APPROXIMATING EQUILIBRIA WITH EX-POST HETEROGENEITY AND AGGREGATE RISK

Elisabeth Pröhl

University of Amsterdam

October 10, 2018

Dynamic stochastic general equilibrium models with ex-post heterogeneity due to idiosyncratic risk have to be solved numerically. This is a nontrivial task as the cross-sectional distribution of endogenous variables becomes an element of the state space due to aggregate risk. Existing global solution methods often assume bounded rationality in terms of a parametric law of motion of aggregate variables in order to reduce dimensionality. I do not make this assumption and tackle dimensionality by polynomial chaos expansions, a projection technique for square-integrable random variables. This approach results in a nonparametric law of motion of aggregate variables. Moreover, I establish convergence of the proposed algorithm to the rational expectations equilibrium. Economically, I find that higher levels of idiosyncratic risk sharing lead to higher systemic risk, i.e., higher volatility within the ergodic state distribution, and second, heterogeneity leads to an amplification of aggregate risk for sufficiently high levels of risk sharing.

Keywords: Dynamic stochastic general equilibrium, Incomplete markets, Heterogeneous agents, Aggregate uncertainty, Convergence, Numerical solutions, Polynomial chaos.

JEL Classification: C62, C63, D31, D52, E21.

Roeterstraat 11, 1018 WB Amsterdam, Netherlands, e.proehl@uva.nl

I thank Johannes Brumm, René Carmona, Chris Carroll, David Childers, Jérôme Detemple, Rajna Gibson Brandon, Daniel Harenberg, Kenneth Judd, Greg Kaplan, Felix Kübler, Semyon Malamud, Benjamin Moll, Jean-Charles Rochet, Juan Sagredo, Yuliy Sannikov, Olivier Scaillet, Karl Schmedders, Harald Uhlig, Thomas Winberry, Victor Zhorin, and participants at the 2018 Tinbergen Macro Research Day, PASC18, RES PhD Meeting 2017, SAEE 2017, ESEWM 2017, 13th Dynare Conference, 6th Lindau Nobel Laureate Meeting, NBER Summer Institute 2017, WEHIA 2017, 2017 Workshop on Recent Developments in Numerical Methods with Applications in Statistics and Finance, SNDE 2017, RGS Doctoral Conference 2017, ESEM 2016, EcoMod 2016, CEF 2016, ESNASM 2016, MFM Summer Camp 2016, 12th Cowles Conference on General Equilibrium and its Applications, ISCEF 2016, MFM Winter Meeting 2016, CEF 2015, SFI Research Days 2015, PASC 2015, and seminar participants at the University of Amsterdam, Bank of Spain, Bonn University, University of Chicago, Collegio Carlo Alberto, Erasmus University, University of Geneva, University of Lausanne/EPFL, Princeton University, Stockholm School of Economics, Warwick Business School and the University of Zurich for fruitful comments on this work. An earlier version of this work was disseminated under the title "Discretizing the Infinite-Dimensional Space of Distributions to Approximate Markov Equilibria with Ex-Post Heterogeneity and Aggregate Risk". Lastly, I would like to acknowledge that this paper was written when I received support by the Macro Financial Modeling Initiative which included a grant funded by the Alfred P. Sloan Foundation, the CME Group Foundation and Fidelity Management and Research.

1. INTRODUCTION

Economies consist of heterogeneous agents who are exposed to various idiosyncratic risks, the most prominent example of which is labor income risk for households. This was first modeled in a dynamic stochastic general equilibrium (DSGE) model by Bewley (1977) where agents face idiosyncratic income shocks affecting their wealth, and, extended by Aiyagari (1994) to include a production technology. They show that individual precautionary savings contribute to aggregate savings because idiosyncratic risk cannot be fully insured. Other examples of idiosyncratic risks are firm-specific productivity shocks in models of firm exit and entry as in Hopenhayn (1992) or uninsurable persistent income shocks as in Constantinides and Duffie (1996), who show that these shock have a strong impact on the equity premium. Lately, there has been a renewed interest in heterogeneous agent models. Many recent studies reevaluate the impact of heterogeneity in the economy and find strong implications. Let me mention just a few examples. Khan and Thomas (2008) model heterogeneity to investigate the effect of non-convex adjustment costs and find that they are important to produce a realistic investment rate distribution. Apart from achieving a realistic wealth distribution using heterogeneous households trading in two assets, Kaplan et al. (2018) investigate the effect of monetary policy on the consumption of households. They find that the indirect effects significantly outweigh the direct ones. In finance, Storesletten et al. (2007) find a moderate effect of idiosyncratic risk on the Sharpe ratio, but a significant negative impact on inter-generational risk sharing. Overall, there is plenty of evidence that idiosyncratic risks have a sizable impact on the economy.

Many of these models do not feature aggregate risk, however, to avoid the corresponding difficulties in solving the model. The challenge in constructing a solution algorithm lies in handling the cross-sectional distribution of the agents' idiosyncratic variables, which becomes an infinite-dimensional element of the state space. Moreover, this distribution changes stochastically over time depending on the realization of the aggregate shocks. The aggregate variables evolve, in turn, depending on how the cross-sectional distribution changes. Models as in Storesletten et al. (2007) and Khan and Thomas (2008) which do include aggregate risk, on the other hand, are typically solved using the Krusell-Smith algorithm.

In their seminal paper, Krusell and Smith (1998) were the first to propose a global solution algorithm for the Aiyagari growth model with aggregate risk. They handle the dimensionality problem in assuming bounded rationality, which means that agents are not required to observe the whole cross-sectional distribution to predict the movement of aggregate variables. They rather use a parametric law of motion for the aggregate variables depending on a finite number of moments. Given that assumption, they then solve the model by iterating on the following two steps: Solving for the optimal policies given a guess of parameters for the aggregate variables' law of motion, and second, estimating new parameters for the law of motion given a set of simulated data from the new optimal policy. The main economic result from this seminal work is that, given the bounded rationality assumption, adding moments higher than the mean to the parametric law of motion does not change the equilibrium solution. Hence, the idiosyncratic risk does not matter for aggregation. Various more recent papers improve the original algorithm mainly by eliminating the agent dimension in the simulation step, and, by varying the parametric form of the law of motion. However, these works still more or less rely on the bounded rationality assumption.

The existing methodology of global solution methods for heterogeneous agent models with aggregate risk has several drawbacks. First, it is not clear whether the assumed parametric law of motion for the aggregates in the bounded rational expectations equilibrium is indeed close to its equivalent in the fully rational expectations equilibrium. Generally, it is unknown whether the bounded rational solution is at all close to the fully rational solution as convergence results are lacking and there is no theory on measuring their distance. Second, it is not clear a priori how many moments are necessary for the bounded rational equilibrium to exist. In fact, Kubler and Schmedders (2002) show that there are models, for which recursive equilibria depending only on aggregate wealth, i.e., the first moment of the cross-sectional distribution, do not exist.

The main contribution of this paper is a novel global solution method for DSGE models featuring both idiosyncratic and aggregate risk. Most importantly, I prove convergence to the theoretical rational expectations equilibrium. What distinguishes this algorithm from existing solution procedures is that it discretizes the space of cross-sectional distributions rather than assuming a parametric law of motion for the aggregate variables. To discretize, I use a projection method called generalized polynomial chaos which extends the polynomial projection of real functions to a projection of square-integrable random variables. It can, hence, be interpreted as a probabilistic polynomial projection. This method has several advantages over standard polynomial projection. First of all, polynomial chaos does not require smoothness and can, therefore, handle distributions with mass points. This is relevant for discrete-time models with borrowing constraints as these models feature mass points in the cross-sectional distribution. Moreover, the location of these mass points is endogenous such that one cannot simply treat them separately. Another advantage is that polynomial chaos converges quite fast. Lastly, by approximating the full distribution, the aggregate variables emerge automatically in a nonparametric fashion. Therefore, I do not require a separate step in the solution algorithm to estimate their law of motion. No simulation is necessary.

