta}\hat{\Omega}_0$. We then proceed by induction by substituting for $\overline{a}_{Z,t}$

$$\begin{aligned}
\frac{d}{d\theta}\hat{\Omega}_{t+1} &= \mathcal{L}^{(a)} \cdot \frac{d}{d\theta}\hat{\Omega}_t - \sum_{j=0}^{\infty} (\mathcal{M} \cdot \mathsf{a}_j)\, \overline{Y}_{Z,t+j} \\
&= \mathcal{L}^{(a)} \cdot \left(-\sum_{s=0}^{\infty} \mathsf{A}_{t,s}\overline{Y}_{Z,s}\right) - \sum_{s=0}^{\infty} (\mathcal{M} \cdot \mathsf{a}_{s-t})\, \overline{Y}_{Z,s} \\
&= \sum_{s=0}^{\infty} - \left(\mathcal{L}^{(a)} \cdot \mathsf{A}_{t,s} + \mathcal{M} \cdot \mathsf{a}_{s-t}\right) \overline{Y}_{Z,s} \\
&\equiv \sum_{s=0}^{\infty} \mathsf{A}_{t+1,s}\overline{Y}_{Z,s}
\end{aligned}$$

where the second equality is achieved by letting $s = t + j$ and WLOG setting $\mathsf{a}_k = 0$ for $k < 0$. □

Applying integration by parts we have[29]

$$\int \overline{x} d\hat{\Omega}_t = -\iint \overline{x}_a \frac{d}{d\theta}\hat{\Omega}_t da d\theta := -\mathcal{I}^{(a)} \cdot \frac{d}{d\theta}\hat{\Omega}_t.$$

From the proof of Lemma 3^{FO} we know that $\frac{d}{d\theta}\hat{\Omega}_t$ is a density with mass points at a finite number of points a_n^* which implies that the set of points where $\overline{x}_a$ is not defined is measure zero under $\frac{d}{d\theta}\hat{\Omega}_t da d\theta$ so $\mathcal{I}^{(a)} \cdot \frac{d}{d\theta}\hat{\Omega}_t$ is well defined. Therefore

$$\int \overline{x} d\hat{\Omega}_t = -\mathcal{I}^{(a)} \cdot \left(-\sum_{s=0}^{\infty} \mathsf{A}_{t,s}\overline{Y}_{Z,s}\right) = \sum_{s=0}^{\infty} \left(\mathcal{I}^{(a)} \cdot \mathsf{A}_{t,s}\right) \overline{Y}_{Z,s}$$

[29] Again, we can formally define this as integration over the sub intervals of a where $\overline{x}$ is smooth and then apply integration by parts. Continuity of $\overline{x}$ at the kinks implies that those limit terms drop out.

We conclude by substituting for $\overline{x}_{Z,t}$ and $\int \overline{x} d\hat{\Omega}_t$ in equation (19) to conclude that

$$\left(\int \overline{x} d\Omega\right)_{Z,t} = \sum_{j=0}^{\infty} \int \mathsf{x}_j d\Omega^* \overline{Y}_{Z,t+s} + \sum_{s=0}^{\infty} \left(\mathcal{I}^{(a)} \cdot \mathsf{A}_{t,s}\right) \overline{Y}_{Z,s}$$
$$= \sum_{s=0}^{\infty} \underbrace{\left(\int \mathsf{x}_{t-s} d\Omega^* + \mathcal{I}^{(a)} \cdot \mathsf{A}_{t,s}\right)}_{\mathsf{J}_{t,s}} \overline{Y}_{Z,s}$$

as desired.

A.7 Proof Of Proposition 1^{FO}

This is a direct result of combining Corollary 1^{FO} with equation (17).

A.8 Proof of Lemma 1^{SO}

We proceed by taking a second-order derivatives of (16) and (15) w.r.t. σ to find $\overline{Z}_{0,\sigma\sigma}\left(\mathcal{E}^0\right) = 0$ and [30]

$$\overline{Z}_{t+1,\sigma\sigma}\left(\mathcal{E}^{t+1}\right) = \overline{Z}_Z \cdot \overline{Z}_{t,\sigma\sigma}\left(\mathcal{E}^t\right) + \overline{Z}_{ZZ} \cdot \left(\overline{Z}_{t,\sigma}\left(\mathcal{E}^t\right), \overline{Z}_{t,\sigma}\left(\mathcal{E}^t\right)\right) + \overline{Z}_{\sigma\sigma} \tag{81}$$

$$\overline{X}_{t,\sigma\sigma}\left(\mathcal{E}^t\right) = \overline{X}_Z \cdot \overline{Z}_{t,\sigma\sigma}\left(\mathcal{E}^t\right) + \overline{X}_{ZZ} \cdot \left(\overline{Z}_{t,\sigma}\left(\mathcal{E}^t\right), \overline{Z}_{t,\sigma}\left(\mathcal{E}^t\right)\right) + \overline{X}_{\sigma\sigma} \tag{82}$$

where $\overline{Z}_{ZZ}$ is defined in the main text and $\overline{Z}_{\sigma\sigma} = \left[0, \overline{\Omega}_{\sigma\sigma}\right]^T$. We begin by showing the following claim relating $\overline{Z}_{t,\sigma\sigma}\left(\mathcal{E}^t\right)$ to the directions $\hat{Z}_{t,k}$ and $\hat{Z}_{\sigma\sigma,t}$.

Claim 7. For all t

$$\overline{Z}_{t,\sigma\sigma}\left(\mathcal{E}^t\right) = \hat{Z}_{\sigma\sigma,t} + \sum_{s=0}^{t}\sum_{m=0}^{t} \hat{Z}_{t-s,t-m}\mathcal{E}_s\mathcal{E}_m \tag{83}$$

Proof. We proceed by induction. As $\hat{Z}_{\sigma\sigma,0} = \hat{Z}_{0,0} = 0$ we conclude that equation (15) holds for $t = 0$ since $\overline{Z}_{0,\sigma\sigma}\left(\mathcal{E}^0\right) = 0$. Assuming (83) holds for $t-1$ we have

$$\begin{aligned}
\overline{Z}_{t,\sigma\sigma}\left(\mathcal{E}^t\right) &= \overline{Z}_Z \cdot \left(\hat{Z}_{\sigma\sigma,t-1} + \sum_{s=0}^{t-1}\sum_{m=0}^{t-1} \hat{Z}_{t-1-s,t-1-m}\mathcal{E}_s\mathcal{E}_m\right) \\
&\quad + \overline{Z}_{ZZ} \cdot \left(\sum_{s=0}^{t-1} \hat{Z}_{t-1-s}\mathcal{E}_s, \sum_{m=0}^{t-1} \hat{Z}_{t-1-m}\mathcal{E}_m\right) + \overline{Z}_{\sigma\sigma} \\
&= \overline{Z}_Z \cdot \hat{Z}_{\sigma\sigma,t-1} + \overline{Z}_{\sigma\sigma} + \sum_{s=0}^{t-1}\sum_{m=0}^{t-1} \left(\overline{Z}_Z \cdot \hat{Z}_{t-1-s,t-1-m} + \overline{Z}_{ZZ} \cdot \left(\hat{Z}_{t-1-s}, \hat{Z}_{t-1-m}\right)\right)\mathcal{E}_s\mathcal{E}_m \\
&= \hat{Z}_{\sigma\sigma,t} + \sum_{s=0}^{t}\sum_{m=0}^{t} \hat{Z}_{t-s,t-m}\mathcal{E}_s\mathcal{E}_m
\end{aligned}$$

where in the second equality we used the fact that $\overline{Z}_{ZZ}$ is a bi-linear mapping and in the third equality we use the recursive definitions of $\hat{Z}_{\sigma\sigma,t}$ and $\hat{Z}_{t,k}$, and $\hat{Z}_{0,0} = 0$. □

[30] There are also $\overline{X}_{\sigma Z}$ and $\overline{Z}_{\sigma Z}$ terms but they are 0 following the same logic as $\overline{X}_\sigma$ and $\overline{Z}_\sigma$ being 0 in the proof of Lemma 1

Finally we plug in for $\overline{Z}_{t,\sigma\sigma}\left(\mathcal{E}^t\right)$ and $\overline{Z}_{t,\sigma}\left(\mathcal{E}^t\right)$ in equation (82) to find

$$\begin{aligned}\overline{X}_{t,\sigma\sigma}\left(\mathcal{E}^t\right) &= \overline{X}_Z\cdot\left(\hat{Z}_{\sigma\sigma,t}+\sum_{s=0}^{t}\sum_{m=0}^{t}\hat{Z}_{t-s,t-m}\mathcal{E}_s\mathcal{E}_m\right)+\overline{X}_{ZZ}\cdot\left(\sum_{s=0}^{t}\hat{Z}_{t-s}\mathcal{E}_s,\sum_{m=0}^{t}\hat{Z}_{t-m}\mathcal{E}_m\right)\\ &= \overline{X}_Z\cdot\hat{Z}_{\sigma\sigma,t}+\sum_{s=0}^{t}\sum_{m=0}^{t}\left(\overline{X}_Z\cdot\hat{Z}_{t-s,t-m}+\overline{X}_{ZZ}\cdot\left(\hat{Z}_{t-s},\hat{Z}_{t-m}\right)\right)\mathcal{E}_s\mathcal{E}_m\\ &= \overline{X}_{\sigma\sigma,t}+\sum_{s=0}^{t}\sum_{m=0}^{t}\overline{X}_{ZZ,t-s,t-m}\mathcal{E}_s\mathcal{E}_m\end{aligned}$$

which completes the proof.

A.9 Proof of Lemma 2a^{SO}

We assume knowledge of $\overline{x}_{ZZ,0,0}(a,\theta)$. To find $\overline{x}_{\sigma\sigma}(a,\theta)$, for any (a,θ) not on a kink, differentiate the F mapping twice with respect to σ and add to it the derivative of F in direction $\hat{Z}_{\sigma\sigma,t}$

$$0=\mathsf{F}_x(a,\theta)\overline{x}_{\sigma\sigma,t}(a,\theta)+\mathsf{F}_Y(a,\theta)\overline{X}_{\sigma\sigma,t}+\mathsf{F}_{x^e}(a,\theta)\left(\mathbb{E}_{\varepsilon,\mathcal{E}}\tilde{x}\right)_{\sigma\sigma,t}$$

where

$$\mathbb{E}_{\varepsilon,\mathcal{E}}\widetilde{x}=\sum_{j=0}^{N^\vee}\int_{\tilde{\theta}^{\vee,j}(\tilde{a}(a,\theta,Z),\tilde{Z}(\sigma\mathcal{E},Z;\sigma))}^{\tilde{\theta}^{\vee,j+1}(\tilde{a}(a,\theta,Z),\tilde{Z}(\sigma\mathcal{E},Z;\sigma))}\widetilde{x}\left(\tilde{a}(a,\theta,Z;\sigma),\theta',\tilde{Z}(\sigma\mathcal{E},Z;\sigma)\right)\mu\left(\theta'-\rho_\theta\theta\right)d\theta' d\Pr\left(\mathcal{E}\right)$$

and $\tilde{Z}(\sigma\mathcal{E},Z;\sigma)=\left[\sigma\mathcal{E},\tilde{A}(Z;\sigma),\tilde{\Omega}(Z;\sigma)\right]^T$ .Taking the second derivative of this object with respect to σ and adding to it the derivative in direction $\hat{Z}_t$ yields (after exploiting continuity of $\overline{x}$ and $\overline{x}_\sigma=0$) yields[31]

$$\begin{aligned}\left(\mathbb{E}_{\varepsilon,\mathcal{E}}\tilde{x}\right)_{\sigma\sigma,t} &= \mathbb{E}_\varepsilon\left[\mathring{\overline{x}}_{ZZ,0,0}|a,\theta\right]\operatorname{var}(\mathcal{E})+\mathbb{E}_\varepsilon\left[\overline{x}_{\sigma\sigma,t+1}|a,\theta\right]+\mathbb{E}_\varepsilon\left[\overline{x}_a|a,\theta\right]\mathsf{P}\overline{x}_{\sigma\sigma,t}(a,\theta)\\ &\quad+\sum_{j=1}^{N^\vee}\overline{x}_{Z,0}\left(\overline{a}(a,\theta),\overline{\theta}^{\vee,j}(\overline{a}(a,\theta))\right)\overline{\theta}^{\vee,j}_{Z,0}(\overline{a}(a,\theta))\mu\left(\overline{\theta}^{\vee,j}_{Z,0}(\overline{a}(a,\theta))\right)\end{aligned}$$

using the distributional derivative from $\overline{x}_{ZZ,0,0}$ and exploiting that $\tilde{\theta}^{\vee,j}$ is the inverse of $\tilde{a}^{\vee,j}$ we can write this more succinctly as

$$\left(\mathbb{E}_{\varepsilon,\mathcal{E}}\tilde{x}\right)_{\sigma\sigma,t}=\mathbb{E}_\varepsilon\left[\overline{x}_{ZZ,0,0}|a,\theta\right]\operatorname{var}(\mathcal{E})+\mathbb{E}_\varepsilon\left[\overline{x}_{\sigma\sigma,t+1}|a,\theta\right]+\mathbb{E}_\varepsilon\left[\overline{x}_a|a,\theta\right]\mathsf{P}\overline{x}_{\sigma\sigma,t}(a,\theta)$$

Let $\mathsf{x}_{\sigma\sigma}(a,\theta)$ be the function that solves the following linear functional equation

$$0=\mathsf{F}_x(a,\theta)\mathsf{x}_{\sigma\sigma}(a,\theta)+\mathsf{F}_{x^e}(a,\theta)\left(\mathbb{E}_\varepsilon\left[\overline{x}_{ZZ,0,0}|a,\theta\right]\operatorname{var}(\mathcal{E})+\mathbb{E}_\varepsilon\left[\mathsf{x}_{\sigma\sigma}|a,\theta\right]+\mathbb{E}_\varepsilon\left[\overline{x}_a|a,\theta\right]\mathsf{x}_{\sigma\sigma}(a,\theta)\right).$$

Subtracting these two equations and defining $\hat{x}_{\sigma\sigma,t}(a,\theta)=\overline{x}_{\sigma\sigma,t}(a,\theta)-\mathsf{x}_{\sigma\sigma}(a,\theta)$ we see that

$$0=\mathsf{F}_x(a,\theta)\hat{x}_{\sigma\sigma,t}+\mathsf{F}_X(a,\theta)\overline{X}_{\sigma\sigma,t}+\mathsf{F}_{x'}(a,\theta)\left(\mathbb{E}\left[\hat{x}_{\sigma\sigma,t+1}|a,\theta\right]+\mathbb{E}\left[\overline{x}_a|a,\theta\right]\mathsf{P}\hat{x}_{\sigma\sigma,t}\right).$$

[31] As $\overline{x}_\sigma$ is uniformly zero, the same steps that show distributional derivatives are equal to the classical derivatives to first order can be used to show $\overline{x}_{\sigma\sigma}=\mathring{\overline{x}}_{\sigma\sigma}$.

This is identical to system of equations solved by $\overline{x}_{Z,t}$ which allows us to conclude that

$$\hat{x}_{\sigma\sigma,t}(a,\theta) = \sum_{s=0}^{\infty} \mathsf{x}_s(a,\theta)\overline{X}_{\sigma\sigma,t+s}$$

which implies (34).

A.10 Proof of Proposition 1a^{SO}

The same steps as the proof of Lemma 3^{FO} can be used to show that $\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t}$ satisfies the recursive equation

$$\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t+1} = \mathcal{L}^{(a)} \cdot \frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t} - \mathcal{M} \cdot \overline{a}_{\sigma\sigma,t}.$$

We can then proceed to prove the following Claim

Claim 8. $\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t}$ is given by

$$\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t} = -\sum_{s=0} \mathsf{A}_{t,s}\overline{Y}_{\sigma\sigma,s} - \mathsf{B}_{\sigma\sigma,t}$$

where $\mathsf{A}_{t,s}$ is as defined in Corollary 1^{FO} and $\mathsf{B}_{\sigma\sigma,0} = 0$ and $\mathsf{B}_{\sigma\sigma,t+1} = \mathcal{L}^{(a)} \cdot \mathsf{B}_{\sigma\sigma,t} + \mathcal{M} \cdot \mathsf{a}_{\sigma\sigma}$

Proof. We proceed by induction. It's trivially true from $t = 0$ as $\mathsf{A}_{0,s} = 0, \mathsf{B}_{\sigma\sigma,0} = 0$, and $\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,0} = 0$. We then proceed by induction

$$\begin{aligned}
\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t} &= \mathcal{L}^{(a)} \cdot \frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t} - \sum_{j=0}^{\infty} \mathcal{M} \cdot \mathsf{a}_j \overline{Y}_{\sigma\sigma,t+j} - \mathcal{M} \cdot \mathsf{b}_{\sigma\sigma} \\
&= \mathcal{L}^{(a)} \cdot \left(-\sum_{s=0} \mathsf{A}_{t,s}\overline{Y}_{\sigma\sigma,s} - \mathsf{B}_{\sigma\sigma,t}\right) - \sum_{s=0}^{\infty} \mathcal{M} \cdot \mathsf{a}_{s-t}\overline{Y}_{\sigma\sigma,s} - \mathcal{M} \cdot \mathsf{b}_{\sigma\sigma} \\
&= \sum_{s=0}^{\infty} -\left(\mathcal{L}^{(a)} \cdot \mathsf{A}_{t,s} + \mathcal{M} \cdot \mathsf{a}_{s-t}\right)\overline{Y}_{\sigma\sigma,s} - \left(\mathcal{L}^{(a)} \cdot \mathsf{B}_{\sigma\sigma,t} + \mathcal{M} \cdot \mathsf{b}_{\sigma\sigma}\right) \\
&\equiv -\sum_{s=0}^{\infty} \mathsf{A}_{t+1,s}\overline{Y}_{\sigma\sigma,s} - \mathsf{B}_{\sigma\sigma,t+1}
\end{aligned}$$

where the second equality is achieved by letting $s = t + j$ and WLOG setting $\mathsf{a}_k = 0$ for $k < 0$. □

Integration by parts then implies that

$$\int \overline{x} d\hat{\Omega}_{\sigma\sigma,t} = -\int \overline{x}_a \frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t} da d\theta = \sum_{s=0}^{\infty} \left(\mathcal{I}^{(a)} \cdot \mathsf{A}_{t,s}\right)\overline{Y}_{\sigma\sigma,s} + \mathcal{I}^{(a)} \cdot \mathsf{B}_{\sigma\sigma,t}$$

where the same arguments as in the first-order guarantee that $\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t}$ is a generalized function with mass-points at $\{a_n^*\}$ and thus the operation $\mathcal{I}^{(a)} \cdot \frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t}$ is well defined. This implies that

$$\begin{aligned}
\left(\int \overline{x} d\Omega\right)_{\sigma\sigma,t} &= \int \overline{x}_{\sigma\sigma,t} d\Omega^* + \int \overline{x} d\hat{\Omega}_{\sigma\sigma,t} \\
&= \sum_{s=0}^{\infty} \underbrace{\left(\int \mathsf{x}_{s-t} d\Omega^* + \mathcal{I}^{(a)} \cdot \mathsf{A}_{t,s}\right)}_{\mathsf{J}_{t,s}} \overline{Y}_{\sigma\sigma,s} + \underbrace{\int \mathsf{x}_{\sigma\sigma} d\Omega^* + \mathcal{I}^{(a)} \cdot \mathsf{B}_{\sigma\sigma,t}}_{:=\mathsf{H}_{\sigma\sigma,t}}
\end{aligned}$$

the LOM for $\mathsf{B}_{\sigma\sigma,t}$ in Claim 8 implies that $\mathsf{H}_{\sigma\sigma,t}$ satisfies the recursion $\mathsf{H}_{\sigma\sigma,0} = \int \mathsf{x}_{\sigma\sigma} d\Omega^*$ and $\mathsf{H}_{\sigma\sigma,t+1} = \mathsf{H}_{\sigma\sigma,t} + \mathcal{I}^{(a)} \cdot \left(\mathcal{L}^{(a)}\right)^{t-1} \cdot \mathcal{M} \cdot \mathsf{a}_{\sigma\sigma}$. Substituting for $\left(\int \overline{x} d\Omega\right)_{\sigma\sigma,t}$ in (32) completes the proof.