To illustrate the advantages of the proposed algorithm numerically, I compute two example models: The Aiyagari-Bewley model with aggregate risk which has become a benchmark model in the computational economics literature, and, the Huggett model with aggregate risk which in contrast to the Aiyagari-Bewley economy does not feature approximate aggregation. When comparing the numerical solution of the proposed algorithm to existing methods, I find a significant improvement in precision for individual policies in terms of Euler equation errors as well as a significant improvement in the prediction of the law of motion of aggregate variables. Solutions with sufficient precision are achieved by approximations of order two in the case of the Aiyagari-Bewley economy and approximations of order three in the case of the Huggett economy. Note that an order zero approximation leads to optimal policies, which depend solely on the mean of the distribution, whereas, order one and higher lead to policies, which depend on an approximation of the whole cross-sectional distribution. The fact that orders higher than zero are needed in the example models implies that idiosyncratic risk matters in the respective rational expectations equilibrium. Nevertheless, it is possible to retrace why Krusell and Smith (1998) obtain their approximate aggregation result for the Aiyagari-Bewley model. An alternative definition of approximate aggregation can be given by measuring the change in the expected ergodic cross-sectional capital distribution when the approximation order increases. This change is insignificant for the benchmark calibration of the Aiyagari-Bewley economy in Krusell and Smith (1998). Surprisingly, approximate aggregation does not persist in the Aiyagari-Bewley model when a sufficiently high but not unrealistic unemployment benefit is introduced. Furthermore, the Krusell-Smith algorithm fails to converge in that case which confirms the failure of approximate aggregation.

Even though the two example models are standard heterogeneous agent models in growth theory and asset pricing, the newly solutions yield interesting novel economic implications. The first economic result emerges by comparing low and high levels of idiosyncratic risk sharing. I find that the expected stationary distribution of individual capital in the Aiyagari-Bewley model has fatter tails, whereas, the stationary distribution of equilibrium bond prices in the Huggett model has larger volatility the higher the level of idiosyncratic risk sharing. This means that systemic risk in the equilibrium problem increases with risk sharing, a result which recovers for these standard models what was coined the volatility paradox in Brunnermeier and Sannikov (2014). Second, I find that sufficiently high levels of risk sharing of idiosyncratic risk among agents amplifies aggregate risk. The amplification effect is due to the interaction of the borrowing constraint with the heterogeneity and thus, it is stronger in the Huggett model than in the Aiyagari-Bewley model.

This paper relates to several strands of literature. First of all, it is clearly connected to the literature on numerical algorithms. In general, there are two types of algorithms: Local solution methods are based on perturbation techniques, whereas, global solution methods are based on projection techniques or a mixture of projection and simulation techniques. My algorithm and the aforementioned seminal algorithm by Krusell and Smith (1998) belong to the latter group. The algorithm by Krusell and Smith (1998) has also been the subject of a special issue of the Journal of Economic Dynamics and Control in January 2010. This special issue presents various alternative algorithms which are compared in den Haan (2010). They have in common that they rely more or less on the assumption of bounded rationality by using a small finite number of parameters instead of the full cross-sectional distribution to approximate the policy function and the law of motion of aggregate variables. One problem, which is addressed by Algan et al. (2008); Young (2010); Ríos-Rull (1997); den Haan (1997) and summarized in Algan et al. (2010), is the cross-sectional variation due to the simulation of a finite number of agents in Krusell and Smith (1998) when estimating the law of motion parameters. They use parametric and nonparametric procedures to get around this issue. However, the variation due to simulating over aggregate exogenous shocks remains. In contrast to the simulation approach, den Haan and Rendahl (2010) use direct aggregation to obtain the law of motion. Interestingly, Algan et al. (2008) and Reiter (2010a) parameterize the cross-sectional distribution itself to obtain a better prediction of the law of motion, but their parametric functional forms are somewhat ad hoc. They do not span the space of square-integrable random variables. I use the algorithm by Reiter (2010a) in my numerical comparison and find that it performs significantly worse than the algorithm proposed herein.

Local solution methods based on perturbations do not assume bounded rationality. To reduce dimensionality, they first solve for the optimal policy and stationary distribution of the model without aggregate shocks using projection methods, and then, perturb this solution to accommodate aggregate shocks. The most prominent perturbation algorithm goes back to Reiter (2009, 2010b). Childers (2015) investigates the theoretical underpinning of these perturbations. Mertens and Judd (2013) use perturbations for the law of motion, whereas, Boppart et al. (2018) use MIT shocks to approximate the numerical derivatives. Winberry (2018) combines the law of motion approach of Algan et al. (2008) with the perturbation of Reiter (2009) to take the cross-sectional distribution into account. He also presents a model where the aggregation result by Krusell and Smith (1998) does not hold. There are two major drawbacks for perturbation methods: First, the perturbation in aggregate shocks is often only linear, or at most quadratic. Therefore, any higher-order nonlinear effects of aggregate shocks like risk are not accounted for. Second, as for all perturbation methods, the solutions are only accurate for small aggregate shocks. Crises scenarios consisting of a large aggregate shock or a long series of aggregate shocks in one direction cannot be analyzed with confidence.

It is also worth pointing out the relation to the literature on mean field games and their numerical solutions because they are essentially continuous-time versions of DSGE models with ex-post heterogeneity. Achdou et al. (2017) show how to use partial differential equations to solve heterogeneous agent models. Kaplan et al. (2018) put forward a very interesting application of this methodology to monetary policy questions. However, their models incorporate only idiosyncratic shocks without aggregate risk. Applying generalized polynomial chaos, as in the algorithm presented herein, to extend their framework to aggregate risk could yield interesting results.

The paper proceeds as follows. In the next section, I present the two illustrative example models. In Section 3, I introduce the algorithm and its convergence result. Section 4 evaluates the algorithm and its approximation error numerically. In Section 5, I investigate whether approximate aggregation holds in the example models. Lastly, I analyze which economic insights this novel computational method yields. The appendix contains all proofs. The online appendix contains supplementary information and robustness checks. The code and other supplementary material can be found on https://github.com/zschorli/Proehl_SolvingHeterogAgentModels.

2. EXAMPLE MODELS

To demonstrate the numerical method developed in this paper, I use two illustrative models featuring both idiosyncratic as well as aggregate risk. First, I use the Aiyagari-Bewley growth model with aggregate risk from Krusell and Smith (1998) as this model has become a benchmark test case. More specifically, I rely on the model description in den Haan et al. (2010), which was used for a comparison of Krusell-Smith-style algorithms in the special issue of the Journal of Economic Dynamics and Control of January 2010. As this model is widely seen to lack a sufficient effect of the heterogeneity on aggregate variables and to illustrate the flexibility of my method, I additionally consider the Huggett (1993) economy where the effect of heterogeneity turns out to be stronger. I first introduce the general structure which is common to both models.

2.1. *Generic Model Structure*

Consider a discrete-time infinite-horizon models with a continuum of agents of measure one. There are two kinds of exogenous shocks, an aggregate shock and an idiosyncratic shock. The aggregate shock characterizes the state of the economy with outcomes in $\mathcal{Z}^{ag} = \{0, 1\}$ standing for a bad and good state, respectively. Similarly, the idiosyncratic shock with outcomes in $\mathcal{Z}^{id} = \{0, 1\}$ indicates a good or a bad shock for the individual agent. The idiosyncratic shock can be interpreted as an agent being unemployed or employed in the Aiyagari-Bewley growth model or as an agent receiving low or high endowment in the Huggett economy. It is i.i.d. across agents conditional on the aggregate shock. I denote the compound exogenous process $\left(z_t^{ag}, z_t^{id}\right)_{t\geq 0}$ by $(z_t)_{t\geq 0} \in \mathcal{Z}$ with $\mathcal{Z} = \mathcal{Z}^{ag} \times \mathcal{Z}^{id}$. The transition probabilities are exogenously given by a four-by-four matrix.

The market setting consists of the agents' security claims $(x_t)_{t\geq 0}$. The aggregate security holdings $(X_t)_{t\geq 0}$, which are an element of the model specific market-clearing condition, are computed by taking the mean over all agents

$$X_t = \sum_{z^{id}=0}^{1} \int_{-\infty}^{\infty} x d\mu_t\left(z^{id}, x\right) \; \forall\, t \geq 0, \tag{1}$$

where μ_t is the cross-sectional distribution of idiosyncratic exogenous and endogenous variables at the beginning of time t, i.e. before the agents choose their optimal security holdings. It is simply the probability distribution of individual security holdings across agents given the trajectory of aggregate shocks

$$\mu_t\left(z^{id}, x\right) = \mathbb{P}\left(\left\{z_t^{id} = z^{id}\right\} \cap \{x_{t-1} \leq x\} \,\middle|\, z_t^{ag}, \ldots, z_0^{ag}\right) \tag{2}$$

for all $t \geq 0$, $z^{id} \in \mathcal{Z}^{id}$ and $x \in \mathbb{R}$.[1] The aggregate shocks cause the cross-sectional

[1] Note that we can use the methodology of Fubini extension by Sun (2006) to ensure the validity of the law of large numbers when aggregating over the set of agents. In particular, let us denote the atomless measure space of agents by $(I, \mathcal{I}, \lambda)$ with $\lambda(I) = 1$ and the sample probability space by $(\mathcal{Z}^{id}, \sigma(\mathcal{Z}^{id}), p^{z^{id}|z^{ag}})$. Let f be a measurable function mapping the Fubini extension $(I \times \mathcal{Z}^{id}, \mathcal{I} \boxtimes \sigma(\mathcal{Z}^{id}), \lambda \boxtimes p^{z^{id}|z^{ag}})$ into $\mathbb{R}$. If the random variables $f(i, .)$ are essentially

distribution to vary over time, which is indicated by the time subscript of μ_t.

Each agent chooses her security holdings and consumption such that they satisfy certain constraints. First, individual consumption must be positive at all times $c_t > 0$, $t \geq 0$, and security holdings are subject to a hard borrowing constraint $x_t \geq \bar{x}$, $t \geq 0$, $\bar{x} \leq 0$. Second, given an initial security claim $x_{-1} \geq \bar{x}$ and an initial cross-sectional distribution μ_{-1} with non-negative support, each agent adheres to a budget constraint, which equates individual consumption and current security holdings to the incoming cash flows of the current period and potential savings.

$$k_t + c_t = B\left(z_t, x_{t-1}, X_t\right) \; \forall \, t \geq 0. \tag{3}$$

The right-hand side of the budget constraint denoted by the function B is model specific and is defined for the two example models subsequently. The time line underlying this equation is clarified in Figure 1.[2]

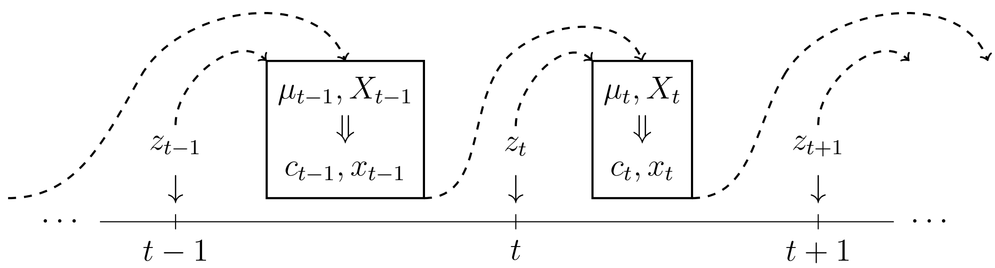

FIGURE 1. **Time line of events.** Before period t, the agent observes how much of the security everybody held in the previous period x_{t-1}. At period t, the agent observes the exogenous shocks z_t, which also determine the beginning-of-period cross-sectional distribution μ_t and hence, the aggregate holdings X_t. The agent then decides how much to consume c_t and how much of the security x_t to hold in that period.

All agents have time-separable CRRA utility with a risk aversion coefficient

---

pairwise independent, then $f(i, .)$ have a common distribution μ for λ-almost all $i \in I$. The same holds for the samples $f(., z^{id})$. When f represents individual security holdings, we get

$$X = \int_I x(i) d\lambda(i) = \sum_{z^{id}=0}^{1} \int_{-\infty}^{\infty} x d\mu\left(z^{id}, x\right).$$

[2]Note that I specify the time line slightly differently than den Haan et al. (2010) and Huggett (1993). These authors substitute x_t with x_{t+1} in the budget constraint because the security pays out in the next period. In the Aiyagari-Bewley economy it is the capital which is put forward as start capital to period $t+1$, whereas, in the Huggett economy it is the claim to a one period bond which pays out one unit in $t+1$. In contrast to that notation, however, I want to emphasize the time period, at which the agent optimally chooses the magnitude of her security holdings. Taking this view, the optimal consumption and security holdings choice have the same time subscript. My time line therefore indicates, which filtration the endogenous variables are adapted to.

$\gamma > 1$ and time preference parameter $\beta \in (0, 1)$. Then, given an agent's initial security claim $x_{-1} \geq \bar{x}$ and the initial cross-sectional distribution μ_{-1} with support in $[\bar{x}, \infty)$, the individual optimization problem reads

$$
\begin{aligned}
\max_{\{c_t, x_t\} \in \mathbb{R}^2} \quad & \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\gamma} - 1}{1-\gamma}\right] \\
\text{s.t.} \quad & x_t + c_t = B\left(z_t, x_{t-1}, X_t\right) \, \forall \, t \geq 0 \\
& c_t > 0, \; x_t \geq \bar{x} \, \forall \, t \geq 0.
\end{aligned} \tag{4}
$$

In a competitive equilibrium, the individual problems are solved subject to markets clearing. This paper focuses on a competitive equilibrium of recursive form. To define a recursive equilibrium, let us switch to prime-notation for convenience, where a prime denotes variables in the current period and variables with no prime refer to the previous period.

Definition 1 (Recursive equilibrium) *A solution to the agents' individual optimization problems (4) given an initial cross-sectional distribution of individual capital μ_{-1} with non-negative support, which ensures that markets clear, is called recursive if there exist functions $g_i : \mathcal{Z} \times \mathbb{R} \times \mathcal{P}(\mathcal{Z}^{id} \times \mathbb{R}) \to \mathbb{R}$, $i \in \{c, x\}$, such that, for any point in time, the current optimal consumption and security holdings choices equal $c' = g_c(z', x, \mu')$ and $x' = g_x(z', x, \mu')$ for any agent with previous-period holdings x who observes the current-period exogenous shock $z' = (z^{ag'}, z^{id'})$ and the beginning-of-current-period cross-sectional distribution μ'. In models where the security is priced, there additionally exists a function $g_p : \mathcal{Z} \times \mathcal{P}(\mathcal{Z}^{id} \times \mathbb{R}) \to \mathbb{R}$ such that, for any point in time, the equilibrium price equals $p' = g_p(z', \mu')$.*

Remark Note that I will solely work with the investment policy $x' = g_x(z', x, \mu')$ in the following as the consumption policy follows directly due to the budget constraint. The subscript x is dropped for convenience, i.e., I denote $x' = g(z', x, \mu')$.