A.11 Derivation of Equation (39) and (37)

Differentiating the G map twice in directions $\hat{Z}_t$ and $\hat{Z}_k$ and adding to it the derivative in direction $\hat{Z}_{t,k}$ yields

$$\mathsf{G}_x \left(\int \overline{x} d\Omega\right)_{ZZ,t,k} + G_Y \overline{Y}_{ZZ,t,k} + \mathsf{G}_{\Theta\Theta,t,k} = 0$$

where

$$\begin{aligned}\mathsf{G}_{\Theta\Theta,t,k} =& \mathsf{G}_{xx} \cdot \left(\left(\int \overline{x} d\Omega\right)_{Z,t}, \left(\int \overline{x} d\Omega\right)_{Z,k}\right) + \mathsf{G}_{xY} \cdot \left(\left(\int \overline{x} d\Omega\right)_{Z,t}, \overline{Y}_{Z,k}\right) + \mathsf{G}_{Yx} \cdot \left(\overline{Y}_{Z,t}, \left(\int \overline{x} d\Omega\right)_{Z,k}\right) \\ &+ \mathsf{G}_{YY} \cdot \left(\overline{Y}_{Z,t}, \overline{Y}_{Z,k}\right)\end{aligned}$$

Finding $\left(\int \overline{x} d\Omega\right)_{ZZ,t,k}$ requires differentiating

$$\int \sum_{j=0}^{N^\vee} \int_{\overline{a}^{\vee,j}(\theta,Z)}^{\overline{a}^{\vee,j+1}(\theta,Z)} \overline{x}_{Z,t}(a,\theta,Z) d\Omega\langle a,\theta\rangle + \int \sum_{j=0}^{N^\vee} \int_{\overline{a}^{\vee,j}(\theta,Z)}^{\overline{a}^{\vee,j+1}(\theta,Z)} \overline{x}(a,\theta,Z) d\hat{\Omega}_t\langle a,\theta\rangle$$

in direction $\hat{Z}_k$ and adding to it

$$\left(\int \overline{x} d\Omega\right)_Z \cdot \hat{Z}_{t,k} = \int \overline{x}_Z \cdot \hat{Z}_{t,k} d\Omega^* + \int \overline{x} d\hat{\Omega}_{t,k}.$$

This implies

$$\begin{aligned}\left(\int \overline{x} d\Omega\right)_{ZZ,t,k} =& \int \mathring{\overline{x}}_{ZZ,t,k} d\Omega^* + \int \overline{x}_{Z,t} d\hat{\Omega}_k + \int \overline{x}_{Z,k} d\hat{\Omega}_k + \int \overline{x} d\hat{\Omega}_{t,k} \\ &- \int \sum_{j=1}^{N^\vee} \overline{x}^{\Delta,j}_{Z,t}(\theta) \overline{a}^{\vee,j}_{Z,k}(\theta) \omega^*(\overline{a}^{\vee,j}(\theta),\theta) d\theta\end{aligned}$$

Exploiting $\overline{x}^{\Delta,j}_{Z,t}(\theta)\overline{a}^{\vee,j}_{Z,k}(\theta) = -\overline{x}^{\Delta,j}_a(\theta)\overline{a}^{\vee,j}_{Z,t}(\theta)\overline{a}^{\vee,j}_{Z,k}(\theta)$ we can write this more concisely with the distributional derivative notation

$$\left(\int \overline{x} d\Omega\right)_{ZZ,t,k} = \int \overline{x}_{ZZ,t,k} d\Omega^* + \int \overline{x}_{Z,t} d\hat{\Omega}_k + \int \overline{x}_{Z,k} d\hat{\Omega}_k + \int \overline{x} d\hat{\Omega}_{t,k}$$

A.12 Proof of Lemma (2b^{SO})

To determine $\mathring{\overline{x}}_{ZZ,t,k}(a,\theta)$ at points away from the kinks we start with the derivative of the F mapping in the direction $\hat{Z}_{t,k}$ and then add to it the second derivative of the F mapping in directions $\hat{Z}_t$ and

$\hat{Z}_k$. Doing so yields

$$0 = \mathsf{F}_x(a,\theta)\mathring{\overline{x}}_{ZZ,t,k}(a,\theta) + \mathsf{F}_Y(a,\theta)\overline{Y}_{ZZ,t,k} + \mathsf{F}_{x^e}(a,\theta)\left(\mathbb{E}_\varepsilon\left[\overline{x}|a,\theta,Z\right]\right)_{ZZ,t,k}$$
$$+ \mathsf{F}_{xx}(a,\theta)\cdot\left(\overline{x}_{Z,t}(a,\theta),\overline{x}_{Z,k}(a,\theta)\right) + \mathsf{F}_{xY}(a,\theta)\cdot\left(\overline{x}_{Z,t}(a,\theta),\overline{Y}_{Z,k}\right) + \mathsf{F}_{xx^e}(a,\theta)\cdot\left(\overline{x}_{Z,t}(a,\theta),\overline{x}^e_{Z,k}(a,\theta)\right)$$
$$+ \mathsf{F}_{Yx}(a,\theta)\cdot\left(\overline{Y}_{Z,t},\overline{x}_{Z,k}(a,\theta)\right) + \mathsf{F}_{YY}(a,\theta)\cdot\left(\overline{Y}_{Z,t},\overline{Y}_{Z,k}\right) + \mathsf{F}_{Yx^e}(a,\theta)\cdot\left(\overline{Y}_{Z,t},\overline{x}^e_{Z,k}(a,\theta)\right)$$
$$+ \mathsf{F}_{x^e x}(a,\theta)\cdot\left(\overline{x}^e_{Z,t}(a,\theta),\overline{x}_{Z,k}(a,\theta)\right) + \mathsf{F}_{x^e Y}(a,\theta)\cdot\left(\overline{x}^e_{Z,t}(a,\theta),\overline{Y}_{Z,k}\right) + \mathsf{F}_{x^e x^e}(a,\theta)\cdot\left(\overline{x}^e_{Z,t}(a,\theta),\overline{x}^e_{Z,k}(a,\theta)\right)$$

where $\overline{x}^e_{Z,t}(a,\theta) := \mathbb{E}_\varepsilon\left[\overline{x}_a|a,\theta\right]\overline{x}_{Z,i}(a,\theta) + \mathbb{E}_\varepsilon\left[\overline{x}_{Z,t+1}|a,\theta\right]$. The term $\left(\mathbb{E}_\varepsilon\left[\overline{x}|a,\theta,Z\right]\right)_{ZZ,t,k}$ is obtained by differentiating

$$\sum_{j=0}^{N^\vee}\int_{\overline{\theta}^{\vee,j}(\overline{a}(a,\theta,Z),\overline{Z}(Z))}^{\overline{\theta}^{\vee,j+1}(\overline{a}(a,\theta,Z),\overline{Z}(Z))}\overline{x}\left(\overline{a}(a,\theta,Z),\theta',\overline{Z}(Z)\right)\mu(\theta'-\rho_\theta\theta)d\theta'$$

twice in directions $\hat{Z}_t$ and $\hat{Z}_k$ and then adding to it the derivative of in direction $\hat{Z}_{t,k}$. This implies

$$\left(\mathbb{E}_\varepsilon\left[\overline{x}|a,\theta,Z\right]\right)_{ZZ,t,k} = \mathbb{E}_\varepsilon\left[\mathring{\overline{x}}_{aa}|a,\theta\right]\overline{a}_{Z,t}(a,\theta)\overline{a}_{Z,k}(a,\theta) + \mathbb{E}\left[\mathring{\overline{x}}_{aZ,k+1}\right]\overline{a}_{Z,t}(a,\theta) + \mathbb{E}\left[\mathring{\overline{x}}_{aZ,t+1}\right]\overline{a}_{Z,k}(a,\theta) + \mathbb{E}_\varepsilon\left[\mathring{\overline{x}}_{ZZ,t,k}|a,\theta\right]$$
$$+\sum_{j=1}^{N^\vee}\Bigg(\left(\overline{x}^{\Delta,j}_a\left(\overline{\theta}^{\vee,j}(\overline{a}(a,\theta))\right)\overline{a}_{Z,t}(a,\theta) + \overline{x}^{\Delta,j}_{Z,t+1}\left(\overline{\theta}^{\vee,j}(\overline{a}(a,\theta))\right)\right)$$
$$\times\left(\overline{\theta}^{\vee,j}_a(\overline{a}(a,\theta))\overline{a}_{Z,k}(a,\theta) + \overline{\theta}^{\vee,j}_{Z,k+1}(\overline{a}(a,\theta))\right)\mu\left(\overline{\theta}^{\vee,j}(\overline{a}(a,\theta)) - \rho_\theta\theta\right)\Bigg)$$

using the distributional derivatives and exploiting that $\overline{\theta}^{\vee,j}$ is the inverse of $\overline{a}^{\vee,j}$ we can write this more succinctly as

$$\left(\mathbb{E}_\varepsilon\left[\overline{x}|a,\theta,Z\right]\right)_{ZZ,t,k} = \mathbb{E}_\varepsilon\left[\overline{x}_{aa}|a,\theta\right]\overline{a}_{Z,t}(a,\theta)\overline{a}_{Z,k}(a,\theta) + \mathbb{E}\left[\overline{x}_{aZ,k+1}\right]\overline{a}_{Z,t}(a,\theta) + \mathbb{E}\left[x_{aZ,t+1}\right]\overline{a}_{Z,k}(a,\theta) + \mathbb{E}_\varepsilon\left[\overline{x}_{ZZ,t,k}|a,\theta\right].$$

As we can write $\mathbb{E}_\varepsilon\left[\overline{x}_{ZZ,t,k}|a,\theta\right] = \mathbb{E}_\varepsilon\left[\mathring{\overline{x}}_{ZZ,t,k}|a,\theta\right] + \mathbb{E}_\varepsilon\left[\overline{x}_{ZZ,t,k} - \mathring{\overline{x}}_{ZZ,t,k}|a,\theta\right]$ with $\overline{x}_{ZZ,t,k}$ being defined in Claim 1. We can write the following recursive expression for $\mathring{\overline{x}}_{ZZ,t,k}(a,\theta)$ as

$$\mathsf{F}_x(a,\theta)\mathring{\overline{x}}_{ZZ,t,k}(a,\theta) + \mathsf{F}_Y(a,\theta)\overline{Y}_{ZZ,t,k} + \mathsf{F}_{x^e}(a,\theta)\mathbb{E}_\varepsilon\left[\mathring{\overline{x}}_{ZZ,t,k}|a,\theta\right] + \mathsf{F}_{t,k}(a,\theta)$$

where

$$\mathsf{F}_{t,k}(a,\theta) = \mathsf{F}_{x^e}(a,\theta)\Bigg(\mathbb{E}_\varepsilon\left[\overline{x}_{aa}|a,\theta\right]\overline{a}_{Z,t}(a,\theta)\overline{a}_{Z,k}(a,\theta) + \mathbb{E}_\varepsilon\left[\overline{x}_{aZ,k+1}\right]\overline{a}_{Z,t}(a,\theta)$$
$$+ \mathbb{E}_\varepsilon\left[x_{aZ,t+1}\right]\overline{a}_{Z,k}(a,\theta) + \mathbb{E}_\varepsilon\left[\mathring{\overline{x}}_{ZZ,t,k}|a,\theta\right]\Bigg)$$
$$+ \mathsf{F}_{xx}(a,\theta)\cdot\left(\overline{x}_{Z,t}(a,\theta),\overline{x}_{Z,k}(a,\theta)\right) + \mathsf{F}_{xY}(a,\theta)\cdot\left(\overline{x}_{Z,t}(a,\theta),\overline{Y}_{Z,k}\right) + \mathsf{F}_{xx^e}(a,\theta)\cdot\left(\overline{x}_{Z,t}(a,\theta),\overline{x}^e_{Z,k}(a,\theta)\right)$$
$$+ \mathsf{F}_{Yx}(a,\theta)\cdot\left(\overline{Y}_{Z,t},\overline{x}_{Z,k}(a,\theta)\right) + \mathsf{F}_{YY}(a,\theta)\cdot\left(\overline{Y}_{Z,t},\overline{Y}_{Z,k}\right) + \mathsf{F}_{Yx^e}(a,\theta)\cdot\left(\overline{Y}_{Z,t},\overline{x}^e_{Z,k}(a,\theta)\right)$$
$$+ \mathsf{F}_{x^e x}(a,\theta)\cdot\left(\overline{x}^e_{Z,t}(a,\theta),\overline{x}_{Z,k}(a,\theta)\right) + \mathsf{F}_{x^e Y}(a,\theta)\cdot\left(\overline{x}^e_{Z,t}(a,\theta),\overline{Y}_{Z,k}\right) + \mathsf{F}_{x^e x^e}(a,\theta)\cdot\left(\overline{x}^e_{Z,t}(a,\theta),\overline{x}^e_{Z,k}(a,\theta)\right)$$

with all of these objects easily constructed from first order terms.

A.13 Proof of Lemma 3b^{SO}

We start by differentiating the LOM in direction $\hat{Z}_{t,k}$. The same arguments as the first order gives

$$\frac{d}{d\theta}\overline{\Omega}_Z \cdot \hat{Z}_{t,k} = \mathcal{L}^{(a)} \cdot \hat{\Omega}_{t,k} - \mathcal{M} \cdot \left(\overline{a}_Z \cdot \hat{Z}_{t,k}\right).$$

To get $\frac{d}{d\theta}\hat{\Omega}_{t+1,k+1}$ we add to it the derivative direction $\hat{Z}_k$ of

$$\begin{aligned}\frac{d}{d\theta}\overline{\Omega}_Z(Z) \cdot \hat{Z}_t\langle a',\theta'\rangle = &\int \sum_{j=0}^{N^\vee} \int_{\overline{a}^{\vee,j}(\theta,Z)}^{\overline{a}^{\vee,j+1}(\theta,Z)} \overline{\Lambda}(a',\theta',a,\theta,Z)\overline{a}_a(a,\theta,Z)\frac{d}{d\theta}\hat{\Omega}_t\langle a,\theta\rangle da d\theta \\ &- \int \sum_{j=0}^{N^\vee} \int_{\overline{a}^{\vee,j}(\theta,Z)}^{\overline{a}^{\vee,j+1}(\theta,Z)} \overline{\Lambda}(a',\theta',a,\theta,Z)\overline{a}_{Z,t}(a,\theta,Z)d\Omega\langle a,\theta\rangle\end{aligned}$$

where $\overline{\Lambda}(a',\theta',a,\theta,Z) \equiv \delta(\overline{a}(a,\theta,Z)-a')\mu(\theta'-\rho_\theta\theta)$. Doing so gives

$$\begin{aligned}\frac{d}{d\theta}\hat{\Omega}_{t+1,k+1}\langle a',\theta'\rangle = &\iint \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_a(a,\theta)\frac{d}{d\theta}\hat{\Omega}_{t,k}\langle a,\theta\rangle da d\theta - \iint \overline{\Lambda}(a',\theta',a,\theta)\mathring{\overline{a}}_{ZZ,t,k}(a,\theta)d\Omega^* \\ &+ \iint \overline{\Lambda}\left(a',\theta',a,\theta\right)\mathring{\overline{a}}_{aZ,k}(a,\theta)\frac{d}{d\theta}\hat{\Omega}_t\langle a,\theta\rangle da d\theta - \frac{d}{da'}\iint \overline{\Lambda}\left(a',\theta',a,\theta\right)\overline{a}_a(a,\theta)\overline{a}_{Z,k}(a,\theta)\frac{d}{d\theta}\hat{\Omega}_t\langle a,\theta\rangle da d\theta \\ &- \iint \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_{Z,t}(a,\theta)d\hat{\Omega}_k + \frac{d}{da'}\iint \overline{\Lambda}\left(a',\theta',a,\theta\right)\overline{a}_{Z,k}(a,\theta)\overline{a}_{Z,t}(a,\theta)d\Omega^* \\ &- \int \sum_{j=1}^{N^\vee} \overline{\Lambda}\left(a',\theta',\overline{a}^{\vee,j}(\theta),\theta\right)\underbrace{\overline{a}_a^{\Delta,j}(\theta)\overline{a}_{Z,k}^{\vee,j}(\theta)}_{-\overline{a}_{Z,k}^{\Delta,j}(\theta)}\frac{d}{d\theta}\hat{\Omega}_t\langle\overline{a}^{\vee,j}(\theta),\theta\rangle da d\theta \\ &+ \int \sum_{j=1}^{N^\vee} \overline{\Lambda}\left(a',\theta',\overline{a}^{\vee,j}(\theta),\theta\right)\underbrace{\overline{a}_{Z,t}^{\Delta,j}(\theta)\overline{a}_{Z,k}^{\vee,j}(\theta)}_{-\overline{a}_a^{\Delta,j}(\theta)\overline{a}_{Z,k}^{\vee,j}(\theta)\overline{a}_{Z,k}^{\Delta,j}(\theta)} d\Omega^*\langle\overline{a}^{\vee,j}(\theta),\theta\rangle.\end{aligned}$$

Finally, integration by parts implies

$$\begin{aligned}\iint \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_{Z,t}(a,\theta)d\hat{\Omega}_k = &-\iint \overline{\Lambda}(a',\theta',a,\theta)\mathring{\overline{a}}_{aZ,t}(a,\theta)\frac{d}{d\theta}\hat{\Omega}_t\langle a,\theta\rangle da d\theta \\ &+ \frac{d}{da'}\iint \overline{\Lambda}\left(a',\theta',a,\theta\right)\overline{a}_a(a,\theta)\overline{a}_{Z,t}(a,\theta)\frac{d}{d\theta}\hat{\Omega}_k\langle a,\theta\rangle da d\theta \\ &- \int \sum_{j=1}^{N^\vee} \overline{\Lambda}\left(a',\theta',\overline{a}^{\vee,j}(\theta),\theta\right)\overline{a}_{Z,t}^{\Delta,j}(\theta)\frac{d}{d\theta}\hat{\Omega}_k\langle\overline{a}^{\vee,j}(\theta),\theta\rangle da d\theta.\end{aligned}$$

All combined, using distributional derivatives to absorb the derivatives of the kinks, we have

$$\begin{aligned}\frac{d}{d\theta}\hat{\Omega}_{t+1,k+1}\langle a',\theta'\rangle = &\iint \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_a(a,\theta)\frac{d}{d\theta}\hat{\Omega}_{t,k}\langle a,\theta\rangle da d\theta - \iint \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_{ZZ,t,k}(a,\theta)d\Omega^* \\ &+ \iint \overline{\Lambda}\left(a',\theta',a,\theta\right)\overline{a}_{aZ,k}(a,\theta)\frac{d}{d\theta}\hat{\Omega}_t\langle a,\theta\rangle da d\theta - \frac{d}{da'}\iint \overline{\Lambda}\left(a',\theta',a,\theta\right)\overline{a}_a(a,\theta)\overline{a}_{Z,k}(a,\theta)\frac{d}{d\theta}\hat{\Omega}_t\langle a,\theta\rangle da d\theta \\ &- \iint \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_{aZ,t}(a,\theta)\frac{d}{d\theta}\hat{\Omega}_t\langle a,\theta\rangle da d\theta + \frac{d}{da'}\iint \overline{\Lambda}\left(a',\theta',a,\theta\right)\overline{a}_a(a,\theta)\overline{a}_{Z,t}(a,\theta)\frac{d}{d\theta}\hat{\Omega}_k\langle a,\theta\rangle da d\theta \\ &+ \frac{d}{da'}\iint \overline{\Lambda}\left(a',\theta',a,\theta\right)\overline{a}_{Z,k}(a,\theta)\overline{a}_{Z,t}(a,\theta)d\Omega^*\end{aligned}$$

which can be written more concisely as (42).