In order to obtain a full description of equilibrium, we need to define a consistent law of motion of μ' to μ''. Given a fixed distribution μ' over the cross-section of individual capital at the beginning of the current period and a recursive equilibrium, the distribution changes in two steps $\mu' \to \tilde{\mu}' \to \mu''$. In the first step, the agents implement their optimal security holdings, which leads to the end-of-current-period distribution

$$
\tilde{\mu}'\left(z^{id'}, x\right) = \mathbb{P}\left(\left\{\zeta' = z^{id'}\right\} \cap \left\{g\left(z^{ag'}, \zeta', \chi, \mu'\right) \leq x\right\} \middle| \, z^{ag'}\right), \tag{5}
$$

where $(\zeta', \chi) \sim \mu'$ is a random variable distributed according to the cross-sectional distribution. In the second step, the next-period shocks z'' for all agents realize and shift the sizes of the agent groups identified by a particular idiosyncratic shock realization depending on the outcome of the aggregate shock. Agents with a previous positive shock either stay within their group, i.e., they keep being lucky, or they now face a negative shock, the same holds for agents with a previous negative shock. Therefore, the distribution at the beginning of the next period μ'' is given by

$$
\begin{aligned}
(6) \qquad \mu''\left(z^{id''}, x\right) &= \sum_{z^{id'} \in \mathcal{Z}^{id}} \frac{p^{(z^{ag''}, z^{id''})|(z^{ag'}, z^{id'})}}{p^{z^{ag''}|z^{ag'}}} \tilde{\mu}'\left(z^{id'}, x\right) \\
&= \sum_{z^{id'} \in \mathcal{Z}^{id}} \frac{p^{(z^{ag''}, z^{id''})|(z^{ag'}, z^{id'})}}{p^{z^{ag''}|z^{ag'}}} \\
&\quad \cdot \mathbb{P}\left(\left\{\zeta' = z^{id'}\right\} \cap \left\{g\left(z^{ag'}, \zeta', \chi, \mu'\right) \leq x\right\} \middle| z^{ag'}\right)
\end{aligned}
$$

for all $z^{id''} \in \mathcal{Z}^{id}$ and $x \in \mathbb{R}$. The multipliers in front of the end-of-current-period distribution are the probabilities that the idiosyncratic state changes from $z^{id'}$ to $z^{id''}$ given the observed trajectory of $z^{ag'}$ to $z^{ag''}$. From this definition of the new distribution, the new aggregate security holdings X'' follows immediately due to (1).

2.2. *The Aiyagari-Bewley Economy*

The security market consists of a claim to aggregate capital $(K_t)_{t \geq 0}$. An agent's share of physical capital is denoted by $(k_t)_{t \geq 0}$. For the capital market to clear, it thus has to hold that

$$
(7) \qquad K_t = \sum_{z^{id}=0}^{1} \int_{-\infty}^{\infty} k d\mu_t\left(z^{id}, k\right) \forall t \geq 0.
$$

Given an initial capital holding $k_{-1} \geq 0$ and an initial cross-sectional distribution μ_{-1}[3] with zero mean and nonnegative support, each agent adheres to a budget constraint which equates individual consumption and current capital stock to productive income and saved capital stock

$$
(8) \qquad k_t + c_t = B^{\mathrm{AB}}\left(z_t, k_{t-1}, K_t\right) := I\left(z_t, k_{t-1}, K_t\right) + [1-\rho]\, k_{t-1} \forall t \geq 0.
$$

[3] The initial cross-sectional distribution μ_{-1} does not only imply the initial aggregate capital K_{-1}, but also the initial aggregate economic state as it is pinned down by the employment rate $\mathbb{P}(z_{-1}^{id} = 1 | z_{-1}^{ag}) = (1/K_{-1}) \int_0^\infty k d\mu_{-1}(1, k)$.

The parameters in this budget constraint are defined as follows. The capital stock brought forward from period $t-1$ depreciates by a rate $\rho \in (0,1)$. The productive income is given by

$$
\begin{aligned}
(9) \qquad I\left(z_t, k_{t-1}, K_t\right) \quad = \quad & R\left(z_t^{ag}, K_t\right) k_{t-1} \\
& +z_t^{id}\pi\left[1-\tau_t\right] W\left(z_t^{ag}, K_t\right) + \left[1-z_t^{id}\right] \nu W\left(z_t^{ag}, K_t\right).
\end{aligned}
$$

It is composed of, first, the return on capital stock, and second, labor income, which equals the individual's wage W when the agent is employed and a proportional unemployment benefit νW otherwise. The agent's wage is subject to a tax rate $\tau_t = \nu(1-p_t^e)/(\pi p_t^e)$ whose sole purpose it is to redistribute money from the employed to the unemployed. The parameter $\nu \in (0,1)$ denotes the unemployment benefit rate, whereas, $p_t^e = \mathbb{P}(z_t^{id} = 1|z_t^{ag})$ is the employment rate at time t and $\pi > 0$ is a time endowment factor. It is reasonable to assume $\nu/\pi < 1-\tau_t \Leftrightarrow \nu < \pi p_t^e$ for all $t \geq 0$. The wage W and the rental rate R are derived from a Cobb-Douglas production function for the consumption good

$$
\begin{aligned}
W\left(z_t^{ag}, K_t\right) \quad &= \quad (1-\alpha)\left(1+z_t^{ag}a-(1-z_t^{ag})a\right)\left[\frac{K_t}{\pi p_t^e}\right]^{\alpha} \\
R\left(z_t^{ag}, K_t\right) \quad &= \quad \alpha\left(1+z_t^{ag}a-(1-z_t^{ag})a\right)\left[\frac{K_t}{\pi p_t^e}\right]^{\alpha-1},
\end{aligned}
$$

where $a \in (0,1)$ is the absolute aggregate productivity rate and $\alpha \in (0,1)$ is the output elasticity parameter. Labor supply is defined by the employment rate p_t^e scaled by the time endowment factor π.

I make the following technical assumption on the model parameters which ensures existence and uniqueness of a solution according to Pröhl (2018).

Assumption 2 Suppose that $\beta(1-\rho)^{1-\gamma} < 1$.

The model's parameter values used in this paper are set according to the calibration of den Haan et al. (2010) and can be found in Online Appendix OA.1.1.

2.3. *The Huggett Economy*

The security market in this model consists of a risk-free single-period bond which can be bought or sold for a price p and pays out one unit of the consumption good to the buyers one period later. An agent's asset holding is denoted by $(a_t)_{t\geq 0}$. The bond's price at which the security market clears is indirectly defined by the zero-net supply condition

$$
(10) \qquad 0 = \sum_{z^{id}=0}^{1} \int_{-\infty}^{\infty} a d\mu_t\left(z^{id}, a\right) \; \forall \, t \geq 0.
$$

Given an initial bond holding $a_{-1} \geq \bar{a}$ and an initial cross-sectional distribution μ_{-1} with zero mean and support in $[\bar{a}, \infty)$, each agent adheres to a budget constraint, which equates individual consumption and the current value of bond holdings to the agent's current endowment and settlement of the previous period bond holdings

$$(11) \qquad a_t p_t + c_t = B^{\mathrm{H}}(z_t, a_{t-1}) := e(z_t) + a_{t-1} \, \forall \, t \geq 0.$$

The endowment e is a positive constant whose level depends on the exogenous shock. It is higher in the good economic state and it is higher for agents with $z^{id} = 1$ compared to the other agent group. The model's parameter values used in this paper are set according to the calibration of a quantitative example in Krusell et al. (2011) and can be found in Online Appendix OA.1.2.