A.14 Proof of Corollary 1b^{SO}

We start with the following claim regarding $\frac{d}{da}$ and $\mathcal{L}^{(a)}$

Claim 9. Suppose that y is a generalized function with a finite number of mass points at $\{a_n^*\}$ then

$$\mathcal{L}^{(a)} \cdot \frac{d}{da}\mathsf{y} = -\mathcal{L}^{(aa)} \cdot \mathsf{y} + \frac{d}{da}\mathcal{L}^{(a,a)} \cdot \mathsf{y}$$

where $\mathcal{L}^{(aa)}$ and $\mathcal{L}^{(a,a)}$ are defined as $\mathcal{L}^{(a)}$ with $\overline{a}_a$ replaced with $\overline{a}_{aa}$ and $\overline{a}_a\overline{a}_a$ respectively.

Proof. We have

$$\begin{aligned}
\left(\mathcal{L}^{(a)} \cdot \frac{d}{da}\mathsf{y}\right)(a^{'},\theta^{'}) &= \int \sum_{j=0}^{N^{\vee}} \int_{\overline{a}^{\vee,j}(\theta)}^{\overline{a}^{\vee,j+1}(\theta)} \overline{\Lambda}(a^{'},\theta^{'},a,\theta)\overline{a}_a(a,\theta)\frac{d}{da}\mathsf{y}(a,\theta)da d\theta \\
&= -\int \overline{\Lambda}(a^{'},\theta^{'},a,\theta)\mathring{\overline{a}}_{aa}(a,\theta)\mathsf{y}(a,\theta)dad\theta - \int \overline{\Lambda}_a(a^{'},\theta^{'},a,\theta)\overline{a}_a(a,\theta)\mathsf{y}(a,\theta)dad\theta \\
&\quad - \int \sum_{j=0}^{N^{\vee}} \overline{\Lambda}(a^{'},\theta^{'},\overline{a}^{\vee,j}(\theta),\theta)\overline{a}_a^{\Delta,j}(\theta)\frac{d}{da}\mathsf{y}(\overline{a}^{\vee,j}(\theta),\theta)d\theta \\
&= -\int \underbrace{\overline{\Lambda}(a^{'},\theta^{'},a,\theta)\overline{a}_{aa}(a,\theta)\mathsf{y}(a,\theta)dad\theta}_{:=\left(\mathcal{L}^{(aa)}\cdot\mathsf{y}\right)(a',\theta')} + \frac{d}{da'}\underbrace{\int \overline{\Lambda}_a(a^{'},\theta^{'},a,\theta)\overline{a}_a(a,\theta)\overline{a}_a(a,\theta)\mathsf{y}(a,\theta)dad\theta}_{:=\left(\mathcal{L}^{(a,a)}\cdot\mathsf{y}\right)(a',\theta')}
\end{aligned}$$

where the second equality uses integration by parts and the third equality applies the relationship between $\overline{a}_{aa}$ and $\mathring{\overline{a}}_{aa}$. That y is a generalized function with a finite number of mass points at $\{a_n^*\}$ guarantees that these integrals are well defined using the same logic as the first order. □

Claim 9 allows us to prove the following claim on $\hat{\Omega}_{t,k}$

Claim 10. $\frac{d}{d\theta}\hat{\Omega}_{ZZ,t,k}$ is given by

$$\frac{d}{d\theta}\hat{\Omega}_{ZZ,t,k} = -\sum_{s=0}^{\infty}\mathsf{A}_{t,s}\overline{Y}_{ZZ,s,k-t+s} - \mathsf{B}_{t,k} + \frac{d}{da}\mathsf{C}_{t,k}$$

where $\mathsf{A}_{t,s}$ is as defined in Corollary 1^{FO}, and $\mathsf{B}_{t,k}$ and $\mathsf{C}_{t,k}$ are defined in Corollary 1b^{SO}.

Proof. We proceed by induction. The case when $t=0$ is trivial as $\hat{\Omega}_{0,k-t}=0$ and $\mathsf{A}_{0,s}=\mathsf{B}_{0,k-t}=$

$\mathsf{C}_{0,k-t}$. We then proceed by induction

$$\begin{aligned}
\frac{d}{d\theta}\hat{\Omega}_{ZZ,t+1,k+1} &= \mathcal{L}^{(a)}\cdot\frac{d}{d\theta}\hat{\Omega}_{ZZ,t,k} - \sum_{j=0}^{\infty}\mathsf{a}_j\overline{Y}_{ZZ,t+j,k+j} - \mathsf{b}_{t,k} + \frac{d}{da}\mathsf{c}_{t,k} \\
&= \mathcal{L}^{(a)}\cdot\left(-\sum_{s=0}^{\infty}\mathsf{A}_{t,s}\overline{Y}_{ZZ,s,k-t+s} - \mathsf{B}_{t,k} + \frac{d}{da}\mathsf{C}_{t,k}\right) \\
&\quad - \sum_{s=0}^{\infty}\mathsf{a}_{s-t}\overline{Y}_{ZZ,s,k-t+s} - \mathsf{b}_{t,k} + \frac{d}{da}\mathsf{c}_{t,k} \\
&= -\sum_{s=0}^{\infty}\left(\mathcal{L}^{(a)}\cdot\mathsf{A}_{t,s} + \mathsf{a}_{s-t}\right)\overline{Y}_{ZZ,s,k-t+s} - \left(\mathcal{L}^{(a)}\cdot\mathsf{B}_{t,k} + \mathsf{b}_{t,k}\right) \\
&\quad + \mathcal{L}^{(a)}\cdot\frac{d}{da}\mathsf{C}_{t,k} + \frac{d}{da}\mathsf{c}_{t,k} \\
&= -\sum_{s=0}^{\infty}\left(\mathcal{L}^{(a)}\cdot\mathsf{A}_{t,s} + \mathsf{a}_{s-t}\right)\overline{Y}_{ZZ,s,k-t+s} - \left(\mathcal{L}^{(a)}\cdot\mathsf{B}_{t,k} + \mathsf{b}_{t,k}\right) \\
&\quad - \mathcal{L}^{(aa)}\cdot\mathsf{C}_{t,k} + \frac{d}{da}\mathcal{L}^{(a,a)}\cdot\mathsf{C}_{t,k} + \frac{d}{da}\mathsf{c}_{t,k} \\
&= -\sum_{s=0}\mathsf{A}_{t+1,s}\overline{Y}_{ZZ,s,k-t+s} - \mathsf{B}_{t+1,k} + \frac{d}{da}\mathsf{C}_{t+1,k}
\end{aligned}$$

as desired. □

Next, by applying integration by parts we have

$$\begin{aligned}
\int\overline{x}d\hat{\Omega}_{t,k} &= -\int\overline{x}_a\left(-\sum_{s=0}\mathsf{A}_{t,s}\overline{Y}_{ZZ,s,k-t+s} - \mathsf{B}_{t,k} + \frac{d}{da}\mathsf{C}_{t,k}\right)dad\theta \\
&= \sum_{s=0}^{\infty}\left(\mathcal{I}\cdot\mathsf{A}_{t,s}\right)\overline{Y}_{ZZ,s,k-t+s} + \mathcal{I}\cdot\mathsf{B}_{t,k} - \int\sum_{j=0}^{N^{\vee}}\int_{\overline{a}^{\vee,j}(\theta)}^{\overline{a}^{\vee,j+1}(\theta)}\overline{x}_a(a,\theta)\frac{d}{da}\mathsf{C}_{t,k}(a,\theta)dad\theta \\
&= \sum_{s=0}^{\infty}\left(\mathcal{I}\cdot\mathsf{A}_{t,s}\right)\overline{Y}_{ZZ,s,k-t+s} + \mathcal{I}\cdot\mathsf{B}_{t,k} + \int\mathring{\overline{x}}_{aa}\mathsf{C}_{t,k}dad\theta + \int\sum_{j=1}^{N^{\vee}}\overline{x}_a^{\Delta,j}(\overline{a}^{\vee,j}(\theta),\theta)\mathsf{C}_{t,k}(\overline{a}^{\vee,j}(\theta),\theta)d\theta \\
&= \sum_{s=0}^{\infty}\left(\mathcal{I}\cdot\mathsf{A}_{t,s}\right)\overline{Y}_{ZZ,s,k-t+s} + \mathcal{I}\cdot\mathsf{B}_{t,k} + \underbrace{\int\overline{x}_{aa}\mathsf{C}_{t,k}dad\theta}_{:=\mathcal{I}^{(aa)}\cdot\mathsf{C}_{t,k}}
\end{aligned}$$

where once again we can use the same arguments as the first order to guarantee that both $\mathsf{B}_{t,k}$ and $\mathsf{C}_{t,k}$ are generalized functions with a finite number of mass points at $\{a_n^*\}$ so all the integrals are well defined.

Finally, tuning to equation 39, we note that

$$\begin{aligned}\int \overline{x}_{Z,k} d\hat{\Omega}_t &= \int \sum_{j=0}^{N^\vee} \int_{\overline{a}^{\vee,j}(\theta)}^{\overline{a}^{\vee,j+1}(\theta)} \overline{x}_{Z,k}(a,\theta) d\hat{\Omega}_t\langle a,\theta\rangle \\ &= -\iint \mathring{\overline{x}}_{aZ,k} \frac{d}{d\theta}\hat{\Omega}_t da d\theta - \int \sum_{j=1}^{N^\vee} \overline{x}^{\Delta,j}_{Z,k}(\overline{a}^{\vee,j}(\theta),\theta) \frac{d}{d\theta}\hat{\Omega}_t\langle \overline{a}^{\vee,j}(\theta),\theta\rangle \\ &= -\underbrace{\iint \overline{x}_{aZ,k} \frac{d}{d\theta}\hat{\Omega}_t da d\theta}_{:=\mathcal{I}^{(a)}_{Z,k}\cdot\frac{d}{d\theta}\hat{\Omega}_t}\end{aligned}$$

and similarly for $\int \overline{x}_{Z,t} d\hat{\Omega}_k$. Substituting all of these results into equation 39 yields the result of corollary $1b^{SO}$.

A.15 Proof of Proposition $1b^{SO}$

The Proposition is a direct result of Corollary $1b^{SO}$ and equation 37.

B Proofs of Section 5

In this section we present the proofs for the extensions presented in Section 5. As needed we will use distributional derivatives in place of the classical derivatives,

B.1 Proofs for Section 5.1

B.1.1 Proof of Lemma 1^{TD}

The path of aggregates, $X_t(\mathcal{E}^t;\Omega_0,\sigma)$, depends on the history of aggregate shocks, $\mathcal{E}^t$, and the initial state Ω_0. It can be constructed from the recursive representation $\tilde{X}(Z;\sigma)$ and $\tilde{\Omega}(Z;\sigma)$ by defining $Z_t(\mathcal{E}^t;\Omega_0,\sigma) = \left[\Theta_t\left(\mathcal{E}^t;\sigma\right), \Omega_t\left(\mathcal{E}^t;\Omega_0,\sigma\right)\right]^T$ recursively as follows: let $Z_0(\mathcal{E}_0;\Omega_0,\sigma) = [\sigma\mathcal{E}_t,\Omega_0]^T$ and for $t \geq 1$

$$Z_t\left(\mathcal{E}^t;\Omega_0,\sigma\right) = \left[\rho_\Theta\Theta_{t-1}\left(\mathcal{E}^{t-1};\sigma\right) + \sigma\mathcal{E}_t, \tilde{\Omega}\left(Z_{t-1}\left(\mathcal{E}^{t-1};\Omega_0,\sigma\right);\sigma\right)\right], \tag{84}$$

The path of aggregates can then be defined as

$$X_t(\mathcal{E}^t;\Omega_0,\sigma) = \tilde{X}\left(Z_t\left(\mathcal{E}^t;\Omega_0,\sigma\right);\sigma\right). \tag{85}$$

Defining $\overline{Z}_{t,\sigma}\left(\mathcal{E}^t\right)$ and $\overline{X}_{t,\sigma}\left(\mathcal{E}^t\right)$ as the derivatives of $Z_t\left(\mathcal{E}^t;\Omega_0,\sigma\right)$ and $X_t(\mathcal{E}^t;\Omega_0,\sigma)$ w.r.t σ evaluated at $\sigma = 0$ and $\Omega_0 = \Omega^*$. The same steps as in the proof of Lemma (1^{FO}) show that

$$\overline{X}_{t,\sigma}\left(\mathcal{E}^t\right) = \sum_{s=0}^{t} \overline{X}_{Z,t-s}\mathcal{E}_s.$$

Next taking the derivative of $Z_t(\mathcal{E}^t;\Omega_0,\sigma)$ and $X_t(\mathcal{E}^t;\Omega_0,\sigma)$ w.r.t. Ω_0 in the direction $\hat{\Omega}_0=\Omega_0-\Omega^*$, we have $\overline{Z}_{0,\Omega}\cdot\hat{\Omega}_0=\hat{Z}_{\Omega,0}$ and for $t\geq 0$

$$\overline{Z}_{t+1,\Omega}\left(\mathcal{E}^{t+1}\right)\cdot\hat{\Omega}_0=\overline{Z}_Z\cdot\overline{Z}_{t,\Omega}\left(\mathcal{E}^t\right)\cdot\hat{\Omega}_0 \tag{86}$$

$$\overline{X}_{t,\Omega}\left(\mathcal{E}^t\right)\cdot\hat{\Omega}_0=\overline{X}_Z\cdot\overline{Z}_{t,\sigma}\left(\mathcal{E}^t\right)\cdot\hat{\Omega}_0 \tag{87}$$

which implies that

$$\overline{Z}_{t,\Omega}\left(\mathcal{E}^t\right)\cdot\hat{\Omega}_0=\left(\overline{Z}_Z\right)^t\cdot\overline{Z}_{0,\Omega}\cdot\hat{\Omega}_0=\hat{Z}_{\Omega,t}$$

and

$$\overline{X}_{t,\Omega}\left(\mathcal{E}^t\right)\cdot\hat{\Omega}_0=\overline{X}_Z\cdot\overline{Z}_{t,\sigma}\left(\mathcal{E}^t\right)\cdot\hat{\Omega}_0=\overline{X}_{\Omega,t}.$$

All put together we have

$$\begin{aligned}X_t(\mathcal{E}^t;\Omega_0)&=\overline{X}_{t,\sigma}\left(\mathcal{E}^t\right)+\overline{X}_{t,\Omega}\left(\mathcal{E}^t\right)\cdot\hat{\Omega}_0+O\left(\|\mathcal{E},\hat{\Omega}_0\|^2\right)\\&=\sum_{s=0}^{t}\overline{X}_{Z,t-s}\mathcal{E}_s+\overline{X}_{\Omega,t}+O\left(\|\mathcal{E},\hat{\Omega}_0\|^2\right).\end{aligned}$$

Taking expectations completes the proof.

B.1.2 Proof of Proposition 1^{TD}

The proofs of Lemma 2^{FO}, and Lemma 3^{FO} go through unchanged with $\hat{Z}_{\Omega,t}$ replacing $\hat{Z}_t$, $\overline{X}_{\Omega,t}$ replacing $\overline{X}_{Z,t}$, $\overline{x}_{\Omega,t}$ replacing $\overline{x}_{Z,t}$, and $\hat{\Omega}_{\Omega,t}$ replacing $\hat{\Omega}_t$.

Rolling forward LoM allows us to prove the follow claim

Claim 11. $\frac{d}{d\theta}\hat{\Omega}_{\Omega,t}$ is given by

$$\frac{d}{d\theta}\hat{\Omega}_t=-\sum_{s=0}\mathsf{A}_{t,s}\overline{Y}_{Z,s}-\mathsf{B}_{\Omega,t}$$

where $\mathsf{A}_{t,s}$ is as defined in Corollary 1^{FO} where $\mathsf{A}_{\Omega,t}$ satisfies $\mathsf{A}_{\Omega,t+1}=\mathcal{L}^{(a)}\cdot\mathsf{A}_{\Omega,t}$ and $\mathsf{A}_{\Omega,0}=-\frac{d}{d\theta}\hat{\Omega}_0$

Proof. We proceed by induction. It's trivially true from $t=0$ as $\mathsf{A}_{0,s}=0$ and $\frac{d}{d\theta}\hat{\Omega}_{\Omega,0}=\frac{d}{d\theta}\hat{\Omega}_0$. We then proceed by induction

$$\begin{aligned}\frac{d}{d\theta}\hat{\Omega}_{t+1}&=\mathcal{L}^{(a)}\cdot\frac{d}{d\theta}\hat{\Omega}_t-\sum_{j=0}^{\infty}\mathcal{M}\cdot\mathsf{a}_j\overline{Y}_{Z,t+j}\\&=\mathcal{L}^{(a)}\cdot\left(-\sum_{s=0}^{\infty}\mathsf{A}_{t,s}\overline{Y}_{Z,s}-\mathsf{B}_{\Omega,t}\right)-\sum_{s=0}^{\infty}\mathcal{M}\cdot\mathsf{a}_{s-t}\overline{Y}_{Z,s}\\&=-\sum_{s=0}^{\infty}\left(\mathcal{L}^{(a)}\cdot\mathsf{A}_{t,s}+\mathcal{M}\cdot\mathsf{a}_{s-t}\right)\overline{X}_{Z,s}-\mathcal{L}^{(a)}\cdot\mathsf{B}_{\Omega,t}\\&\equiv-\sum_{s=0}^{\infty}\mathsf{A}_{t+1,s}\overline{Y}_{Z,s}-\mathsf{B}_{\Omega,t+1}\end{aligned}$$

where the second equality is achieved by letting $s=t+j$ and WLOG setting $\mathsf{a}_k=0$ for $k<0$. □

Next we use integration by parts to obtain

$$\int \overline{x} d\hat{\Omega}_{\Omega,t} = -\int \overline{x}_z \frac{d}{d\theta}\hat{\Omega}_{\Omega,t} dz d\theta = -\mathcal{I}^{(a)} \cdot \left(-\sum_{s=0}^{\infty} \mathsf{A}_{t,s} \overline{Y}_{Z,s} - \mathsf{B}_{\Omega,t+1} \right)$$
$$= \sum_{s=0}^{\infty} \left(\mathcal{I}^{(a)} \cdot \mathsf{A}_{t,s} \right) \overline{Y}_{Z,s} + \mathcal{I}^{(a)} \cdot \mathsf{A}_{\Omega,t}$$

This allows us to directly conclude that

$$\mathsf{G}_x \sum_{s=0}^{\infty} \mathsf{J}_{t,s} \overline{Y}_{\Omega,s} + \mathsf{G}_X \overline{Y}_{\Omega,t} + \mathsf{G}_x \mathsf{J}_{\Omega,t} = 0$$

where $\mathsf{J}_{\Omega,t} = \mathcal{I} \cdot \mathsf{A}_{\Omega,t}$.

B.2 Proofs for Section 5.2

B.2.1 Proof of Lemma 1^{SV}

We proceed in the same manner as the proof of Lemma 1^{SO}. The only difference is that the derivatives w.r.t. σ now depend on the level of risk Υ_t. As Υ_t only affects $\mathcal{E}_t$ which is scaled by σ all other derivatives are independent of Υ_t. Second-order derivatives of (16) and (15) w.r.t. σ to find $\overline{Z}_{0,\sigma\sigma}\left(\boldsymbol{\mathcal{E}}^0\right) = 0$ and [32]

$$\overline{Z}_{t+1,\sigma\sigma}\left(\boldsymbol{\mathcal{E}}^{t+1}\right) = \overline{Z}_Z \cdot \overline{Z}_{t,\sigma\sigma}\left(\boldsymbol{\mathcal{E}}^t\right) + \overline{Z}_{ZZ} \cdot \left(\overline{Z}_{t,\sigma}\left(\boldsymbol{\mathcal{E}}^t\right), \overline{Z}_{t,\sigma}\left(\boldsymbol{\mathcal{E}}^t\right)\right) + \overline{Z}_{\sigma\sigma}(\Upsilon_t) \tag{88}$$

$$\overline{X}_{t,\sigma\sigma}\left(\boldsymbol{\mathcal{E}}^t\right) = \overline{X}_Z \cdot \overline{Z}_{t,\sigma\sigma}\left(\boldsymbol{\mathcal{E}}^t\right) + \overline{X}_{ZZ} \cdot \left(\overline{Z}_{t,\sigma}\left(\boldsymbol{\mathcal{E}}^t\right), \overline{Z}_{t,\sigma}\left(\boldsymbol{\mathcal{E}}^t\right)\right) + \overline{X}_{\sigma\sigma}(\Upsilon_t) \tag{89}$$

where $\overline{Z}_{ZZ}$ is defined in the main text and $\overline{Z}_{\sigma\sigma} = \left[0, \overline{\Omega}_{\sigma\sigma}(\Upsilon_t)\right]^T$. We begin by showing the following claim relating $\overline{Z}_{t,\sigma\sigma}\left(\boldsymbol{\mathcal{E}}^t\right)$ to the directions $\hat{Z}_{t,k}$ and $\hat{Z}_{\sigma\sigma,t}(\mathcal{E}^t_\Upsilon)$.

Claim 12. For all t

$$\overline{Z}_{t,\sigma\sigma}\left(\boldsymbol{\mathcal{E}}^t\right) = \hat{Z}_{\sigma\sigma,t}(\mathcal{E}^t_\Upsilon) + \sum_{s=0}^{t}\sum_{m=0}^{t} \hat{Z}_{t-s,t-m}\mathcal{E}_s\mathcal{E}_m \tag{90}$$

Proof. We proceed by induction. As $\hat{Z}_{\sigma\sigma,0} = \hat{Z}_{0,0} = 0$ we conclude that equation (15) holds for $t = 0$ since $\overline{Z}_{0,\sigma\sigma}\left(\mathcal{E}^0\right) = 0$. Assuming (90) holds for $t-1$ we have

$$\begin{aligned}
\overline{Z}_{t,\sigma\sigma}\left(\boldsymbol{\mathcal{E}}^t\right) = {} & \overline{Z}_Z \cdot \left(\hat{Z}_{\sigma\sigma,t-1}(\mathcal{E}^{t-1}_\Upsilon) + \sum_{s=0}^{t-1}\sum_{m=0}^{t-1} \hat{Z}_{t-1-s,t-1-m}\mathcal{E}_s\mathcal{E}_m \right) \\
& + \overline{Z}_{ZZ} \cdot \left(\sum_{s=0}^{t-1} \hat{Z}_{t-1-s}\mathcal{E}_s, \sum_{m=0}^{t-1} \hat{Z}_{t-1-m}\mathcal{E}_m \right) + \overline{Z}_{\sigma\sigma}(\Upsilon_{t-1}) \\
& = \overline{Z}_Z \cdot \hat{Z}_{\sigma\sigma,t-1}(\mathcal{E}^{t-1}_\Upsilon) + \overline{Z}_{\sigma\sigma}(\Upsilon_{t-1}) + \sum_{s=0}^{t-1}\sum_{m=0}^{t-1} \left(\overline{Z}_Z \cdot \hat{Z}_{t-1-s,t-1-m} + \overline{Z}_{ZZ} \cdot \left(\hat{Z}_{t-1-s}, \hat{Z}_{t-1-m} \right) \right) \mathcal{E}_s\mathcal{E}_m \\
& = \hat{Z}_{\sigma\sigma,t}(\mathcal{E}^t_\Upsilon) + \sum_{s=0}^{t}\sum_{m=0}^{t} \hat{Z}_{t-s,t-m}\mathcal{E}_s\mathcal{E}_m
\end{aligned}$$

[32] There are also $\overline{X}_{\sigma Z}$ and $\overline{Z}_{\sigma Z}$ terms but they are 0 following the same logic as $\overline{X}_\sigma$ and $\overline{Z}_\sigma$ being 0 in the proof of Lemma 1

where in the second equality we used the fact that $\overline{Z}_{ZZ}$ is a bi-linear mapping and in the third equality we use the recursive definitions of $\hat{Z}_{\sigma\sigma,t}$ and $\hat{Z}_{t,k}$, and $\hat{Z}_{0,0} = 0$. □

Finally we plug in for $\overline{Z}_{t,\sigma\sigma}\left(\boldsymbol{\mathcal{E}}^t\right)$ and $\overline{Z}_{t,\sigma}\left(\boldsymbol{\mathcal{E}}^t\right)$ in equation (89) to find

$$\begin{aligned}
\overline{X}_{t,\sigma\sigma}\left(\boldsymbol{\mathcal{E}}^t\right) &= \overline{X}_Z \cdot \left(\hat{Z}_{\sigma\sigma,t}(\mathcal{E}^t_\Upsilon) + \sum_{s=0}^{t}\sum_{m=0}^{t} \hat{Z}_{t-s,t-m}\mathcal{E}_s\mathcal{E}_m\right) + \overline{X}_{ZZ} \cdot \left(\sum_{s=0}^{t} \hat{Z}_{t-s}\mathcal{E}_s, \sum_{m=0}^{t} \hat{Z}_{t-m}\mathcal{E}_m\right) \\
&= \overline{X}_Z \cdot \hat{Z}_{\sigma\sigma,t}(\mathcal{E}^t_\Upsilon) + \sum_{s=0}^{t}\sum_{m=0}^{t}\left(\overline{X}_Z \cdot \hat{Z}_{t-s,t-m} + \overline{X}_{ZZ} \cdot \left(\hat{Z}_{t-s}, \hat{Z}_{t-m}\right)\right)\mathcal{E}_s\mathcal{E}_m \\
&= \overline{X}_{\sigma\sigma,t}(\mathcal{E}^t_\Upsilon) + \sum_{s=0}^{t}\sum_{m=0}^{t} \overline{X}_{ZZ,t-s,t-m}\mathcal{E}_s\mathcal{E}_m
\end{aligned}$$

which completes the proof.

B.2.2 Proof of Lemma 3^{SV}

To find $\overline{x}_{\sigma\sigma}(a,\theta,\mathcal{E}^t_\Upsilon)$ differentiate the F mapping twice with respect to σ and add to it the derivative of F in direction $\hat{Z}_{\sigma\sigma,t}\left(\mathcal{E}^t_\Upsilon\right)$

$$\begin{aligned}
0 &= \mathsf{F}_x(a,\theta)\overline{x}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_\Upsilon) + \mathsf{F}_Y(a,\theta)\overline{Y}_{\sigma\sigma,t}(\mathcal{E}^t_\Upsilon) \\
&\quad + \mathsf{F}_{x^e}(a,\theta)\left(\mathbb{E}\left[\overline{x}_{ZZ,0,0}|a,\theta\right](1+\Upsilon_t)\operatorname{var}(\mathcal{E}) + \mathbb{E}\left[\overline{x}_{\sigma\sigma,t+1}|a,\theta,\mathcal{E}^t_\Upsilon\right] + \mathbb{E}\left[\overline{x}_a|a,\theta\right]\mathsf{P}\overline{x}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_\Upsilon)\right).
\end{aligned}$$

Let $\mathsf{x}_{\sigma\sigma}(a,\theta)$ be the same as in the proof of Lemma $2\mathrm{b}^{SO}$ and $\mathsf{x}_\Upsilon(a,\theta)$ be defined by

$$0 = \mathsf{F}_x(a,\theta)\mathsf{x}_\Upsilon(a,\theta) + \mathsf{F}_{x^e}(a,\theta)\left(\mathbb{E}\left[\overline{x}_{ZZ,0,0}|a,\theta\right]\operatorname{var}(\mathcal{E}) + \rho_\Upsilon\mathbb{E}\left[\mathsf{x}_\Upsilon|a,\theta\right] + \mathbb{E}\left[\overline{x}_a|a,\theta\right]\mathsf{x}_\Upsilon(a,\theta)\right).$$

If we then define $\hat{x}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_\Upsilon) = \overline{x}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_\Upsilon) - \mathsf{x}_{\sigma\sigma}(a,\theta) - \Upsilon_t\mathsf{x}_\Upsilon(a,\theta)$ we see that $\hat{x}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_\Upsilon)$ solves

$$0 = \mathsf{F}_x(a,\theta)\hat{x}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_\Upsilon) + \mathsf{F}_Y(a,\theta)\overline{Y}_{\sigma\sigma,t}(\mathcal{E}^t_\Upsilon) + \mathsf{F}_{x^e}(a,\theta)\left(\mathbb{E}\left[\hat{x}_{\sigma\sigma,t+1}|a,\theta,\mathcal{E}^t_\Upsilon\right] + \mathbb{E}\left[\overline{x}_a|a,\theta\right]\mathsf{P}\hat{x}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_\Upsilon)\right).$$

This linear system of equations is identical to the one solved by $\overline{x}_{Z,t}$ which allows us to conclude that

$$\hat{x}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_\Upsilon) = \sum_{s=0}^{\infty}\mathsf{x}_s(a,\theta)\mathbb{E}\left[\overline{Y}_{\sigma\sigma,t+s}|\mathcal{E}^t_\Upsilon\right]$$

which implies (49).

B.2.3 Proof of Proposition (1^{SV})

We begin by deriving the recursive LoM for $\hat{\Omega}_{\sigma\sigma,t+1}\left(\mathcal{E}^{t+1}_\Upsilon\right)$. Differentiating the LoM twice with respect to σ and adding to it the derivative of the LOM in direction $\hat{Z}_{\sigma\sigma,t}\left(\mathcal{E}^t_\Upsilon\right)$, after applying integration by parts, yields

$$\begin{aligned}
\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t+1}\left(\mathcal{E}^{t+1}_\Upsilon\right)\langle a',\theta'\rangle &= \int \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_a(a,\theta)\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t}\left(\mathcal{E}^t_\Upsilon\right)\langle a,\theta\rangle d\theta da \\
&\quad - \int \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_\Upsilon)d\Omega^*
\end{aligned}$$

Substituting for $\overline{a}_{\sigma\sigma,t}\overline{a}_{\sigma\sigma,t}(a,\theta,\mathcal{E}^t_\Upsilon)$ using Lemma 3^{SV} immediately obtains the LoM

$$\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t+1}\left(\mathcal{E}^{t+1}_\Upsilon\right) = \mathcal{L}^{(a)}\cdot\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t}\left(\mathcal{E}^{t}_\Upsilon\right) - \mathcal{M}\cdot\left(\sum_{s=0}^{\infty}\mathsf{a}_s\mathbb{E}\left[\overline{Y}_{\sigma\sigma,t+s}|\mathcal{E}^t_\Upsilon\right] + \mathsf{a}_{\sigma\sigma} + \Upsilon_t\mathsf{a}_\Upsilon\right)$$
$$\equiv \mathcal{L}^{(a)}\cdot\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t}\left(\mathcal{E}^{t}_\Upsilon\right) - \sum_{s=0}^{\infty}\mathcal{M}\cdot\mathsf{a}_s\mathbb{E}\left[\overline{Y}_{\sigma\sigma,t+s}|\mathcal{E}^t_\Upsilon\right] + \mathcal{M}\cdot\mathsf{a}_{\sigma\sigma} + \Upsilon_t\mathcal{M}\cdot\mathsf{a}_\Upsilon(a,\theta).$$

We then are able to prove the following Claim about $\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t+1}\left(\mathcal{E}^{t+1}_\Upsilon\right)$

Claim 13. $\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t}$ satisfies

$$\mathbb{E}\left[\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t}\right] = -\sum_{s=0}^{\infty}\mathsf{A}_{t,s}\mathbb{E}\left[\overline{Y}_{\sigma\sigma,s}\right] - \mathsf{B}_{\sigma\sigma,t}$$

and for $k \geq 0$

$$\Delta\mathbb{E}\left[\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,\tau+k}|\mathcal{E}^\tau_\Upsilon\right] = -\sum_{k=0}^{\infty}\mathsf{A}_{k,s}\Delta\mathbb{E}\left[\overline{Y}_{\sigma\sigma,\tau+s}|\mathcal{E}^\tau_\Upsilon\right] - \mathsf{B}_{\Upsilon,k}\mathcal{E}_{\Upsilon,\tau}$$

where $\Delta\mathbb{E}\left[Y|\mathcal{E}^\tau_\Upsilon\right] \equiv \mathbb{E}\left[Y|\mathcal{E}^\tau_\Upsilon\right] - \mathbb{E}\left[Y|\mathcal{E}^{\tau-1}_\Upsilon\right]$ and $\mathsf{B}_{\Upsilon,k}$ is defined by $\mathsf{B}_{\Upsilon,k+1} = \mathcal{L}^{(a)}\cdot\mathsf{B}_{\Upsilon,k} + \rho^k_\Upsilon\mathcal{M}\cdot\mathsf{a}_\Upsilon$ with $\mathsf{B}_{\Upsilon,0} = 0$.

Proof. We proceed by induction. The first equation is trivially true for $t = 0$ as $\mathsf{A}_{0,s} = 0, \mathsf{B}_{\sigma\sigma,0} = 0$ and $\frac{d}{d\theta}\hat{\Omega}_0 = 0$. We then proceed by induction (exploiting $\mathbb{E}\left[\Upsilon_t\right] = 0$)

$$\mathbb{E}\left[\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t+1}\right] = \mathcal{L}^{(a)}\cdot\mathbb{E}\left[\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t}\right] - \sum_{j=0}^{\infty}\mathcal{M}\cdot\mathsf{a}_j\mathbb{E}\left[\overline{Y}_{\sigma\sigma,t+j}\right] - \mathcal{M}\cdot\mathsf{a}_{\sigma\sigma}$$
$$= \mathcal{L}^{(a)}\cdot\left(-\sum_{s=0}^{\infty}\mathsf{A}_{t,s}\mathbb{E}\left[\overline{Y}_{\sigma\sigma,s}\right] - \mathsf{B}_{\sigma\sigma,t}\right) - \sum_{s=0}^{\infty}\mathcal{M}\cdot\mathsf{a}_{s-t}\mathbb{E}\left[\overline{Y}_{\sigma\sigma,s}|\mathcal{E}^t_\Upsilon\right] - \mathcal{M}\cdot\mathsf{a}_{\sigma\sigma}$$
$$= \sum_{s=0}^{\infty} - \left(\mathcal{L}^{(a)}\cdot\mathsf{A}_{t,s} + \mathcal{M}\cdot\mathsf{a}_{s-t}\right)\mathbb{E}\left[\overline{Y}_{\sigma\sigma,s}\right] - \left(\mathcal{L}^{(a)}\cdot\mathsf{B}_{\sigma\sigma,t} + \mathcal{M}\cdot\mathsf{a}_{\sigma\sigma}\right)$$
$$= -\sum_{s=0}^{\infty}\mathsf{A}_{t+1,s}\mathbb{E}\left[\overline{Y}_{\sigma\sigma,s}\right] - \mathsf{B}_{\sigma\sigma,t+1}$$

as desired.

For the second equation we note that it holds for $k = 0$ as $\Delta\mathbb{E}\left[\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,\tau}|\mathcal{E}^\tau_\Upsilon\right] = 0$ (since $\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,\tau}$ is measurable w.r..t $\mathcal{E}^{\tau-1}_\Upsilon$) and $\mathsf{A}_{0,s} = 0, \mathsf{B}_{\Upsilon,0} = 0$. Taking expectations of the LOM implies

$$\Delta\mathbb{E}\left[\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,\tau+k+1}|\mathcal{E}^\tau_\Upsilon\right] = \mathcal{L}^{(a)}\cdot\Delta\mathbb{E}\left[\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,\tau+k}|\mathcal{E}^\tau_\Upsilon\right] - \sum_{j=0}^{\infty}\mathcal{M}\cdot\mathsf{a}_j\Delta\mathbb{E}\left[\overline{Y}_{\sigma\sigma,\tau+k+j}|\mathcal{E}^\tau_\Upsilon\right] - \rho^k_\Upsilon\mathcal{M}\cdot\mathsf{a}_\Upsilon\mathcal{E}_{\Upsilon,\tau}$$
$$= \mathcal{L}^{(a)}\cdot\left(-\sum_{s=0}^{\infty}\mathsf{A}_{k,s}\Delta\mathbb{E}\left[\overline{Y}_{\sigma\sigma,\tau+s}|\mathcal{E}^\tau_\Upsilon\right] - \mathsf{B}_{\Upsilon,k}\mathcal{E}_{\Upsilon,\tau}\right) - \sum_{s=0}^{\infty}\mathcal{M}\cdot\mathsf{a}_{s-k}\Delta\mathbb{E}\left[\overline{Y}_{\sigma\sigma,\tau+s}|\mathcal{E}^\tau_\Upsilon\right]$$
$$- \rho^k_\Upsilon\mathcal{M}\cdot\mathsf{a}_\Upsilon\mathcal{E}_{\Upsilon,\tau}$$
$$= \sum_{s=0}^{\infty} - \left(\mathcal{L}^{(a)}\cdot\mathsf{A}_{k,s} + \mathcal{M}\cdot\mathsf{a}_{s-t}\right)\Delta\mathbb{E}\left[\overline{Y}_{\sigma\sigma,\tau+s}|\mathcal{E}^\tau_\Upsilon\right] - \left(\mathcal{L}\cdot\mathsf{B}_{\Upsilon,k} + \rho^k_\Upsilon\mathcal{M}\cdot\mathsf{a}_\Upsilon\right)\mathcal{E}_{\Upsilon,\tau}$$
$$= -\sum_{s=0}^{\infty}\mathsf{A}_{t+1,s}\Delta\mathbb{E}\left[\overline{Y}_{\sigma\sigma,\tau+s}|\mathcal{E}^\tau_\Upsilon\right] - \mathsf{B}_{\Upsilon,k+1}\mathcal{E}_{\Upsilon,\tau}$$

□

Differentiating the G mapping twice with respect to σ and adding to it the derivative in direction $\hat{Z}_{\sigma\sigma,t}\left(\mathcal{E}_{\Upsilon}^{t}\right)$ yields, after applying integration by parts,

$$\mathsf{G}_x \int \overline{x}_{\sigma\sigma}\left(\mathcal{E}_{\Upsilon}^{t}\right) d\Omega^* - \mathsf{G}_x \int \overline{x}_a \frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t}\left(\mathcal{E}_{\Upsilon}^{t}\right) da d\theta + \mathsf{G}_Y \overline{Y}_{\sigma\sigma,t}\left(\mathcal{E}_{\Upsilon}^{t}\right) = 0. \tag{91}$$

Taking expectations of both sides and substituting for $\mathbb{E}\left[\overline{x}_{\sigma\sigma}\right]$ and $\mathbb{E}\left[\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t}\right]$ using Lemma 3^{SV} and Claim 13 we have

$$\mathsf{G}_x \sum_{s=0}^{\infty} \mathsf{J}_{t,s}\mathbb{E}\left[\overline{Y}_{\sigma\sigma,s}\right] + \mathsf{G}_x \mathsf{H}_{\sigma\sigma,t} + \mathsf{G}_Y \mathbb{E}\left[\overline{Y}_{\sigma\sigma,t}\right] = 0$$

which implies that $\mathbb{E}\left[\overline{X}_{\sigma\sigma,t}\right]$ solves the same system of equations as the $\overline{X}_{\sigma\sigma,t}$ terms in Proposition $1b^{SO}$.