3. THE NUMERICAL METHOD AND ITS THEORETICAL UNDERPINNINGS

There are two important aspects of solving models which combine both idiosyncratic as well as aggregate risk numerically. The first aspect is the iteration procedure with which the equilibrium policy functions are computed. The difficulty here is that the equilibrium problem consists of a continuum of coupled Euler equations which cannot be decoupled due to the aggregate risk. The coupling is due to market clearing. The fact that the cross-sectional distribution and, thus, the aggregate variables move over time in accordance with the aggregate risk makes the decoupling of these Euler equations impossible such that they have to be solved jointly. The problem is that standard value-function iteration cannot be used to solve an infinite-dimensional set of coupled equations. An alternative iteration procedure which does converge to the equilibrium solution has been derived in Pröhl (2018). However, the convergence only holds for the theoretical recursive policies. In this work, I extend the convergence result to discretized policies. The second important aspect of solving these models has a similar nature as we have to keep track of the cross-sectional distribution. The distribution space is infinite-dimensional which yields the question of how to discretize this space and still ensure convergence of the algorithm. Subsequently, I introduce a projection technique for distributions for which I derive approximation error bounds ensuring convergence.

3.1. *The Iterative Procedure*

The alternative iterative method in Pröhl (2018) to handle the continuum of coupled Euler equation does, in fact, turn out to be similar to the value-function iteration of a specific social planner. The reason is that the continuum of coupled Euler equations can be reduced to one generalized Euler equation which maps random variables instead of real numbers. The value function of the social planner, or in other words her Lagrangian, is then constructed in a way such that the social planner's first-order condition coincides with this generalized Euler equation. The key to this approach is to rewrite the recursive equilibrium in terms of square-integrable random variables, i.e., $g(z', x, \mu') = h(z', x, \chi)$ where $\chi \sim \mu$.[4] As the security holdings of a single agent x are one realization of the random variable χ, the continuum of Euler equations can be reduced by replacing x with χ in the Euler equation of one agent. This defines the generalized Euler equation of a social planner which I denote here by $\mathbf{T}(h)$.[5]

As in Pröhl (2018), the planner's objective function is defined by $L^A : L^2(\mathcal{Z}^{id} \times \mathbb{R}, \mathcal{B}(\mathcal{Z}^{id} \times \mathbb{R}), \mu) \times L^2(\mathcal{Z}^{id} \times \mathbb{R}, \mathcal{B}(\mathcal{Z}^{id} \times \mathbb{R}), \mu) \to [-\infty, \infty]$ where

$$
\begin{aligned}
(12) \qquad L^A(h, y; h^n) = {} & \langle \mathbf{T}(h), h\rangle + \frac{1}{2\lambda}\|h - h^n\|^2 \\
& + \left\langle y\mathbb{1}_{\{\lambda(\bar{x}-h)\geq y\}}, \bar{x} - h\right\rangle + \frac{\lambda}{2}\left\|(\bar{x} - h)\mathbb{1}_{\{\lambda(\bar{x}-h)\geq y\}}\right\|^2 \\
& - \frac{1}{2\lambda}\left\|y\mathbb{1}_{\{\lambda(\bar{x}-h)<y\}}\right\|^2
\end{aligned}
$$

with $\langle ., .\rangle$ denoting the inner product and $\|.\|$ denoting the norm within the square-integrable random variables. The argument y is the Lagrange multiplier to enforce the borrowing constraint. The parameter $\lambda > 0$ determines the step size of the iteration. Note that the objective function takes the previous guess of the optimal policy function h^n into account. With the Lagrangian as above, I now state the algorithm to approximate the equilibrium policy function in Algorithm 1. Note that Rockafellar (1976a) shows that the proximal point algorithm converges to an optimal policy even if the update of the optimal consumption and individual capital in line 5 is only approximate. This is important as the minimizer of the Lagrangian is often not known in closed form, but can be approximated with standard nonlinear solvers.

[4] Note that $\chi = \chi(z^{id'})$ is in fact a collection of conditional random variables depending on the idiosyncratic shock. This conditional random variable can be interpreted as the collection of individual security holdings of the set of agents facing the same idiosyncratic shock realization. Therefore, to be more precise one can write $\chi(z^{id'}) \sim \mu/\mathbb{P}(z^{id'}|z^{ag'})$.

[5] The generalized Euler equation $\mathbf{T}(h)$ for the two example models is defined in Online Appendix OA.2.

Algorithm 1 Proximal point algorithm

▷ A INITIALIZATION

1: Set $n = 0$. Initialize the agents' policy function and the Lagrange multiplier $H^n = (h^n, y^n)$.

2: Set the parameter $\lambda > 0$.

3: Set the termination criterion small $\tau > 0$ and the initial distance larger $d > \tau$.

▷ B ITERATIVE PROCEDURE

4: **while** $d > \tau$ **do**

5: Update H^{n+1} by

$$\begin{aligned} h^{n+1} &\approx \arg\min_h L^A\left(h, y^n; h^n\right) \\ y^{n+1} &= \max\left\{0, y^n + \lambda\left(\bar{x} - h^{n+1}\right)\right\} \end{aligned}$$

where L is defined as in (12).

6: Compute the distance $d = \|H^{n+1} - H^n\|_{L^2_{\mathbf{P}}}$.

7: Set $n = n + 1$.

8: **end while**

I make the form of approximation of the policy update precise to show convergence in the following theorem.

LEMMA 3 (Convergence under approximate minimization[6]) *Compute the iterate h^{n+1} in Algorithm 1, line 5, as an approximate solution satisfying*

$$\begin{aligned} \frac{\epsilon^2}{2\lambda} \geq &\left(B\left(z', x, X'\right) - \bar{x}\right) \\ &\cdot \left\| \mathbf{T}\left(h^{n+1}\right) + \frac{1}{\lambda}\left(h^{n+1} - h^n\right) - \left(y^n + \lambda\left[\bar{x} - h^{n+1}\right]\right) \mathbb{1}_{\{\lambda(\bar{x} - h^{n+1}) \geq y^n\}} \right\|_1 \end{aligned} \tag{13}$$

for any $(z', x, X') \in \mathcal{Z} \times \mathbb{R}^2_{\geq \bar{x}}$, where B as in (4) and $\mathbf{T}$ denotes the generalized Euler equation. Then Algorithm 1 converges to the optimal policy.

REMARK (i) The term inside the norm is the first-order condition corresponding to the social planner's objective L^A in (12).

(ii) Equation (13) is easily implemented by any root solver using a tolerance level of $\epsilon^2/(2\lambda)$.

(iii) The previous theorem is stated w.r.t. the theoretical policy. However, to implement the algorithm, we need to discretize the policy function. Therefore, to obtain convergence for the discretized policies, sufficiently small discretization error bounds are necessary. This issue is addressed in Section 3.2.

The convergence rate of Algorithm 1 is $O(n^{-1})$ as is shown by Güler (1991). The

[6]The mathematical details are explained in Appendix A.1, which also contains the proof.

proximal point algorithm can, however, be accelerated, which goes back to Güler (1992). The convergence rate of the accelerated algorithm is $O(n^{-2})$, which was proven in Salzo and Villa (2012). I explain the acceleration in Online Appendix OA.3.1.

Note that the proximal point algorithm corresponds to what is typically referred to as an inner loop. For instance, in the Huggett economy the outer and inner loop construction of the iteration algorithm works as follows. In the outer loop, we search for the equilibrium bond prices, whereas, in the inner loop, we search for the optimal policy corresponding to the current guess of the bond price. This inner loop will be implemented as in Algorithm 1. The outer loop will be implemented using a standard nonlinear solver. The pseudo-algorithm of the outer-inner loop construction can be found in Online Appendix OA.3.2. In contrast, the Aiyagari-Bewley economy does not require an outer loop. It is shown in Pröhl (2018) that the inner loop already solves for the aggregate variables in equilibrium. The reason is that the interest rate is defined explicitly in terms of aggregate capital due to the optimizing firm. This simplification does not work in the Huggett economy because the bond price is implicitly defined by the zero-net supply condition.