If we instead take expectations of 69 conditional on $\mathcal{E}_{\Upsilon}^{\tau}$ and subtract off the expectation conditional on $\mathcal{E}_{\Upsilon}^{t-1}$ we find, for $t = \tau + k$,

$$\begin{aligned} 0 = \mathsf{G}_x &\left(\sum_{j=0}^{\infty} \int \mathsf{x}_j d\Omega^* \Delta\mathbb{E}\left[\overline{Y}_{\sigma\sigma,\tau+k+j}|\mathcal{E}_{\Upsilon}^{\tau}\right] + \rho_{\Upsilon}^{k} \int \mathsf{x}_{\Upsilon} d\Omega^* \mathcal{E}_{\Upsilon,\tau} \right) \\ &+ \mathsf{G}_x \left(\sum_{s=0}^{\infty} \left(\mathcal{I}\cdot \mathsf{A}_{t,s}\right) \Delta\mathbb{E}\left[\overline{Y}_{\sigma\sigma,\tau+s}|\mathcal{E}_{\Upsilon}^{\tau}\right] + \left(\mathcal{I}\cdot \mathsf{B}_{\Upsilon,k}\right)\mathcal{E}_{\Upsilon,\tau} \right) + \mathsf{G}_Y \Delta\mathbb{E}\left[\overline{Y}_{\sigma\sigma,\tau+k}|\mathcal{E}_{\Upsilon}^{\tau}\right] \end{aligned}$$

or

$$\mathsf{G}_x \sum_{j=0} \mathsf{J}_{k,j} \Delta\mathbb{E}\left[\overline{Y}_{\sigma\sigma,\tau+k+j}|\mathcal{E}_{\Upsilon}^{\tau}\right] + \mathsf{H}_{\Upsilon,k}\mathcal{E}_{\Upsilon,\tau} + \mathsf{G}_Y \Delta\mathbb{E}\left[\overline{Y}_{\sigma\sigma,\tau+k}|\mathcal{E}_{\Upsilon}^{\tau}\right] = 0$$

where $\mathsf{H}_{\Upsilon,k} = \mathcal{I}\cdot\mathsf{B}_{\Upsilon,k} + \rho_{\Upsilon}^{k}\int \mathsf{x}_{\Upsilon} d\Omega^*$. This implies $\Delta\mathbb{E}\left[\overline{Y}_{\sigma\sigma,\tau+k}|\mathcal{E}_{\Upsilon}^{\tau}\right] = \overline{Y}_{\Upsilon,k}\mathcal{E}_{\Upsilon,\tau}$ where $\overline{Y}_{\Upsilon,k}$ solves (51). Our knowledge of $\mathbb{E}\left[\overline{X}_{\sigma\sigma,t}\right]$ and $\Delta\mathbb{E}\left[\overline{X}_{\sigma\sigma,\tau+k}|\mathcal{E}_{\Upsilon}^{\tau}\right]$ immediately implies

$$\overline{X}_{\sigma\sigma,t}\left(\mathcal{E}_{\Upsilon}^{t}\right) = \overline{X}_{\sigma\sigma,t} + \sum_{s=0}^{\infty} \overline{X}_{\Upsilon,t-s}\mathcal{E}_{\Upsilon,s}.$$

B.3 Proofs for Section 5.3

For this section we will allow $\mathcal{E}$ to be multivariate. This implies that all the derivatives $\overline{X}_{Z,t}$ and $\overline{x}_{Z,t}(a,\theta)$ should be interpreted as matrices. We let $\Sigma_{\mathcal{E}}$ represent the covariance matrix of $\mathcal{E}$.

B.3.1 Proof of Lemma 2^{PF}

We begin by differentiation equation (61) in direction $\hat{Z}_t$. For $t \geq 1$ this implies that

$$\mathsf{S}\mathbb{E}\left[\overline{x}_{Z,t}|a,\theta_{\text{-}}\right]\mathsf{R}\overline{Y} + \mathbb{E}\left[\mathsf{S}\overline{x}|a,\theta_{\text{-}}\right]\mathsf{R}\overline{Y}_{Z,t} = 0.$$

The steady state implies $\mathsf{R}\overline{Y} = 0$ which implies that this equation can only hold if $\mathsf{R}\overline{Y}_{Z,t} = 0$ when $t \geq 1$. As equation (61) does not depend on Θ it places no restrictions on $\mathsf{R}\overline{Y}_{Z,t} \equiv \overline{R}_{Z,0}$.

Next we differentiate equation (60) in direction $\hat{Z}_t$. Doing so yields

$$\mathsf{F}_x(a,\theta)\overline{x}_{Z,t}(a,\boldsymbol{\theta}) + \mathsf{F}_Y(a,\theta)\overline{Y}_{Z,t} + \mathsf{F}_{x^e}(a,\theta)\left(\mathbb{E}_\varepsilon\left[\overline{x}_a|a,\boldsymbol{\theta}\right]\mathsf{p}\overline{x}_{Z,t}(a,\boldsymbol{\theta}) + \mathbb{E}_\varepsilon\left[\overline{x}_{Z,t+1}|a,\boldsymbol{\theta}\right]\right) \\ +\mathsf{F}_k(a,\theta)\overline{k}(a,\theta_-)^T\mathsf{R}\overline{Y}_{Z,t} = 0.$$

For $t \geq 1$ this simplifies to

$$\mathsf{F}_x(a,\theta)\overline{x}_{Z,t}(a,\boldsymbol{\theta}) + \mathsf{F}_Y(a,\theta)\overline{Y}_{Z,t} + \mathsf{F}_{x^e}(a,\theta)\left(\mathbb{E}_\varepsilon\left[\overline{x}_a|a,\boldsymbol{\theta}\right]\overline{x}_{Z,t}(a,\boldsymbol{\theta}) + \mathbb{E}_\varepsilon\left[\overline{x}_{Z,t+1}|a,\boldsymbol{\theta}\right]\right) = 0$$

which is equivalent to (80) and is solved by

$$\overline{x}_{Z,t}(a,\boldsymbol{\theta}) = \sum_{s=0}^{\infty}\mathsf{x}_s(a,\theta)\overline{Y}_{Z,t+s}.$$

For $t = 0$ we have

$$\mathsf{F}_x(a,\theta)\overline{x}_{Z,0}(a,\boldsymbol{\theta}) + \mathsf{F}_Y(a,\theta)\overline{Y}_{Z,0} + \mathsf{F}_{x^e}(a,\theta)\left(\mathbb{E}_\varepsilon\left[\overline{x}_a|a,\boldsymbol{\theta}\right]\mathsf{p}\overline{x}_{Z,0}(a,\boldsymbol{\theta}) + \mathbb{E}_\varepsilon\left[\overline{x}_{Z,1}|a,\boldsymbol{\theta}\right]\right) \\ +\mathsf{F}_k(a,\theta)\overline{k}(a,\theta_-)^T\mathsf{R}\overline{Y}_{Z,0} = 0.$$

Substituting for $\overline{x}_{Z,1}(a,\boldsymbol{\theta})$ and solving for $\overline{x}_{Z,0}(a,\boldsymbol{\theta})$ implies

$$\overline{x}_{Z,0}(a,\boldsymbol{\theta}) = \sum_{s=0}^{\infty}\mathsf{x}_s(a,\theta)\overline{Y}_{Z,s} + \mathsf{r}(a,\theta)\overline{k}(a,\theta_-)^T\overline{R}_{Z,0},$$

as desired with

$$\mathsf{r}(a,\theta) = -\left(\mathsf{F}_x(a,\theta) + \mathsf{F}_{x^e}(a,\theta)\mathbb{E}_\varepsilon\left[\overline{x}_a|a,\boldsymbol{\theta}\right]\mathsf{p}\right)^{-1}\mathsf{F}_k(a,\theta)$$

Finally, to determine $\overline{k}(a,\theta_-)$ we differentiate (61) twice with respect to σ to get

$$\mathbb{E}\left[\mathsf{S}\overline{x}|a,\theta_-\right]\overline{R}_{\sigma\sigma,0} + \mathbb{E}\left[\overline{R}_{Z,0}\mathcal{E}\mathcal{E}^T\left(\mathsf{S}\overline{x}_{Z,0}\right)^T|a,\theta_-\right] = 0.$$

As $\mathcal{E}$ is independent of θ we conclude that this is equivalent to

$$\mathbb{E}\left[\mathsf{S}\overline{x}|a,\theta_-\right]\overline{R}_{\sigma\sigma,0} + \sum_{s=0}^{\infty}\overline{R}_{Z,0}\Sigma_{\mathcal{E}}\overline{Y}_{Z,s}^T\mathbb{E}\left[\mathsf{S}\mathsf{x}_s|a,\theta_-\right]^T + \mathbb{E}\left[\mathsf{S}\mathsf{x}^k|a,\theta_-\right]\overline{R}_{Z,0}\Sigma_{\mathcal{E}}\overline{R}_{Z,0}^T\overline{k}(a,\theta_-) = 0,$$

where we have exploited our knowledge that $\mathbb{E}\left[\mathsf{S}\mathsf{x}^k|a,\theta_-\right]$ is a real number. Defining $\mathfrak{S}\left(\overline{R}_{Z,0}\right) = \left(\overline{R}_{Z,0}\Sigma_{\mathcal{E}}\overline{R}_{Z,0}^T\right)^{-1}$ and solving for $\overline{k}(a,\theta_-)$ gives

$$\overline{k}(a,\theta_-) = \mathsf{k}_{\sigma\sigma}(a,\theta_-)\mathfrak{S}\left(\overline{R}_{Z,0}\right)\overline{R}_{\sigma\sigma,0} + \mathfrak{S}\left(\overline{R}_{Z,0}\right)\overline{R}_{Z,0}\Sigma_{\mathcal{E}}\sum_{s=0}^{\infty}\left(\mathsf{k}_s(a,\theta_-)\overline{Y}_{Z,s}\right)^T$$

where

$$\mathsf{k}_{\sigma\sigma}(a,\theta_-) \equiv -\frac{\mathbb{E}\left[\mathsf{S}\overline{x}|a,\theta_-\right]}{\mathbb{E}\left[\mathsf{S}\mathsf{x}^k|a,\theta_-\right]} \text{ and } \mathsf{k}_s(a,\theta_-) \equiv -\frac{\mathbb{E}\left[\mathsf{S}\mathsf{x}_s|a,\theta_-\right]}{\mathbb{E}\left[\mathsf{S}\mathsf{x}^k|a,\theta_-\right]}.$$

B.3.2 Proof Of Corollary 1^{PF}

The same steps as in the proof of Lemma 3^{FO} implies that differentiating with $\hat{Z}_t$ yields

$$\frac{d}{d\boldsymbol{\theta}}\hat{\Omega}_{t+1} = \mathcal{L}^{(a)} \cdot \frac{d}{d\boldsymbol{\theta}}\hat{\Omega}_t - \mathcal{M} \cdot \overline{a}_{Z,t}.$$

where the $\mathcal{L}$ and $\mathcal{M}$ operators are extended to include integrating over θ_{-}. For $t \geq 1$ substituting for $\overline{a}_{Z,t}$ yields

$$\frac{d}{d\boldsymbol{\theta}}\hat{\Omega}_{t+1} = \mathcal{L} \cdot \frac{d}{d\boldsymbol{\theta}}\hat{\Omega}_t - \sum_{s=0}^{\infty} \mathcal{M} \cdot \mathsf{a}_s \overline{Y}_{Z,t+s}$$

as desired. For $t = 0$ we exploit that

$$\begin{aligned}\overline{a}_{Z,0}(a,\boldsymbol{\theta}) = &\sum_{s=0}^{\infty} \mathsf{a}_s(a,\theta)\overline{X}_{Z,s} + \mathsf{pr}(a,\theta)\mathsf{k}_{\sigma\sigma}(a,\theta_{-})\overline{R}^T_{\sigma\sigma,0}\mathfrak{S}\left(\overline{R}_{Z,0}\right)\overline{R}_{Z,0}\\ &\sum_{s=0}^{\infty} \mathsf{pr}(a,\theta)\mathsf{k}_s(a,\theta_{-})\overline{X}_{Z,s}\Sigma_{\mathcal{E}}\overline{R}^T_{Z,0}\mathfrak{S}\left(\overline{R}_{Z,0}\right)\overline{R}_{Z,0}.\end{aligned}$$

Defining $\mathsf{a}^{PF}_{\sigma\sigma}(a,\boldsymbol{\theta}) \equiv \mathsf{pr}(a,\theta)\mathsf{k}_{\sigma\sigma}(a,\theta_{-})$ and $\mathsf{a}^{PF}_s(a,\boldsymbol{\theta}) \equiv \mathsf{pr}(a,\theta)\mathsf{k}_s(a,\theta_{-})$ to get

$$\frac{d}{d\boldsymbol{\theta}}\hat{\Omega}_1 = \mathcal{L}\cdot\frac{d}{d\boldsymbol{\theta}}\hat{\Omega}_0 - \sum_{s=0}^{\infty}\mathsf{a}_s\overline{Y}_{Z,s} - \left(\mathcal{M}\cdot\mathsf{a}^{PF}_{\sigma\sigma}\right)\overline{R}^T_{\sigma\sigma,0}\mathfrak{S}\left(\overline{R}_{Z,0}\right)\overline{R}_{Z,0} - \sum_{s=0}^{\infty}\left(\mathcal{M}\cdot\mathsf{a}^{PF}_s\overline{Y}_{Z,s}\right)\Sigma_{\mathcal{E}}\overline{R}^T_{Z,0}\mathfrak{S}\left(\overline{R}_{Z,0}\right)\overline{R}_{Z,0}$$

Next we show the following Claim

Claim 14. $\frac{d}{d\boldsymbol{\theta}}\hat{\Omega}_t$ is given by

$$\begin{aligned}\frac{d}{d\boldsymbol{\theta}}\hat{\Omega}_t = &-\sum_{s=0}\mathsf{A}_{t,s}\overline{Y}_{Z,s} - \left(\left(\mathcal{L}^{(a)}\right)^{t-1}\cdot\mathsf{a}^{PF}_{\sigma\sigma}\right)\overline{R}^T_{\sigma\sigma,0}\mathfrak{S}\left(\overline{R}_{Z,0}\right)\overline{R}_{Z,0}\\ &-\sum_{s=0}^{\infty}\left(\left(\mathcal{L}^{(a)}\right)^{t-1}\cdot\mathsf{a}^{PF}_s\right)\overline{Y}_{Z,s}\Sigma_{\mathcal{E}}\overline{R}^T_{Z,0}\mathfrak{S}\left(\overline{R}_{Z,0}\right)\overline{R}_{Z,0}.\end{aligned}$$

Proof. Time $t = 1$ holds trivially as $\frac{d}{d\boldsymbol{\theta}}\hat{\Omega}_1$ and $\mathsf{A}_{1,s} = \mathsf{a}_s$, For $t > 1$ we proceed by induction as

$$\begin{aligned}\frac{d}{d\boldsymbol{\theta}}\hat{\Omega}_{t+1} &= \mathcal{L}^{(a)}\cdot\frac{d}{d\boldsymbol{\theta}}\hat{\Omega}_t - \sum_{j=0}^{\infty}\mathcal{M}\cdot\mathsf{a}_j\overline{Y}_{Z,t+j}\\ &= \mathcal{L}^{(a)}\cdot\Bigg(-\sum_{s=0}\mathsf{A}_{t,s}\overline{Y}_{Z,s} - \left(\left(\mathcal{L}^{(a)}\right)^{t-1}\cdot\mathsf{a}^{PF}_{\sigma\sigma}\right)\overline{R}^T_{\sigma\sigma,0}\mathfrak{S}\left(\overline{R}_{Z,0}\right)\overline{R}_{Z,0}\\ &\quad -\sum_{s=0}^{\infty}\left(\left(\mathcal{L}^{(a)}\right)^{t-1}\cdot\mathsf{a}^{PF}_s\right)\overline{Y}_{Z,s}\Sigma_{\mathcal{E}}\overline{R}^T_{Z,0}\mathfrak{S}\left(\overline{R}_{Z,0}\right)\overline{R}_{Z,0}\Bigg) - \sum_{s=0}^{\infty}\mathcal{M}\cdot\mathsf{a}_{s-t}\overline{Y}_{Z,s}\\ &= -\sum_{s=0}^{\infty}\left(\mathcal{L}^{(a)}\cdot\mathsf{A}_{t,s} + \mathcal{M}\cdot\mathsf{a}_{s-t}\right)\overline{Y}_{Z,s} - \left(\left(\mathcal{L}^{(a)}\right)^{t}\cdot\mathsf{a}^{PF}_{\sigma\sigma}\right)\overline{R}^T_{\sigma\sigma,0}\mathfrak{S}\left(\overline{R}_{Z,0}\right)\overline{R}_{Z,0}\\ &\quad -\sum_{s=0}^{\infty}\left(\left(\mathcal{L}^{(a)}\right)^{t-1}\cdot\mathsf{a}^{PF}_s\right)\overline{Y}_{Z,s}\Sigma_{\mathcal{E}}\overline{R}^T_{Z,0}\mathfrak{S}\left(\overline{R}_{Z,0}\right)\overline{R}_{Z,0}\end{aligned}$$

which completes the proof. □

Differentiating $\int \tilde{x} d\Omega$ in direction $\hat{Z}_t$ and applying integration by parts yields

$$\left(\int \overline{x} d\Omega\right)_{Z,t} = \int \overline{x}_{Z,t} d\Omega^* - \mathcal{I}^{(a)} \cdot \frac{d}{d\boldsymbol{\theta}} \hat{\Omega}_t.$$

Corollary is completed by substituting for $\overline{x}_{Z,t}$ and $\frac{d}{d\boldsymbol{\theta}}\hat{\Omega}_t$ while noting that $\overline{x}_{Z,t}$ takes a special form at $t = 0$.

B.3.3 Proof of Proposition 1^{PF}

Differentiating the G mapping in direction $\hat{Z}_t$, and applying integration by parts, gives

$$\mathsf{G}_x \left(\int \overline{x} d\Omega\right)_{Z,t} + \mathsf{G}_X \overline{Y}_{Z,t} = 0.$$

Substituting for $\left(\int \overline{x} d\Omega\right)_{Z,t}$ using Corollary 1^{PF} we have

$$\begin{aligned} 0 = {} & \mathsf{G}_x \sum_{s=0}^{\infty} \mathsf{J}_{t,s} \overline{X}_{Z,s} + \mathsf{G}_X \overline{X}_{Z,t} + \mathsf{G}_\Theta \rho_\Theta^t \\ & + \mathsf{G}_s \left(\sum_{s=0}^{\infty} \left(\mathsf{J}_{t,s}^w \overline{X}_{Z,s}\right) \Sigma_{\mathcal{E}} \overline{R}_{Z,0}^T \mathfrak{S}\left(\overline{R}_{Z,0}\right) \overline{R}_{Z,0} + \mathsf{J}_{\sigma\sigma,t}^w \overline{R}_{\sigma\sigma,0}^T \mathfrak{S}\left(\overline{R}_{Z,0}\right) \overline{R}_{Z,0} \right). \end{aligned} \tag{92}$$

In addition to equation (92), it must also be the case that market clearing, equation (63), holds to zeroth-order which implies

$$\int \overline{k} d\Omega^* = \mathsf{K}\overline{X}$$

substituting for $\overline{k}$ using Lemma (2^{PF}) yields

$$\mathsf{K}_{\sigma\sigma} \mathfrak{S}\left(\overline{R}_{Z,0}\right) \overline{R}_{\sigma\sigma,0} + \mathfrak{S}\left(\overline{R}_{Z,0}\right) \overline{R}_{Z,0} \Sigma_{\mathcal{E}} \sum_{s=0}^{\infty} \left(\mathsf{K}_s \overline{X}_{Z,s}\right)^T = \mathsf{K}\overline{X}$$

where $\mathsf{K}_{\sigma\sigma} = \int \mathsf{k}_{\sigma\sigma} d\Omega^*$ and $\mathsf{K}_s = \int \mathsf{k}_s d\Omega^*$. Finally, the measurability constraint in (63) implies $\mathsf{T}\overline{Y}_{Z,0} = 0$ and in Lemma (2^{PF}) we showed that $\mathsf{R}\overline{Y}_{Z,t} = 0$ for $t \geq 1$ which completes the proof.

C Multivariate Extension

Here we extend our analysis to allow for a, θ, and Θ to be multidimensional. For the remainder of this section, we will let a^j represent the j^{th} element of a and θ^j represent the j^{th} element of θ. Almost all of the results extend directly with the caveat that the derivatives with respect to a, such as $\overline{x}_a(a, \theta)$, should now be viewed as matrices as opposed to vectors. In addition, the directions $\hat{Z}_t$ should be viewed as vectors with $\hat{Z}_t^j$ being the directions associated with the shocks Θ^j. Finally, to keep the analysis concise we will use distributional derivatives for all of the second derivatives of the kinked policy functions rather than explicitly working with the limits of the integrals.

C.1 First-order Approximation

Lemma 1^{FO}extends directly

Lemma 1(FO MV). *To the first-order approximation, satisfies*

$$X_t\left(\mathcal{E}^t\right) = \overline{X} + \sum_{s=0}^{t} \overline{X}_{Z,t-s}\mathcal{E}_s + O\left(\|\mathcal{E}\|^2\right).$$

with the only caveat being that now that $\mathcal{E}_t$ is a vector and $\overline{X}_{Z,t}$ is a matrix. Similarly, we are also able to show

Lemma 2(FO MV). *For any t,*

$$\mathsf{G}_x\left(\int \overline{x}_{Z,t} d\Omega^* + \int \overline{x} d\hat{\Omega}_t\right) + \mathsf{G}_Y \overline{Y}_{Z,t} = 0. \tag{93}$$

In a similarly manner that $\hat{Z}_t^j$ represents the change in the aggregate state t periods ahead associated with the shock Θ^j, $\hat{\Omega}_t^j$ represents the associated change in the distribution. Finally, Lemma 2^{FO} remains unchanged as

Lemma 3(FO MV). *For any t,*

$$\overline{x}_{Z,t}\left(a,\theta\right) = \sum_{s=0}^{\infty} \mathsf{x}_s\left(a,\theta\right)\overline{Y}_{Z,t+s}, \tag{94}$$

where matrices $\mathsf{x}_s\left(a,\theta\right)$ are given by

$$\mathsf{x}_0\left(a,\theta\right) = -\left(\mathsf{F}_x(a,\theta) + \mathsf{F}_{x^e}(a,\theta)\mathbb{E}\left[\overline{x}_a|a,\theta\right]\mathsf{P}\right)^{-1}\mathsf{F}_Y(a,\theta), \tag{95}$$

$$\mathsf{x}_{s+1}\left(a,\theta\right) = -\left(\mathsf{F}_x(a,\theta) + \mathsf{F}_{x^e}(a,\theta)\mathbb{E}\left[\overline{x}_a|a,\theta\right]\mathsf{P}\right)^{-1}\mathsf{F}_{x^e}(a,\theta)\mathbb{E}\left[\mathsf{x}_s|a,\theta\right]. \tag{96}$$

The first difference comes with Lemma 3^{FO}. The operators $\mathcal{L}$ and $\mathcal{M}$ remain essentially the same

$$\left(\mathcal{M}\cdot y\right)\langle a',\theta'\rangle := \int \overline{\Lambda}(a',\theta',a,\theta) y\left(a,\theta\right) d\Omega^*\left(a,\theta\right),$$

$$\left(\mathcal{L}^{(a)}\cdot y\right)\langle a',\theta'\rangle := \int \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_a(a,\theta) y\left(a,\theta\right) da d\theta.$$

with the understanding that now y is vector valued and $\overline{a}_a$ is a matrix. For notational simplicity we let $\frac{d}{d\theta}$ represent $\frac{\partial^{n_\theta}}{\partial\theta^1\partial\theta^2\cdots\partial\theta^{n_\theta}}$ and $\frac{d}{da}$ represent $\frac{\partial^{n_a}}{\partial a^1\partial a^2\cdots\partial a^{n_a}}$. For any vector valued function y we define$\nabla_a \mathsf{y} = \sum_j \frac{\partial}{\partial a^j}\mathsf{y}^j$ then we can show

Lemma 4(FO MV). *For any t, $\frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_t = \nabla_a \cdot \hat{\omega}_t$ where $\hat{\omega}_t$ satisfies a recursion*

$$\hat{\omega}_{t+1} = \mathcal{L}^{(a)}\cdot\hat{\omega}_t - \sum_{s=0}^{\infty}\mathcal{M}\cdot\mathsf{a}_s\overline{X}_{Z,t+s}, \tag{97}$$

where $\mathsf{a}_s = \mathcal{M}\cdot\mathsf{a}_s$ and $\hat{\omega}_0 = \mathbf{0}$.

Proof. We proceed by induction. It trivially holds for $t = 0$ as $\hat{\Omega}_0 = 0$. Assuming true for t we can differentiate the LoM in direction $\hat{Z}_t$ to get

$$\hat{\Omega}_{t+1}\langle a', \theta'\rangle = \iint \prod_{i=1}^{n_a} \iota\left(\overline{a}^i(a,\theta) \le a'^i\right) \prod_{k=1}^{n_\theta} \iota\left(\left(\rho_\theta\theta + \varepsilon\right)^k \le \theta'^k\right) \mu(\varepsilon) d\epsilon d\hat{\Omega}_t$$
$$- \sum_{j=1}^{n_a} \iint \delta\left(\overline{a}^j(a,\theta) - a'^j\right) \prod_{i\neq j} \iota\left(\overline{a}^i(a,\theta) \le a'^i\right) \prod_{k=1}^{n_\theta} \iota\left(\left(\rho_\theta\theta + \varepsilon\right)^k \le \theta'^k\right) \mu(\varepsilon) d\epsilon \overline{a}^j_{Z,t}(a,\theta) d\Omega^*.$$

Applying $\frac{d}{da}\frac{d}{d\theta}$ to both sides yields

$$\begin{aligned}
\frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_{t+1}\langle a', \theta'\rangle &= \iint \prod_{i=1}^{n_a} \delta\left(\overline{a}^i(a,\theta) - a'^i\right) \prod_{k=1}^{n_\theta} \delta\left(\left(\rho_\theta\theta + \varepsilon\right)^k - \theta'^k\right) \mu(\varepsilon) d\epsilon d\hat{\Omega}_t \\
&\quad - \sum_{j=1}^{n_a} \frac{\partial}{\partial a'^j} \iint \prod_{i=1}^{n_a} \delta\left(\overline{a}^i(a,\theta) - a'^i\right) \prod_{k=1}^{n_\theta} \delta\left(\left(\rho_\theta\theta + \varepsilon\right)^k - \theta'^k\right) \mu(\varepsilon) d\epsilon \overline{a}^j_{Z,t}(a,\theta) d\hat{\Omega}_t \\
&= \int \overline{\Lambda}(a', \theta', a, \theta) d\hat{\Omega}_t - \sum_{j=1}^{n_a} \frac{\partial}{\partial a'^j} \int \overline{\Lambda}(a', \theta', a, \theta) \overline{a}^j_{Z,t}(a,\theta) d\Omega^* \\
&= \int \overline{\Lambda}(a', \theta', a, \theta) d\hat{\Omega}_t - \sum_{j=1}^{n_a} \frac{\partial}{\partial a'^j} \sum_s \int \overline{\Lambda}(a', \theta', a, \theta) \mathsf{a}^j_s(a,\theta) d\Omega^* \overline{X}_{Z,t+s}.
\end{aligned}$$

where in the second line we used the equality definition

$$\overline{\Lambda}(a', \theta', a, \theta) = \int \prod_{k=1}^{n_a} \delta\left(\overline{a}^k(a,\theta) - a'^k\right) \prod_{l=1}^{n_\theta} \delta\left(\left(\rho_\theta\theta + \varepsilon\right)^l - \theta'^l\right) \mu(\varepsilon) d\epsilon = \prod_{k=1}^{n_a} \delta\left(\overline{a}^k(a,\theta) - a'^k\right) \mu\left(\theta' - \rho_\theta\theta + \varepsilon\right)$$

If we apply $\frac{\partial}{\partial a^j}$ to both sides we find

$$\begin{aligned}
\frac{\partial}{\partial a^j}\overline{\Lambda}(a', \theta', a, \theta) &= \sum_i \delta'\left(\overline{a}^i(a,\theta) - a'^i\right) \prod_{k\neq i} \delta\left(\overline{a}^k(a,\theta) - a'^k\right) \mu\left(\theta' - \rho_\theta\theta + \varepsilon\right) \overline{a}^i_{a^j}(a,\theta) \\
&= -\sum_i \frac{\partial}{\partial a'^i} \delta\left(\overline{a}^i(a,\theta) - a'^i\right) \prod_{k\neq i} \delta\left(\overline{a}^k(a,\theta) - a'^k\right) \mu\left(\theta' - \rho_\theta\theta + \varepsilon\right) \overline{a}^i_{a^j}(a,\theta) \\
&= -\sum_i \frac{\partial}{\partial a'^i} \overline{\Lambda}(a', \theta', a, \theta) \overline{a}^i_{a^j}(a,\theta)
\end{aligned}$$

Next, we have

$$\begin{aligned}
\int \overline{\Lambda}(a', \theta', a, \theta) d\hat{\Omega}_t &= \iint \overline{\Lambda}(a', \theta', a, \theta) \frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_t(a,\theta) da d\theta \\
&= \sum_j \iint \overline{\Lambda}(a', \theta', a, \theta) \frac{\partial}{\partial a^j} \hat{\omega}^j_t(a,\theta) da d\theta \\
&= -\sum_j \iint \frac{\partial}{\partial a^j}\left(\overline{\Lambda}(a', \theta', a, \theta)\right) \hat{\omega}^j_t(a,\theta) da d\theta \\
&= \sum_i \frac{\partial}{\partial a'^i} \iint \sum_j \overline{\Lambda}(a', \theta', a, \theta) \overline{a}^i_{a^j}(a,\theta) \hat{\omega}^j_t(a,\theta) da d\theta.
\end{aligned}$$

All combined this implies that

$$\frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_{t+1} = \nabla_a \mathcal{L}^{(a)} \cdot \hat{\omega}_t - \nabla_a \sum_s \mathcal{M} \cdot \mathsf{a}_s \overline{X}_{Z,t+s} = \nabla_a \hat{\omega}_{t+1}$$

which completes the proof. The same steps as in the proof of Lemma 3^{FO} guarantee that these integrals exist. □

It should be noted that when the dimensionality of a is 1 Lemma 4(FO MV) is equivalent to Lemma 3^{FO} as it states that $\frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_t = \frac{d}{da}\hat{\omega}_t$ and thus $\frac{d}{d\theta}\hat{\Omega}_t = \hat{\omega}_t$ which satisfies the same recursive system as in Lemma 3^{FO}.

For vector valued y, operator $\mathcal{I}^{(a)}$ is extended to be defined by

$$\mathcal{I}^{(a)} \cdot \mathsf{y} = \int \overline{x}_a(a,\theta)\mathsf{y}(a,\theta)dad\theta.$$

We are then able to state the following corollary

Corollary 2(FO MV). *For any t,*

$$\int \overline{x}d\hat{\Omega}_t = \sum_{s=0}^{\infty}\left(\mathcal{I}^{(a)} \cdot \mathsf{A}_{t,s}\right)\overline{X}_{Z,s},$$

where $\{\mathsf{A}_{t,s}\}_{t,s}$ follow a recursion $\mathsf{A}_{0,s} = 0$, and $\mathsf{A}_{t,s} = \mathcal{L} \cdot \mathsf{A}_{t-1,s} + \mathsf{a}_{s-t+1}$.

Proof. Following the same steps as the proof of Corollary 1^{FO} implies

$$\hat{\omega}_t = -\sum_{s=0}^{\infty}\mathsf{A}_{t,s}\overline{X}_{Z,s}.$$

We then have

$$\int \overline{x}d\hat{\Omega}_t = \sum_j \int \overline{x}\frac{\partial}{\partial a^j}\hat{\omega}_t^j dad\theta = -\sum_j \int \overline{x}_{a^j}\hat{\omega}_t^j dad\theta = -\mathcal{I}^{(a)} \cdot \hat{\omega}_t.$$

Substituting for $\hat{\omega}_t$ then completes the proof. □

Finally, we have that Proposition 1^{FO} holds identically for the multivariate case with the understanding that all the derivatives with respect to Z are vector valued.

Proposition 1(FO MV). *$\left\{\overline{X}_{Z,t}\right\}_t$ is the solution to*

$$\mathsf{G}_x\sum_{s=0}^{\infty}\mathsf{J}_{t,s}\overline{Y}_{Z,s} + \mathsf{G}_X\overline{Y}_{Z,t} = 0, \tag{98}$$

where $\left\{\mathsf{J}_{t,s}\right\}_{t,s}$ satisfies $\mathsf{J}_{t,s} = \int \mathsf{x}_{s-t}d\Omega^* + \mathcal{I}^{(a)} \cdot \mathsf{A}_{t,s}$.

C.2 Second-order Approximation

As with the first-order approximation, many of the Lemmas extend directly with the caveat that all derivatives with respect to a and $\hat{Z}_k$ are vector valued. We repeat the corollaries here for conciseness

Lemma 1(SO MV). *To the second-order approximation, X_t satisfies*

$$X_t\left(\mathcal{E}^t\right) = ... + \frac{1}{2}\left(\sum_{s=0}^{t}\sum_{m=0}^{t}\overline{X}_{ZZ,t-s,t-m}\cdot(\mathcal{E}_s,\mathcal{E}_m) + \overline{X}_{\sigma\sigma,t}\right) + O\left(\|\mathcal{E}\|^3\right), \tag{99}$$

where ... are the first-order terms.

Lemma 2(SO MV). *For any t,k*

$$\mathsf{G}_x\left(\int \overline{x}_{\sigma\sigma,t}d\Omega^* + \int \overline{x}d\hat{\Omega}_{\sigma\sigma,t}\right) + \mathsf{G}_X\overline{X}_{\sigma\sigma,t} = 0, \tag{100}$$

$$\mathsf{G}_x\left(\int \overline{x}_{ZZ,t,k}d\Omega^* + \int \overline{x}d\hat{\Omega}_{t,k} + \int \bar{x}_{Z,t}d\hat{\Omega}_k + \int \bar{x}_{Z,k}d\hat{\Omega}_t\right) + \mathsf{G}_X\overline{X}_{ZZ,t,s} + \mathsf{G}_{\Theta\Theta,t,k} = 0, \tag{101}$$

where explicit expression for $\mathsf{G}_{\Theta\Theta,t,k}$ is as before.

Lemma 3(SO MV). *For any t,*

$$\overline{x}_{\sigma\sigma,t}(a,\theta) = \sum_{s=0}^{\infty}\mathsf{x}_s(a,\theta)\overline{X}_{\sigma\sigma,t+s} + \mathsf{x}_{\sigma\sigma}(a,\theta), \tag{102}$$

$$\overline{x}_{ZZ,t,k}(a,\theta) = \sum_{s=0}^{\infty}\mathsf{x}_s(a,\theta)\overline{X}_{ZZ,t+s,k+s} + \mathsf{x}_{t,k}(a,\theta), \tag{103}$$

where explicit expressions for $\mathsf{x}_{\sigma\sigma}$ and $\mathsf{x}_{t,k}$ are provided in the appendix.