3.2. *Discretizing the Space of Distributions*

The optimal policy h depends on the conditional cross-sectional random variable χ. Hence, we need to discretize the space of distributions to compute the optimal policy for any possible current-period cross-sectional distribution of $\chi \sim \mu'$. The existing literature often resorts to using a finite number of moments to characterize the cross-sectional distribution. However, even though there is a one-to-one correspondence between a distribution and its moment-generating function, this function does not exist for all distributions. Hence, any moment-based method cannot span the full space of square-integrable distributions and for some models, especially the ones producing fat-tailed cross-sectional distributions, such an approximation method is bound to fail. Another option would be a histogram representation or a spline interpolation of the distribution, but the discretized state space becomes very large very quickly in this case. Unfortunately, we have to rule out projection on orthogonal polynomials, which is widely used in computational economics as well, because a prerequisite is a sufficiently smooth distribution. Due to the borrowing constraint, however, the cross-sectional distribution exhibits mass points at the constraint and elsewhere. This fact is illustrated in Section OA.4 of the Online Appendix.

There is an efficient way of approximating distributions circumventing the aforementioned problems. To do so, it is important to view the cross-sectional distribution as a random variable. Instead of the polynomial projection of a c.d.f. on the real line, I use polynomial projection in the space of square-integrable random variables. One can interpret this approach as a probabilistic rather than a deterministic polynomial projection. This technique is called polynomial chaos and is well known in the physics and engineering literature. It is a method, which projects a square-integrable distribution onto orthogonal polynomials which have random variables as arguments instead of the real line. The advantage of this approach is that it spans the whole space of square-integrable random variables and hence, one can be sure to approximate any cross-sectional distribution sufficiently well. This includes discrete distributions and mixtures of discrete and continuous distributions. Furthermore, when the basic random variables and their corresponding family of polynomials are chosen carefully, the speed of convergence easily outperforms standard polynomial projection. Hence, a low order of polynomials is enough to obtain a good approximation of the cross-sectional distribution. In the following, I will summarize the method of polynomial chaos in general. Subsequently, I derive approximation error bounds which yield the convergence of this computational approach in this paper.

3.2.1. *Generalized Polynomial Chaos*

The standard polynomial chaos expansion is an approach to represent random variables by a series of polynomials mapping basic random variables into the space of square-integrable random variables L^2. Originally, this approach yields the so-called Wiener-Hermite expansion, i.e., a projection onto Hermite polynomials, which take Gaussians as basic random variables. The well known Cameron-Martin theorem (see e.g., Ernst et al., 2012, Theorem 2.1) shows that this construction spans all square-integrable random variables, which are measurable w.r.t. the basic random variables. Xiu and Karniadakis (2002) extend this concept to sets of orthogonal polynomials mapping more general basic random variables, e.g., uniform, gamma or binomial variables, into L^2. The L^2-convergence result for these generalized polynomial chaos expansions is proven in Ernst et al. (2012). The main purpose of this generalization is the gain in convergence speed when the basic random variables are chosen such that they are similar to the approximated random variable. To summarize, given a basic random variable $\xi \in L^2$ with distribution $\xi \sim F$, which has finite moments of all orders, and a set of orthogonal polynomials $\{\Phi_i\}_{i=0}^{\infty}$, where i denotes the order of each polynomial, we can represent any random variable $\chi \in L^2$ with distribution $\chi \sim \mu$ by

$$\chi = \sum_{i=0}^{\infty} \varphi_i \Phi_i\left(\xi\right), \tag{14}$$

where φ_i are constant projection coefficients.

It is important to note that there is an explicit connection between the basic random variable and the set of orthogonal polynomial used. The orthogonality condition of the polynomials reveals this relation. For polynomials of order $i, j \in \{0, 1, \ldots\}$, it reads

$$\langle \Phi_i, \Phi_j \rangle = \int_{-\infty}^{\infty} \Phi_i\left(\xi\right) \Phi_j\left(\xi\right) dF\left(\xi\right) = \frac{\delta_{ij}}{a_i^2}, \tag{15}$$

where δ_{ij} denotes the Kronecker symbol and $a_i \neq 0$ are constants. One can see that the weighting function, which defines the orthogonal polynomials, has to equal the distribution of the basic random variable.

Once a basic random variable is fixed, we can generate the corresponding orthogonal polynomials by the three-term recurrence relation (see e.g., Gautschi, 1982; Zheng et al., 2015)

$$\Phi_{i+1}\left(\xi\right) = \left(\xi - \theta_i\right) \Phi_i\left(\xi\right) - \omega_i \Phi_{i-1}\left(\xi\right), \, i \in \{0, 1, \ldots\}, \tag{16}$$

where the starting polynomials are defined by $\Phi_{-1}(\xi) = 0$ and $\Phi_0(\xi) = 1$ and $\theta_i, \omega_i \in \mathbb{R}$ are constant parameters with $\omega_i > 0$. In Zheng et al. (2015), different methods for generating polynomials are compared. Of their suggested methods, I use the Stieltjes method, which performs well in terms of precision. It directly computes the parameters θ_i and ω_i in (16) using the standard inner product of L^2 and is explained in detail in Gautschi (1982). The constant parameters are given by

$$\begin{aligned}
\theta_i &= \frac{\langle \Phi_i, \xi \Phi_i \rangle}{\langle \Phi_i, \Phi_i \rangle}, \, i \in \{0, 1, \ldots\} \\
\omega_i &= \frac{\langle \Phi_i, \Phi_i \rangle}{\langle \Phi_{i-1}, \Phi_{i-1} \rangle}, \, i \in \{1, 2, \ldots\}
\end{aligned}$$

with $\langle ., . \rangle$ denoting the standard inner product of L^2 w.r.t. the corresponding basic distribution. The definitions of these parameters follow from inserting the three-term recurrence relation (16) into the orthogonality condition (15).

The projection coefficients in the polynomial chaos expansion (14) are defined as usual by $\varphi_i = \langle \chi, \Phi_i \rangle / \langle \Phi_i, \Phi_i \rangle$ for all $i \in \{0, 1, \ldots\}$. If χ is not a direct function of the basic random variable ξ, one uses the fact that both c.d.f.s $\mu, F \sim \mathcal{U}[0,1]$ are uniform to compute the coefficients

$$\varphi_i = \frac{1}{\langle \Phi_i, \Phi_i \rangle} \int \mu^{-1} \circ F(\xi)\, \Phi_i(\xi)\, dF(\xi)\ \forall\, i \in \{0,1,\ldots\}, \tag{17}$$

where μ^{-1} is the generalized inverse distribution function of χ. Hence, with the polynomial chaos expansion, we can translate any square-integrable random variable $\chi \sim \mu$ into a countable series of constant projection coefficients $\{\varphi_i\}_{i=0}^{\infty}$. For computational reasons, I truncate the series of projection coefficients later on.

For practical reasons, it is important to note that polynomial chaos can be extended to multivariate distributions. First, note that a univariate random variable $\chi \sim \mu$ can be approximated with multiple basic random variables. To do so, we simply need to fix n independent basic random variables $\xi^1, \ldots, \xi^J$ and determine their corresponding univariate orthogonal polynomials $\Phi^{\xi^1}, \ldots, \Phi^{\xi^J}$ separately. Then, the polynomial chaos expansion equals

$$\chi = \sum_{i=0}^{\infty} \varphi_i \Phi_i\left(\xi^1, \ldots, \xi^J\right) = \sum_{i=0}^{\infty} \varphi_i \sum_{\substack{0 \le i_1, \ldots, i_J \le i, \\ i_1 + \ldots + i_J = i}} \Phi_{i_1}^{\xi^1}\left(\xi^1\right) \cdot \ldots \cdot \Phi_{i_J}^{\xi^J}\left(\xi^J\right). \tag{18}$$

The projection coefficient is computed as in (17) for $\xi = (\xi^1, \ldots, \xi^J)$ which reduces to the sum of a composition of integrals due to the independence of the basic random variables.[7] To approximate a joint distribution of m random variables denoted by $(\chi_1, \ldots, \chi_m) \sim \mu$, we can work with the conditional random variables $\chi_i(\chi_1, \ldots, \chi_{i-1}, \chi_{i+1}, \ldots, \chi_m)$. Each of these conditional random variables can then be approximated by $J \geq m$ independent basic random variables. Note that one has to properly define the sigma-algebras of each χ_i. One can for instance choose $\sigma(\xi_i)$ or $\sigma(\xi_1, \ldots, \xi_i)$ for χ_i.