Next we extend the definitions of $\mathcal{L}$, $\mathcal{M}$ and $\mathcal{L}_{Z,t}$ to be of vector valued functions as follows

$$\mathcal{M}\cdot(\mathsf{y},\mathsf{w})\langle a',\theta'\rangle := \int \overline{\Lambda}(a',\theta',a,\theta)\mathsf{y}(a,\theta)\mathsf{w}(a,\theta)^T dad\theta$$

$$\mathcal{L}^{(a)}\cdot(\mathsf{y},\mathsf{w})\langle a',\theta'\rangle := \int \overline{\Lambda}(a',\theta',a,\theta)\left(\overline{a}_a(a,\theta)\mathsf{y}(a,\theta)\right)\mathsf{w}(a,\theta)^T dad\theta$$

$$\mathcal{L}_{Z,t}\cdot\mathsf{y}\langle a',\theta'\rangle := \int \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_{aZ,t}(a,\theta)\mathsf{y}\left(a,\theta\right)dad\theta$$

For a matrix valued function we define $\nabla_a\cdot\mathsf{y}$ as the vector

$$\left(\nabla_a\cdot\mathsf{y}\right)^j(a,\theta) = \sum_i \frac{\partial}{\partial a^i}\mathsf{y}^{i,j}(a,\theta),$$

which implies

$$\nabla_a^2\cdot\mathsf{y} \equiv \nabla_a\cdot\nabla_a\cdot\mathsf{y} = \sum_{i,j}\frac{\partial}{\partial a^i}\frac{\partial}{\partial a^j}\mathsf{y}^{i,j}$$

then we have the following extension of Lemma 3b^{SO}

Lemma 4(SO MV). *For all t, there exists $\hat{\omega}_{\sigma\sigma,t}$ and $\hat{\omega}_{t,k}$ such that $\frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t} = \nabla_a\cdot\hat{\omega}_{\sigma\sigma,t}$ and $\frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_{t,k} = \nabla_a\cdot\hat{\omega}_{t,k}$ where $\hat{\omega}_{\sigma\sigma,t}$ and $\hat{\omega}_{t,k}$ satisfy the following recursive equations*

$$\hat{\omega}_{\sigma\sigma,t+1} = \mathcal{L}\cdot\hat{\omega}_{\sigma\sigma,t} - \sum_{s=0}^{\infty}\mathcal{M}\cdot\mathsf{a}_s\overline{Y}_{\sigma\sigma,t+s} - \mathcal{M}\cdot\mathsf{a}_{\sigma\sigma}, \tag{104}$$

and

$$\hat{\omega}_{t+1,k+1} = \mathcal{L}\cdot\hat{\omega}_{t,k} - \sum_{s=0}^{\infty}\mathcal{M}\cdot\mathsf{a}_s\overline{Y}_{ZZ,t+s,k+s} - \mathcal{M}\cdot\mathsf{a}_{t,k}$$
$$+\nabla_a\cdot\mathcal{M}\cdot(\overline{a}_{Z,t},\overline{a}_{Z,k}) - \nabla_a\cdot\mathcal{L}\cdot(\hat{\omega}_t,\overline{a}_{Z,k}) - \nabla_a\cdot(\mathcal{L}\cdot(\hat{\omega}_k,\overline{a}_{Z,t}))^T \quad (105)$$
$$+\mathcal{L}_{Z,t}\cdot\hat{\omega}_k + \mathcal{L}_{Z,k}\cdot\hat{\omega}_k.$$

with $\hat{\omega}_{\sigma\sigma,0} = \hat{\omega}_{0,k} = \hat{\omega}_{t,0} = 0$

Proof. We proceed by induction which holds trivially for $t = 0$. Differentiating the LoM twice with respect to σ yield, adding the derivative in direction $\hat{Z}_{\sigma\sigma,t}$, and then applying $\frac{d}{da}\frac{d}{d\theta}$ to both sides yields

$$\frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t+1}\langle a',\theta'\rangle = \int\overline{\Lambda}(a',\theta',a,\theta)\frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t}\langle a,\theta\rangle dad\theta - \nabla_a\cdot\int\overline{\Lambda}(a',\theta',a,\theta)\overline{a}_{\sigma\sigma,t}(a,\theta)d\Omega^*.$$

Using the same steps as the proof of Lemma 4(FO MV) we have

$$\int\overline{\Lambda}(a',\theta',a,\theta)\nabla_a\cdot\hat{\omega}_{\sigma\sigma,t}(a,\theta)dad\theta = \nabla_a\cdot\int\overline{\Lambda}(a',\theta',a,\theta)\overline{a}_a(a,\theta)\hat{\omega}_{\sigma\sigma,t}(a,\theta)dad\theta.$$

so substituting for $\overline{a}^j_{\sigma\sigma,t}(a,\theta)$ using (102) and $\frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_{\sigma\sigma,t} = \nabla_a\cdot\hat{\omega}_t$ then gives

$$\frac{d}{da'}\frac{d}{d\theta'}\hat{\Omega}_{\sigma\sigma,t+1} = \nabla_a\cdot\left(\mathcal{L}\cdot\hat{\omega}_t - \sum_s(\mathcal{M}\cdot\mathsf{a}_s)\overline{Y}_{Z,t+s} - \mathcal{M}\cdot\mathsf{a}_{\sigma\sigma}\right).$$

Next we take the second derivative of the LoM in direction $\hat{Z}_t$ and $\hat{Z}_k$ and adding to it the derivative of the LoM in direction $\hat{Z}_{t,k}$ yields, after applying $\frac{d}{da}\frac{d}{d\theta}$ to both sides

$$\frac{d}{da'}\frac{d}{d\theta'}\hat{\Omega}_{t+1,k+1}\langle a',\theta'\rangle = \int\overline{\Lambda}(a',\theta',a,\theta)\frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_{t,k}\langle a,\theta\rangle dad\theta - \sum_i\frac{\partial}{\partial a'^i}\int\overline{\Lambda}(a',\theta',a,\theta)\overline{a}^i_{ZZ,t,k}(a,\theta)d\Omega^*$$
$$+\sum_{i,j}\frac{\partial}{\partial a'^i}\frac{\partial}{\partial a'^j}\int\overline{\Lambda}(a',\theta',a,\theta)\overline{a}^i_{Z,t}(a,\theta)\overline{a}^j_{Z,k}(a,\theta)d\Omega^*$$
$$-\sum_j\frac{\partial}{\partial a'^j}\int\overline{\Lambda}(a',\theta',a,\theta)\overline{a}^j_{Z,k}(a,\theta)\frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_t\langle a,\theta\rangle dad\theta$$
$$-\sum_j\frac{\partial}{\partial a'^j}\int\overline{\Lambda}(a',\theta',a,\theta)\overline{a}^j_{Z,t}(a,\theta)\frac{d}{da}\frac{d}{d\theta}\hat{\Omega}_k\langle a,\theta\rangle dad\theta$$

Written in vectorized form this is equivalent to

$$\frac{d}{da'}\frac{d}{d\theta'}\hat{\Omega}_{t+1,k+1}\langle a',\theta'\rangle = \int\overline{\Lambda}(a',\theta',a,\theta)\nabla_a\cdot\hat{\omega}_{t,k}dad\theta - \nabla_a\cdot\int\overline{\Lambda}(a',\theta',a,\theta)\overline{a}_{ZZ,t,k}(a,\theta)d\Omega^*$$
$$+\nabla_a^2\cdot\int\overline{\Lambda}(a',\theta',a,\theta)\overline{a}_{Z,t}(a,\theta)\overline{a}_{Z,k}(a,\theta)^Td\Omega^* - \nabla_a\cdot\int\overline{\Lambda}(a',\theta',a,\theta)\overline{a}_{Z,k}(a,\theta)\nabla_a\cdot\hat{\omega}_t dad\theta$$
$$-\nabla_a\cdot\int\overline{\Lambda}(a',\theta',a,\theta)\overline{a}_{Z,t}(a,\theta)\nabla_a\cdot\hat{\omega}_k dad\theta.$$

Next, we show the following relationship. For vector valued functions y,w, where y has compact support, we have

$$\begin{aligned}\int \overline{\Lambda}(a',\theta',a,\theta)\mathsf{w}(a,\theta)\nabla_a\cdot\mathsf{y}(a,\theta)d\theta da &= \sum_j\int \overline{\Lambda}(a',\theta',a,\theta)\mathsf{w}(a,\theta)\frac{\partial}{\partial a^j}\mathsf{y}^j(a,\theta)dad\theta\\
&= -\sum_j\int \overline{\Lambda}(a',\theta',a,\theta)\frac{\partial}{\partial a^j}\mathsf{w}(a,\theta)\mathsf{y}^j(a,\theta)dad\theta\\
&\quad -\sum_j\int \frac{\partial}{\partial a^j}\left(\overline{\Lambda}(a',\theta',a,\theta)\right)\mathsf{w}(a,\theta)\mathsf{y}^j(a,\theta)dad\theta\\
&= -\int \overline{\Lambda}(a',\theta',a,\theta)\mathsf{w}_a(a,\theta)\mathsf{y}(a,\theta)dad\theta\\
&\quad +\sum_{i,j}\frac{\partial}{\partial a'^i}\int \overline{\Lambda}(a',\theta',a,\theta)\overline{a}^i_{a^j}(a,\theta)\mathsf{w}(a,\theta)\mathsf{y}^j(a,\theta)dad\theta\\
&= -\int \overline{\Lambda}(a',\theta',a,\theta)\mathsf{w}_a(a,\theta)\mathsf{y}(a,\theta)dad\theta\\
&\quad +\nabla_a\cdot\int \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_a(a,\theta)\mathsf{y}(a,\theta)\mathsf{w}(a,\theta)^T dad\theta.\end{aligned}$$

Applying this relationship implies

$$\begin{aligned}\frac{d}{da'}\frac{d}{d\theta'}\hat{\Omega}_{t+1,k+1}\langle a',\theta'\rangle &= \nabla_a\cdot\int \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_a(a,\theta)\hat{\omega}_{t,k}dad\theta - \nabla_a\cdot\int \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_{ZZ,t,k}(a,\theta)d\Omega^*\\
&+\nabla_a^2\cdot\int \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_{Z,t}(a,\theta)\overline{a}_{Z,k}(a,\theta)^T d\Omega^* + \nabla_a\cdot\int \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_{Za,k}(a,\theta)\hat{\omega}_t dad\theta\\
&-\nabla_a^2\cdot\int \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_a(a,\theta)\hat{\omega}_t\overline{a}_{Z,k}(a,\theta)^T dad\theta + \nabla_a\cdot\int \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_{Za,t}(a,\theta)\hat{\omega}_k dad\theta\\
&-\nabla_a^2\cdot\left(\int \overline{\Lambda}(a',\theta',a,\theta)\overline{a}_a(a,\theta)\hat{\omega}_k\overline{a}_{Z,t}(a,\theta)^T dad\theta\right)^T\end{aligned}$$

which implies that $\frac{d}{da'}\frac{d}{d\theta'}\hat{\Omega}_{t+1,k+1} = \nabla_a\cdot\hat{\omega}_{t,k}$ where $\hat{\omega}_{t,k}$ satisfies (105).

□

Next we extend Corollary 2(FO MV) to the multidimensional case as follows. We first let define

$$\mathsf{b}_{\sigma\sigma} := \mathcal{M}\cdot\mathsf{a}_{\sigma\sigma},$$

$$\mathsf{b}_{t,k} := \mathcal{M}\cdot\mathsf{a}_{t,k} - \mathcal{L}_{Z,t}\cdot\hat{\omega}_k - \mathcal{L}_{Z,k}\cdot\hat{\omega}_t,$$

$$\mathsf{c}_{t,k} = \mathcal{M}\cdot(\overline{a}_{Z,t},\overline{a}_{Z,k}) - \mathcal{L}\cdot(\hat{\omega}_t,\overline{a}_{Z,k}) - \mathcal{L}\cdot(\hat{\omega}_k,\overline{a}_{Z,t})^T,$$

where $\mathsf{b}_{\sigma\sigma}$ and $\mathsf{b}_{t,k}$ are both vector valued while $\mathsf{c}_{t,k}$ is matrix valued. The recursive LoMs, (104) and (105), can be written more succinctly as

$$\hat{\omega}_{\sigma\sigma,t+1} = \mathcal{L}\cdot\hat{\omega}_{\sigma\sigma,t} - \sum_{s=0}^{\infty}\mathcal{M}\cdot\mathsf{a}_s\overline{Y}_{\sigma\sigma,t+s} - \mathsf{b}_{\sigma\sigma} \quad (106)$$

$$\hat{\omega}_{t+1,k+1} = \mathcal{L}\cdot\hat{\omega}_{t,k} - \sum_{s=0}^{\infty}\mathcal{M}\cdot\mathsf{a}_s\overline{Y}_{ZZ,t+s,k+s} - \mathsf{b}_{t,k} + \nabla_a\cdot\mathsf{c}_{t,k}. \quad (107)$$

In a similar manner we extend the operators

$$\mathcal{L}^{(aa)} \cdot \mathsf{y}(a', \theta') = \int \overline{\Lambda}(a', \theta', a, \theta)\overline{a}_{aa}(a, \theta) \cdot \mathsf{y}(a, \theta) da d\theta$$

$$\mathcal{L}^{(a,a)} \cdot \mathsf{y}(a', \theta') = \int \overline{\Lambda}(a', \theta', a, \theta)\overline{a}_{a}(a, \theta)\mathsf{y}(a, \theta)\overline{a}_a(a, \theta)^T da d\theta$$

$$\mathcal{I}^{(aa)} \cdot \mathsf{y}(a', \theta') = \int \overline{x}_{aa}(a, \theta) \cdot \mathsf{y}(a, \theta) da d\theta$$

with $\overline{x}_{aa}(a, \theta) \cdot \mathsf{y}(a, \theta) := \sum_{i,j} \overline{x}_{a^i a^j}(a, \theta)\mathsf{y}^{ij}(a, \theta)$ for matrix valued y.

Corollary 2(SO MV). *For all t,*

$$\int \overline{x} d\hat{\Omega}_{\sigma\sigma,t} = \sum_{s=0}^{\infty} (\mathcal{I} \cdot \mathsf{A}_{t,s}) \overline{Y}_{\sigma\sigma,s} + \mathcal{I} \cdot \mathsf{B}_{\sigma\sigma,t},$$

where $\{\mathsf{B}_{\sigma\sigma,t}\}_t$ follows a recursion $\mathsf{B}_{\sigma\sigma,t+1} = \mathsf{b}_{\sigma\sigma} + \mathcal{L} \cdot \mathsf{B}_{\sigma\sigma,t}$; and

$$\int \overline{x} d\hat{\Omega}_{t,k} = \sum_{s=0}^{\infty} (\mathcal{I} \cdot \mathsf{A}_{t,s}) \overline{Y}_{s,k-t+s} + \mathcal{I} \cdot \mathsf{B}_{t,k} + \mathcal{I}^{(aa)} \cdot \mathsf{C}_{t,k},$$

where $\{\mathsf{B}_{t,k}, \mathsf{C}_{t,k}\}_{t,k}$ follow recursions

$$\mathsf{C}_{t+1,k+1} = \mathsf{c}_{t,k} + \mathcal{L}^{(a,a)} \cdot \mathsf{C}_{t,k},$$

$$\mathsf{B}_{t+1,k+1} = \mathsf{b}_{t,k} + \mathcal{L} \cdot \mathsf{B}_{t,k} + \mathcal{L}^{(aa)} \cdot \mathsf{C}_{t,k}.$$

Proof. Starting with $\hat{\omega}_{\sigma\sigma,0}$ and rolling forward equation (104) implies

$$\hat{\omega}_{\sigma\sigma,t} = -\sum_{s=0}^{\infty} \mathsf{A}_{t,s}\overline{Y}_{\sigma\sigma,s} - \mathsf{B}_{\sigma\sigma,t}$$

Using integration by parts implies

$$\int \overline{x} d\hat{\Omega}_{\sigma\sigma,t} = \int \overline{x}\nabla_a \cdot \hat{\omega}_{\sigma\sigma,t} da d\theta = -\int \overline{x}_a \hat{\omega}_{\sigma\sigma,t} da d\theta = \mathcal{I} \cdot \hat{\omega}_{\sigma\sigma,t}.$$

Combining these two equations gives us our first result.

For the second half of the proof we first note the following. For any matrix valued density c with compact support

$$\begin{aligned}
\mathcal{L} \cdot \nabla_a \cdot \mathsf{c}(a', \theta') &= \sum_{i,j} \int \overline{\Lambda}(a', \theta', a, \theta)\overline{a}_{a^j}(a, \theta)\frac{\partial}{\partial a^i}\mathsf{c}^{ij}(a, \theta) da d\theta \\
&= -\sum_{i,j} \int \overline{\Lambda}(a', \theta', a, \theta)\overline{a}_{a^i a^j}(a, \theta)\mathsf{c}^{ij}(a, \theta) da d\theta \\
&\quad -\sum_{i,j} \int \frac{\partial}{\partial a^i}\left(\overline{\Lambda}(a', \theta', a, \theta)\right)\overline{a}_{a^j}(a, \theta)\mathsf{c}^{ij}(a, \theta) da d\theta \\
&= -\mathcal{L}^{(aa)} \cdot \mathsf{c}(a', \theta') + \sum_{i,j,k} \frac{\partial}{\partial a'^k} \int \overline{\Lambda}(a', \theta', a, \theta)\overline{a}^k_{a^i}\overline{a}_{a^j}(a, \theta)\mathsf{c}^{ij}(a, \theta) da d\theta \\
&= -\mathcal{L}^{(aa)} \cdot \mathsf{c}(a', \theta') + \nabla_a \cdot \mathcal{L}^{(a,a)} \cdot \mathsf{c}(a', \theta').
\end{aligned}$$

We can then proceed by induction. For $t = 0$ we have

$$\hat{\omega}_{0,t-k} = -\sum_{s=0}^{\infty} \mathsf{A}_{0,s}\overline{Y}_{ZZ,s,k-t+s} - \mathsf{B}_{0,k-t} + \nabla_a \cdot \mathsf{C}_{0,k-t}$$

since all terms are 0. If it holds for t then

$$\begin{aligned}
\hat{\omega}_{t+1,k+1} &= \mathcal{L}^{(a)} \cdot \hat{\omega}_{t,k} - \sum_{s=0}^{\infty} \mathcal{M} \cdot \mathsf{a}_s \overline{Y}_{ZZ,t+s,k+s} - \mathsf{b}_{t,k} + \nabla_a \cdot \mathsf{c}_{t,k} \\
&= \mathcal{L}^{(a)} \cdot \left(-\sum_{s=0}^{\infty} \mathsf{A}_{t,s}\overline{Y}_{ZZ,s,k-t+s} - \mathsf{B}_{t,k} + \nabla_a \cdot \mathsf{C}_{t,k} \right) - \sum_{s=0}^{\infty} \mathcal{M} \cdot \mathsf{a}_s \overline{Y}_{ZZ,t+s,k+s} \\
&\quad - \mathsf{b}_{t,k} + \nabla_a \cdot \mathsf{c}_{t,k} \\
&= -\sum_{s=0}^{\infty} \left(\mathcal{L}^{(a)} \cdot \mathsf{A}_{t,s} + \mathsf{a}_{s-t} \right) \overline{X}_{ZZ,s,k-t+s} - \left(\mathcal{L}^{(a)} \cdot \mathsf{B}_{t,k} + \mathsf{b}_{t,k} \right) + \mathcal{L}^{(a)} \cdot \nabla_a \cdot \mathsf{C}_{t,k} + \nabla_a \cdot \mathsf{c}_{t,k} \\
&= -\sum_{s=0}^{\infty} \mathsf{A}_{t+1,s}\overline{Y}_{ZZ,s,k-t+s} - \left(\mathcal{L}^{(a)} \cdot \mathsf{B}_{t,k} + \mathcal{L}^{(aa)} \cdot \mathsf{C}_{t,k} + \mathsf{b}_{t,k} \right) + \nabla_a \cdot \left(\mathcal{L}^{(a,a)} \cdot \mathsf{C}_{t,k} + \mathsf{c}_{t,k} \right) \\
&= -\sum_{s=0}^{\infty} \mathsf{A}_{t+1,s}\overline{Y}_{ZZ,s,k-t+s} - \mathsf{B}_{t+1,k+1} + \nabla_a \cdot \mathsf{C}_{t+1,k+1}.
\end{aligned}$$

Finally, we have that

$$\begin{aligned}
\int \overline{x} d\hat{\Omega}_{t,k} &= \int \overline{x} \nabla_a \cdot \hat{\omega}_{t,k} da d\theta \\
&= -\int \overline{x}_a \hat{\omega}_{t,k} da d\theta \\
&= -\int \overline{x}_a \left(-\sum_{s=0}^{\infty} \mathsf{A}_{t,s}\overline{Y}_{ZZ,s,k-t+s} - \mathsf{B}_{t,k} + \nabla_a \cdot \mathsf{C}_{t,k} \right) da d\theta \\
&= \sum_{s=0}^{\infty} \left(\mathcal{I}^{(a)} \cdot \mathsf{A}_{t,s} \right) \overline{Y}_{ZZ,s,k-t+s} + \mathcal{I}^{(a)} \cdot \mathsf{B}_{t,k} + \int \overline{x}_{aa} \cdot \mathsf{C}_{t,k} da d\theta \\
&= \sum_{s=0}^{\infty} \left(\mathcal{I}^{(a)} \cdot \mathsf{A}_{t,s} \right) \overline{Y}_{ZZ,s,k-t+s} + \mathcal{I}^{(a)} \cdot \mathsf{B}_{t,k} + \mathcal{I}^{(aa)} \cdot \mathsf{C}_{t,k}
\end{aligned}$$

as desired. □

Combining all of these insights yields the multivariate extension of the Proposition 1b^{SO}

Proposition 1(SO). *$\left\{\overline{X}_{ZZ,t,k}\right\}_{t,k}$ and $\left\{\overline{X}_{\sigma\sigma,t}\right\}_t$ are the solutions to linear systems*

$$\mathsf{G}_x \sum_{s=0}^{\infty} \mathsf{J}_{t,s}\overline{Y}_{\sigma\sigma,s} + \mathsf{G}_x \mathsf{H}_{\sigma\sigma,t} + \mathsf{G}_Y \overline{Y}_{\sigma\sigma,t} = 0, \tag{108}$$

and

$$\mathsf{G}_x \sum_{s=0}^{\infty} \mathsf{J}_{t,s}\overline{Y}_{ZZ,s,k-t+s} + \mathsf{G}_x \mathsf{H}_{t,k} + \mathsf{G}_X \overline{Y}_{ZZ,t,k} + \mathsf{G}_{\Theta,t,k} = 0. \tag{109}$$

where $\mathsf{H}_{\sigma\sigma,t} = \int \mathsf{x}_{\sigma\sigma} d\Omega^* + \mathcal{I} \cdot \mathsf{B}_{\sigma\sigma,t}$ *and* $\mathsf{H}_{t,k} = \int \mathsf{x}_{t,k} d\Omega^* - \mathcal{I}^{(a)}_{Z,t} \cdot \hat{\omega}_k - \mathcal{I}^{(a)}_{Z,k} \cdot \hat{\omega}_t + \mathcal{I} \cdot \mathsf{B}_{t,k} + \mathcal{I}^{(aa)} \cdot \mathsf{C}_{t,k}$.