When considering our model setup, we can actually separate the distributions of exogenous and endogenous variables as the former are known. In fact, as the exogenous variables are given, we do not need to use an approximation via polynomial chaos. Instead, we can describe the idiosyncratic shock and it's dependence on the aggregate shock exactly by one basic random variable ξ^1 and a function $z^{id'} = q(z^{ag'}, \xi^1)$. One possibility of defining such a function q results from a dis-

[7]Note that χ has to be measurable w.r.t. $\sigma(\xi^1, \ldots, \xi^J)$.

crete distribution with three states and c.d.f.

$$F\left(\xi^1\right) = \begin{cases} p^{z^{id'}=0|z^{ag'}=1} & ,\xi^1 = 1 \\ p^{z^{id'}=0|z^{ag'}=0} & ,\xi^1 = 2 \\ 1 & ,\xi^1 = 3, \end{cases} \tag{19}$$

which yields

$$z^{id'} = q(z^{ag'}, \xi^1) = \mathbb{1}_{\{\xi^1 > 2 - z^{ag}\}}.$$

In contrast, the distribution of security holdings χ is not known a priori and has to be approximated by polynomial chaos. In fact, it is a conditional distribution depending on the idiosyncratic shock $\chi = \chi(z^{id'})$, i.e., separating the agent groups with the same idiosyncratic outcome. Thus, the polynomial chaos expansion for χ has to depend on ξ^1 and can depend on further basic random variables $\xi^2, \ldots . \xi^J$ as in (18). With this construction, the projection coefficients for χ can be specified in more detail than in (17)

$$\varphi_i = \frac{1}{\langle \Phi_i, \Phi_i \rangle} \sum_{\xi^1=1}^{3} \int \left\{ \left[\mu \left(\mathbb{1}_{\{\xi^1 > 2 - z^{ag}\}}, . \right) \right]^{-1} \circ F\left(\xi\right) \right\} \Phi_i\left(\xi\right) dF\left(\xi\right)$$

for all $i \in \{0, 1, \ldots\}$ and where $\xi = (\xi^1, \ldots . \xi^J)$. Let me now summarize what the necessary steps are to approximate the space of cross-sectional distributions χ with a polynomial chaos expansion.

Before starting the solution algorithm:

1. Determine how many basic random variables are necessary.
2. Fix the distribution of each basic random variable.[8]
3. For each basic random variable, generate its corresponding orthogonal polynomials using the orthogonality condition (15) and the three-term recurrence relation (16).
4. Compute the multivariate orthogonal polynomials by multiplying the univariate polynomials according to (18).

During the solution algorithm:

[8] Any such distribution has to possess finite moments of all orders and be determinate in the Hamburger sense (see Ernst et al., 2012). A distribution is determinate in the Hamburger sense if it uniquely solves the Hamburger moment problem or in other words if it is uniquely determined by the sequence of its moments.

5. Represent any endogenous distribution by projecting it onto the predetermined polynomial chaos expansion according to (17).

The details of implementing Steps 1-3 in the two example models are explained in Online Appendix OA.5.

With the basic random variables defined and the corresponding polynomials generated, the polynomial chaos expansion is fully determined. Any square-integrable distribution measurable w.r.t. the basic random variables can now be projected. The polynomials with different degrees have different effects in this projection as can be seen in Figure 2. In this figure, I consider a polynomial chaos expansion

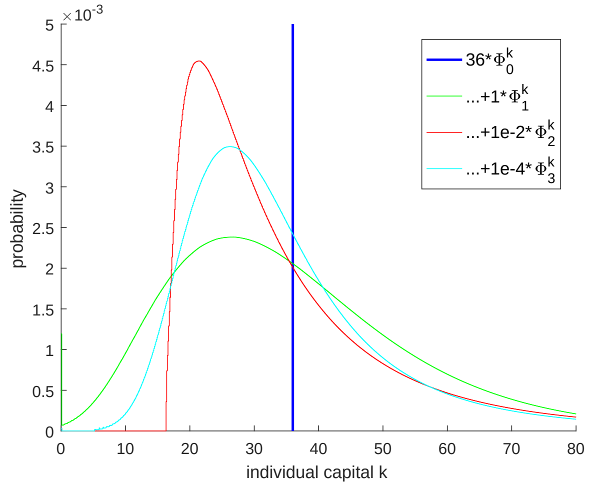

FIGURE 2. **Example distributions resulting from truncated polynomial chaos expansions.** The graph displays distributions resulting from the polynomial chaos expansion truncated at different orders ranging from order 0 to 3. The basic random variable is denoted by ξ^k. Its distribution is defined by a histogram with bin size 0.1 plotted in Figure OA-1 in the online appendix. The projection coefficients for this example are fixed at $[\varphi_0, \ldots, \varphi_3] = [36, 1, 0.01, 0.0002]$.

$\sum_{i=0}^{M} \varphi_i \Phi_i(\xi^k)$ with fixed projection coefficients $\{\varphi_i\}_{i=0}^{M}$. The expansion is truncated at increasing order M. The zeroth-order polynomial results in a mass point at φ_0, which, due to its definition, is the mean of the projected distribution. The first-order polynomial simply stretches or compresses the distribution of the basic random variable depending on its projection coefficient. Summing the zeroth- and first-order term simply centers the stretched/compressed distribution of the basic random variable around the mean of the projected distribution. Further adding the second-order polynomial modifies the skewness of the polynomial chaos expansion, whereas, the third-order polynomial adjusts the kurtosis. Higher orders further refine the tails. Hence, the polynomial chaos expansion gets closer to the projected distribution the higher the order of truncation.

3.2.2. *Convergence of the Discretized Policy*

When using polynomial chaos expansions for the cross-sectional distribution, we can approximate the policy $h(z', x, \chi)$ quite naturally because $\chi \sim \mu'$ is fully defined by the projection coefficients $\{\varphi_i\}_{i=0}^{\infty}$ due to 18. To describe the optimal policy at each realization of the random variable, we can, hence, write down the optimal policy as $h(z', x, \{\varphi_i\}_{i=0}^{\infty})$ where $\xi^1, \ldots, \xi^J$ are fixed. The approximation of this policy occurs in two steps. First, we truncate the polynomial chaos expansion, and second, we discretize all dimensions and apply the finite element method with first-order Lagrange elements, which amounts to linear interpolation. I denote the truncated policy by $h^M = h(z', x, \{\varphi_i\}_{i=0}^{M})$. Its interpolant, denoted by $h^{M,D}$, is defined on a tensor product of finite grids of the state space elements

$$D = \left\{ \left(k_{i_0}, \varphi_{0,i_1}, \ldots, \varphi_{M,i_{M+1}}\right) \middle| \, i_m = 1, \ldots, I_m < \infty \, \forall \, m = 0, \ldots, M+1 \right\}.$$

The question is whether Algorithm 1 converges for such a discretized policy. This is shown in the following.

Convergence follows from a vanishing approximation error. The total policy function approximation error is composed of two parts corresponding to the truncation and interpolation error

$$\|h - h^{M,D}\|_{L^2} \leq \|h - h^M\|_{L^2} + \|h^M - h^{M,D}\|_{L^2}.$$

The following theorem derives bounds on these two parts of the error. The bound on the second part is a well established result from the theory on finite elements (see e.g., Brenner and Scott, 2007), whereas, the bound on the first part is more involved. To derive it, I follow the methodology of the error analysis in Babuška et al. (2007).