D Comparison to Approximating the Distribution with A Histogram

All the terms in this section will implicitly index everything by h: the space between points along each dimension a and θ. We let $a_{[i]}$ be the gridpoints along the a dimension and $\theta_{[j]}$ be the grid points along the asset dimension. To construct the histogram approach we define projection function

$$\mathcal{P}^{i,j}(a,\theta)$$

be the probability of assigning point a,θ to gridpoint $(a,\theta)_{[i,j]}$. Following Young (2010) we project to the closest neighbors:

$$\mathcal{P}^{i,j}(a,\theta) = \mathcal{P}^{i}(a)\mathcal{Q}^{j}(\theta)$$

where

$$\mathcal{P}^{i}(a) = \begin{cases} \frac{a-a_{[i]}}{h} & a \in [a_{[i]}, a_{[i+1]}] \\ \frac{a_{[i]}-a}{h} & a \in [a_{[i-1]}, a_{[i]}] \\ 0 & \text{otherwise} \end{cases}$$

and similarly for $\mathcal{Q}^{j}(\theta)$.

We assume full knowledge of $\tilde{x}(a,\theta,Z)$ and focus purely on the approximation with respect to the histogram. The approximation to the steady state transition density is

$$\overline{\Lambda}(i',j',a,\theta) = \int \mathcal{P}^{i',j'}\left(\overline{a}(a,\theta), \rho_\theta\theta + \varepsilon\right) d\mu(\epsilon)$$

This constructs a steady state transition matrix

$$\overline{\Lambda}(i',j',i,j) = \overline{\Lambda}(i',j',a_{[i]},\theta_{[j]})$$

We let $\overline{\omega}_{[i,j]}$ be the approximation to the steady state density. We assume that all of these objects are well approximated as $h \to 0$ so for any smooth test function $\phi(a,\theta)$

$$\int \phi d\Omega^* = \lim_{h\to 0} \sum_{i,j} \phi\left(a_{[i]},\theta_{[j]}\right) \overline{\omega}_{[i,j]}$$

and

$$\int \phi(a',\theta')\overline{\Lambda}(a',\theta',a,\theta)da'd\theta' = \lim_{h\to 0} \sum_{i',j'} \phi\left(a_{[i']},\theta_{[j']}\right) \overline{\Lambda}(i',j',a,\theta).$$

Given $\tilde{x}(a,\theta,Z)$, the approximated LOM for the distribution is

$$\tilde{\omega}_{[i',j']}(Z) = \sum_{i,j} \int \mathcal{P}^{i',j'}\left(\tilde{a}(a_{[i]},\theta_{[j]},Z), \rho_\theta\theta + \varepsilon\right) d\mu(\epsilon)\omega_{[i,j]}$$

Differentiating with respect to Z in direction $\hat{Z}$ yields

$$\begin{aligned} \overline{\omega}_{Z,[i',j']} \cdot \hat{Z} = &\sum_{i,j} \overline{\Lambda}\left(i',j',i,j\right) \hat{\omega}_{[i,j]} \\ &+ \sum_{i,j} \int \mathcal{P}_a^{i',j'}\left(\overline{a}(a_{[i]},\theta_{[j]}), \rho_\theta\theta + \varepsilon\right) d\mu(\epsilon) \overline{a}_Z\left(a_{[i]},\theta_{[j]}\right) \cdot \hat{Z}\overline{\omega}_{[i,j]} \end{aligned}$$

or

$$\hat{\omega}_{t+1,[i',j']} = \sum_{i,j} \overline{\Lambda}\left(i',j',i,j\right)\hat{\omega}_{t,[i,j]} + \sum_{i,j}\int \mathcal{P}_a^{i',j'}\left(\overline{a}(a_{[i]},\theta_{[j]}),\rho_\theta\theta+\varepsilon\right)d\mu(\epsilon)\overline{a}_{Z,t}\left(a_{[i]},\theta_{[j]}\right)\overline{\omega}_{[i,j]}$$

Which we can write succinctly as

$$\hat{\omega}_{t+1} = \overline{\Lambda}\hat{\omega}_t + \mathcal{M}^h\overrightarrow{a}_{Z,t}$$

where $\overrightarrow{a}_{Z,t}$ is $\overline{a}_{Z,t}$ evaluated at the grid-points. Our first claim is

Claim 15. In the limit as $h \to 0$,

$$\lim_{h\to 0}\sum_{i',j'}\overline{x}\left(a_{[i']},\theta_{[j']}\right)\left(\mathcal{M}^h\overrightarrow{a}_{Z,t}\right)_{[i',j']} = \mathcal{I}^{(a)}\cdot\mathcal{M}\cdot\overline{a}_{Z,t}.$$

Proof. Note that

$$\begin{aligned}\sum_{i',j'}\overline{x}\left(a_{[i']},\theta_{[j']}\right)\left(\mathcal{M}^h\overrightarrow{a}_{Z,t}\right)_{[i',j']} &= \sum_{i',j'}\overline{x}(a_{[i']},\theta_{[j']})\sum_{i,j}\int \mathcal{P}_a^{i',j'}\left(\overline{a}(a_{[i]},\theta_{[j]}),\rho_\theta\theta+\varepsilon\right)d\mu(\epsilon)\overline{a}_{Z,t}\left(a_{[i]},\theta_{[j]}\right)\overline{\omega}_{[i,j]} \\ &= \sum_{i,j}\int\sum_{i',j'}\overline{x}(a_{[i']},\theta_{[j']})\mathcal{P}_a^{i',j'}\left(\overline{a}(a_{[i]},\theta_{[j]}),\rho_\theta\theta+\varepsilon\right)d\mu(\epsilon)\overline{a}_{Z,t}\left(a_{[i]},\theta_{[j]}\right)\overline{\omega}_{[i,j]}\end{aligned}$$

This simplifies as

$$\begin{aligned}\sum_{i',j'}\overline{x}(a_{[i']},\theta_{[j']})\mathcal{P}_a^{i',j'}\left(\overline{a}(a_{[i]},\theta_{[j]}),\rho_\theta\theta+\varepsilon\right) &= \overline{x}_a(\overline{a}(a_{[i]},\theta_{[j]}),\rho_\theta\theta+\varepsilon)+\mathcal{O}\left(h\right) \\ &= \sum_{i',j'}\overline{x}_a(a_{[i']},\theta_{[j']})\mathcal{P}^{i',j'}\left(\overline{a}(a_{[i]},\theta_{[j]}),\rho_\theta\theta+\varepsilon\right)+\mathcal{O}\left(h\right)\end{aligned}$$

and thus

$$\sum_{i',j'}\overline{x}\left(a_{[i']},\theta_{[j']}\right)\left(\mathcal{M}^h\overrightarrow{a}_{Z,t}\right)_{[i',j']} = \sum_{i',j'}\sum_{i,j}\overline{x}_a(a_{[i']},\theta_{[j']})\overline{\Lambda}\left(i',j',i,j\right)\overline{a}_{Z,t}\left(a_{[i]},\theta_{[j]}\right)\overline{\omega}_{[i,j]}$$

Taking limit as $h \to 0$ completes the result. □

Next we claim

Claim 16. In the limit as $h \to 0$

$$\lim_{h\to 0}\sum_{i',j'}\overline{x}\left(a_{[i']},\theta_{[j']}\right)\left(\overline{\Lambda}\mathcal{M}^h\overrightarrow{a}_{Z,t}\right)_{[i',j']} = \mathcal{I}^{(a)}\cdot\mathcal{L}^{(a)}\cdot\mathcal{M}\cdot\overline{a}_{Z,t}.$$

Proof. For this we're going to use that for any smooth function $\phi(a)$

$$\begin{aligned}\sum_{i'',i'}\phi(a_{[i'']})\mathcal{P}^{i''}(\overline{a}(a_{[i']},\theta))\mathcal{P}_a^{i'}\left(a\right) &= \frac{1}{h}\sum_{i''}\phi(a_{[i'']})\left(\mathcal{P}^{i''}(\overline{a}(a_{[\hat{i}+1]},\theta))-\mathcal{P}^{i''}(\overline{a}(a_{[\hat{i}]},\theta))\right) \\ &= \frac{1}{h}\sum_{i''}\phi(a_{[i'']})\left(\mathcal{P}^{i''}(\overline{a}(a_{[\hat{i}]}+h,\theta))-\mathcal{P}^{i''}(\overline{a}(a_{[\hat{i}]},\theta))\right) \\ &= \phi_a(\overline{a}(a,\theta))\overline{a}_a(a,\theta)+\mathcal{O}(h)\end{aligned}$$

$$\frac{1}{h}\sum_{i'}\phi(a_{[i']})\left(\mathcal{P}^{i'}(\overline{a}(a+h,\theta))-\mathcal{P}^{i'}(\overline{a}(a,\theta))\right)=\phi_a(\overline{a}(a,\theta))\overline{a}_a(a,\theta)+\mathcal{O}(h).$$

We then have that

$$\sum_{i',j'}\overline{x}\left(a_{[i']},\theta_{[j']}\right)\left(\overline{\Lambda}\mathcal{M}^h\overrightarrow{a}_{Z,t}\right)_{[i',j']}=\sum_{i'',j'',i',j'}\overline{x}(a_{[i'']},\theta_{[j'']})\overline{\Lambda}(i'',j'',i',j')\sum_{i,j}\int\mathcal{P}_a^{i',j'}\left(\overline{a}(a_{[i]},\theta_{[j]}),\rho_\theta\theta+\varepsilon\right)d\mu(\epsilon)\overline{a}_{Z,t}\left(a_{[i]},\theta_{[j]}\right.$$

We can then exploit the fact that

$$\overline{\Lambda}(i'',j'',i',j')=\int\mathcal{Q}^{j''}\left(\rho_\theta\theta_{[j']}+\varepsilon\right)d\mu(\epsilon)\mathcal{P}^{i''}\left(\overline{a}(a_{[i']},\theta_{[j']})\right)$$

to get

$$\sum_{i',j'}\overline{x}\left(a_{[i']},\theta_{[j']}\right)\left(\overline{\Lambda}\mathcal{M}^h\overrightarrow{a}_{Z,t}\right)_{[i',j']}=\sum_{i'',j'',i',j'}\overline{x}_a(a_{[i'']},\theta_{[j'']})\overline{\Lambda}(i'',j'',i',j')\overline{a}_a(a_{[i']},\theta_{[j']})\overline{\Lambda}(i'',j'',i',j')\overline{a}_{Z,t}\left(a_{[i]},\theta_{[j]}\right)\overline{\omega}_{[i,j]}+\mathcal{O}($$

which in the limit as $h\to 0$ gives

$$\mathcal{I}^{(a)}\cdot\mathcal{L}^{(a)}\cdot\mathcal{M}\cdot\overline{a}_{Z,t}$$

□

This same argument extends to show that

$$\lim_{h\to 0}\sum_{i',j'}\overline{x}\left(a_{[i']},\theta_{[j']}\right)\left(\overline{\Lambda}^t\mathcal{M}^h\overrightarrow{a}_{Z,t}\right)_{[i',j']}=\mathcal{I}^{(a)}\cdot\left(\mathcal{L}^{(a)}\right)^t\cdot\mathcal{M}\cdot\overline{a}_{Z,t}$$

for arbitrary t

D.1 Second Order

Taking a second derivative we have and exploiting that $\mathcal{P}_{aa}^{i,j}=0$

$$\begin{aligned}\hat{\omega}_{ZZ,t+1,k+1,[i',j']}\cdot&=\sum_{i,j}\overline{\Lambda}\left(i',j',i,j\right)\hat{\omega}_{ZZ,t,k,[i,j]}+\sum_{i,j}\int\mathcal{P}_a^{i',j'}\left(\overline{a}(a_{[i]},\theta_{[j]}),\rho_\theta\theta+\varepsilon\right)d\mu(\epsilon)\overline{a}_{Z,t}\left(a_{[i]},\theta_{[j]}\right)\hat{\omega}_{Z,k,[i,j]}\\&+\sum_{i,j}\int\mathcal{P}_a^{i',j'}\left(\overline{a}(a_{[i]},\theta_{[j]}),\rho_\theta\theta+\varepsilon\right)d\mu(\epsilon)\overline{a}_{Z,k}\left(a_{[i]},\theta_{[j]}\right)\hat{\omega}_{Z,t,[i,j]}\\&+\sum_{i,j}\int\mathcal{P}_a^{i',j'}\left(\overline{a}(a_{[i]},\theta_{[j]}),\rho_\theta\theta+\varepsilon\right)d\mu(\epsilon)\overline{a}_{ZZ,t,k}\left(a_{[i]},\theta_{[j]}\right)\overline{\omega}_{[i,j]}\end{aligned}$$

For $t=k=0$ we have

$$\hat{\omega}_{ZZ,1,1,[i',j']}=\sum_{i,j}\int\mathcal{P}_a^{i',j'}\left(\overline{a}(a_{[i]},\theta_{[j]}),\rho_\theta\theta+\varepsilon\right)d\mu(\epsilon)\overline{a}_{ZZ,0,0}\left(a_{[i]},\theta_{[j]}\right)\overline{\omega}_{[i,j]}.$$

Thus in the limit as $h\to 0$ we have

$$\lim_{h\to 0}\sum_{i',j'}\overline{x}(a_{[i']},\theta_{[j']})\hat{\omega}_{ZZ,1,1,[i',j']}=\mathcal{I}^{(a)}\cdot\mathcal{M}\cdot\overline{a}_{ZZ,0,0}=\mathcal{I}^{(a)}\cdot\mathsf{B}_{1,1}\neq\int\overline{x}d\hat{\Omega}_{1,1}$$

as it is missing the $\mathcal{I}^{(aa)}\cdot\mathsf{C}_{1,1}$ term.

E Details for Section 7

E.1 Krusell and Smith with adjustment costs

Household problem Households hold shares in a mutual fund and date t holdings of i denoted by $a_{i,t}$. Let D_t and P_t be the time t dividend and the price per share of the mutual fund. The household problem is given by

$$\max_{c_{i,t},i_{i,t},k_{i,t}} \mathbb{E}\sum_t \beta^t U\left(c_{i,t}\right)$$

subject to

$$c_{i,t} + P_t a_{i,t} = w_t e^{\theta_{i,t}} + \left(D_t + P_t\right) a_{i,t-1}$$

$$a_{i,t} \geq 0$$

The Euler equation of the household is given by

$$1 = \mathbb{E}_t \left(\frac{\beta U_c\left(c_{i,t+1}\right)}{U_c\left(c_{i,t}\right) + \zeta_{i,t}}\right)\left(\frac{D_{t+1} + P_{t+1}}{P_t}\right) \tag{110}$$

where $\zeta_{i,t} \geq 0$ is the Lagrange multiplier on the borrowing constraint.

Stochastic Discount Factors Define a process $\{m_{i,t}\}$ with $m_{i,0} = 1$ and $\frac{m_{i,t+1}}{m_{i,t}} \equiv \frac{\beta U_c(c_{i,t+1})}{U_c(c_{i,t})+\zeta_{i,t}}$. For any positive process $\{o_{i,t}\}$ define M_t with $M_0 = 1$ and $\frac{M_{t+1}}{M_t} = \int o_{i,t}\frac{m_{i,t+1}}{m_{i,t}}di$. Then aggregating (110) we get that the value of the mutual fund satisfies

$$P_t = \mathbb{E}_t \sum_j \frac{M_{t+j}}{M_t} D_{t+j}.$$

Mutual Fund Problem The mutual fund owns physical capital, makes investments subject to quadratic adjustment costs, rents out the capital to the corporate sector, and maximizes present value of dividends. For a given $\{M_t\}$, the problem of the mutual fund is

$$\max_{K_t,D_t} \mathbb{E}_0 \sum_t M_t D_t$$

$$D_t = r_t^k K_t - I_t - \frac{\phi}{2}\left(\frac{I_t}{K_t} - \delta\right)^2 K_t$$
$$K_{t+1} = \left(1-\delta\right) K_t + I_t$$

Let Q_t be the multiplier on the capital accumulation equation. The optimality of the mutual fund with respect to I_t

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

and with respect to K_t is

$$\mathbb{E}_t \frac{M_{t+1}}{M_t} \left\{ \frac{r^k_{t+1} + \phi\left(\frac{I_{t+1}}{K_{t+1}} - \delta\right)\frac{I_{t+1}}{K_{t+1}} - \frac{\phi}{2}\left(\frac{I_{t+1}}{K_{t+1}} - \delta\right)^2 + (1-\delta)Q_{t+1}}{Q_t} \right\} = 1. \tag{111}$$

Its easy to check that

$$\frac{r^k_{t+1} + \phi\left(\frac{I_{t+1}}{K_{t+1}} - \delta\right)\frac{I_{t+1}}{K_{t+1}} - \frac{\phi}{2}\left(\frac{I_{t+1}}{K_{t+1}} - \delta\right)^2 + (1-\delta)Q_{t+1}}{Q_t} = \frac{D_{t+1} + Q_{t+1}K_{t+2}}{Q_t K_{t+1}}$$

and thus iterating on (111) we get

$$Q_t K_{t+1} = P_t = \mathbb{E}_t \sum_j \frac{M_{t+j}}{M_t} D_{t+j}.$$

Equilibrium The equilibrium is given by

$$c_{i,t} + P_t a_{i,t} = w_t e^{\theta_{i,t}} + R_t P_{t-1} a_{i,t-1} \tag{112a}$$

$$1 = \mathbb{E}_t \left(\frac{\beta U_c\left(c_{i,t+1}\right)}{U_c\left(c_{i,t}\right) + \zeta_{i,t}} \right) \left(\frac{D_{t+1} + P_{t+1}}{P_t} \right) \tag{112b}$$

$$a_{i,t}\zeta_{i,t} = 0 \tag{112c}$$

$$W_t - (1-\alpha)\exp\left(\Theta_t\right) K_t^\alpha = 0, \tag{112d}$$

$$R_t = \frac{1-\alpha)\exp\left(\Theta_t\right) K_t^\alpha - I_t - \frac{\phi}{2}\left(\frac{I_t}{K_t} - \delta\right)^2 K_t + P_{t+1}}{P_t}, \tag{112e}$$

$$Q_t = 1 + \phi\left(\frac{I_t}{K_{t-1}} - \delta\right), \tag{112f}$$

$$P_t = Q_t K_{t+1}, \tag{112g}$$

$$\int a_{i,t} di = 1. \tag{112h}$$

Define $k_{i,t} = P_{t-1}a_{i,t}$ and substitute for P_t to get the equations in the main text.