Theorem 4 (Error bounds of the approximation) *Consider the generic model from Section 2.1. Assume that the income function B in the budget constraint (3) is real analytic in the individual and aggregate security holdings and hence, satisfies*

$$\left\| \frac{\partial^p}{[\partial a]^p} B \right\| \leq \mathbf{c}_B^p p!, \; p \in \{1, 2, \ldots\}, \tag{20}$$

for some constants $\mathbf{c}_B$ where a is a handle for x and X.[9] *Consider Algorithm 1 with*

[9] It is easy to check that the budget constraints of the two example models satisfy this assumption.

polynomial chaos expansion as in Section 3.2, i.e., using the basic random variables $\xi^1, \ldots, \xi^J$ *and the corresponding orthogonal polynomials* Φ *to project any square-integrable cross-sectional distribution* $\chi \sim \mu'$ *with* $\chi = \sum_{i=0}^{M} \varphi_i \Phi_i(\xi^1, \ldots, \xi^J)$*. Assume that, for any fixed exogenous shock, start holdings and distribution* (z', x, μ')*, the initial guess of the holdings policy* h^0 *and the Lagrange multiplier* y^0 *for the proximal point algorithm are real analytic in the basic random variables* ξ^j*,* $j \in \{1, \ldots, J\}$*. Furthermore, assume that the initial policy guess* h^0 *is real analytic in security holdings* x*. Consider the following subsets of the complex plane*

$$\Sigma\left(\tau^j_{n+1}, \Gamma^j\right) = \left\{ x \in \mathbb{C} \,\middle|\, \inf_{\xi^j \in \Gamma^j} |x - \xi^j| \leq \tau^j_{n+1} \right\},\ j \in \{1, \ldots, J\},$$

where Γ^j *is the range of* ξ^j *and* $0 < \tau^j_{n+1} < \min(1, \mathbf{L}_{n+1})/2A^1_{n+1,j} < \infty$*.* $\mathbf{L}_{n+1}$ *is the value of the second derivative of the Lagrangian* L^A *in (12) evaluated at the* $(n+1)$*-th policy iterate and* $A^1_{n+1,j}$ *is given in (28) in the proof. Then, the approximation error bound for the* $(n+1)$*-th policy iterate resulting from truncating the polynomial chaos expansion at order* M *and using linear interpolation on a rectangular tensor-product grid*

$$\begin{aligned} D = \{ &(k_{i_0}, \varphi_{0,i_1}, \ldots, \varphi_{M,i_M}) |\, k_{i_n} < k_{i_{n+1}}, \varphi_{m,i_n} < \varphi_{m,i_{n+1}} \\ &\forall\, i_n \in \{1, \ldots, d_n\},\ m \in \{1, \ldots, M\}\} \end{aligned}$$

with maximum mesh-size s *is given by ResApproxImpl*

$$(21) \qquad \begin{aligned} \|h^{n+1} - h^{M,D}\|_{L^2} \leq &\sum_{j=1}^{J} \mathbf{b}_j \frac{2}{\eta^j - 1} e^{-M \log(\eta^j)} \frac{\min(1, \mathbf{L}_{n+1})}{\min(1, \mathbf{L}_{n+1}) - 2\tau^j_{n+1} A^1_{n+1,j}} \\ &+ \mathbf{b}_d s^2 \left(\sum_{m=0}^{M+1} \left\| \frac{\partial^2 h^M}{[\partial D_m]^2} \right\|^2_{L^2} \right)^{\frac{1}{2}}, \end{aligned}$$

where $\mathbf{b}_j$*,* $j \in \{1, \ldots, J, d\}$*, are constants and*

$$\eta^j = \frac{2\tau^j_{n+1}}{|\Gamma^j|} + \sqrt{1 + \frac{4(\tau^j_{n+1})^2}{|\Gamma^j|^2}} > 1,\ j \in \{1, \ldots, J\}.$$

REMARK (i) The theorem implies that the error from the truncation of the polynomial chaos expansion decreases exponentially when increasing the order of the expansion. Furthermore, the error from the interpolation decreases proportionately with the mesh size of the discretization.

(ii) Combining Theorem 4 on the error bounds and Lemma 3 on the approximate minimization results in the convergence of Algorithm 1 with discretized policies. Assume that one implements the nonlinear solver in a way such that the approximate minimization in (13) holds strictly. Then, it is easy to see that, due to the continuity of the right-hand side of (13) in h^{n+1}, one can find a discretization of h^{n+1} which satisfies (13) as well. According to (21), one simply has to choose a fine enough mesh size s and a high enough approximation order M.

(iii) With this proof of convergence, the algorithm is a valid approximation of the rational expectations equilibrium and thus, can be seen as a benchmark for numerical solutions resulting from methods incorporating bounded rationality.

3.3. *Alternative Solution Method*

The proximal point algorithm described in Section 3.1 is similar to value function iteration as it solves a separate optimization problem in each iterative step. Hence, this procedure is expected to be demanding in terms of computational time. In practice, many researchers use policy function iteration which was first described in Coleman (1990) instead of value function iteration for this reason. Li and Stachurski (2014) show that policy function iteration of the agent's problem converges when keeping the aggregate variables fixed. Extending those results to the general equilibrium setup here lies beyond the scope of this paper. However, I do compare policy function iteration and the proximal point algorithm both with polynomial chaos numerically in the following. Note that this also serves as an illustration that the method to discretize the distribution space is versatile and can be combined with various iterative procedures.

4. NUMERICAL EVALUATION OF THE COMPUTATIONAL METHOD

In the following, I show that the theoretical results from the previous section are confirmed numerically. I solely present the results for the Aiyagari-Bewley model with aggregate risk as it is the benchmark model for computational methods for this model type and because the results for the Huggett economy with aggregate risk are qualitatively the same. I report them in Online Appendix OA.6.2.

Using Matlab R2016b, I compute the recursive equilibrium solution of the accelerated proximal point algorithm explained in Online Appendix OA.3.1 as well as the solution resulting from policy function iteration both in combination with polynomial chaos truncated at different orders.[10] For comparison, I also compute

[10] The computations were performed on the Baobab cluster at the University of Geneva. The

the solutions of existing algorithms. I choose the algorithm by Krusell and Smith (1998) as it is the most prominent existing method. I use its Matlab implementation by Maliar et al. (2010). Furthermore, there has been an effort to improve on this original algorithm in a special issue of the Journal of Economic Dynamics and Control in January 2010. From these more recent methods, I use the backward induction algorithm by Reiter (2010a) and the explicit aggregation algorithm by den Haan and Rendahl (2010), both implemented in Matlab, as they perform best in the comparison by den Haan (2010).

To ensure comparability, I run all these methods using the same grid for individual capital and the same termination criterion 5e-5. Additionally, I configure the discretizations of the cross-sectional distribution so that they are as close as possible. The Krusell-Smith and the Reiter method use total aggregate capital, whereas, the den Haan-Rendahl algorithm uses the aggregate capital of the unemployed and employed. In the proximal point algorithm as well as the policy function iteration, the total aggregate capital is equivalent to the projection coefficient ϕ_0 of the polynomial chaos expansion. I use 4 grid points for aggregate capital for the Krusell-Smith, the Reiter and the polynomial chaos based algorithms. Note that the Haan-Rendahl algorithm then has 4×4 grid points in aggregate capital because it differentiates unemployed and employed aggregate capital. Keep in mind that the proximal point algorithm has additional dimensions to discretize the cross-sectional distribution depending on the order of truncation. The different methods are summarized in Table 1. Note that the proximal point algorithm and the policy function iteration is implemented with parallelized Matlab code run on an HPC cluster, whereas, the other three algorithms are implemented with serial code run on a desktop computer. This is the reason for the differences in the number of CPUs used. Comparing the proximal point algorithm with policy function iteration, the compute time is higher for the former algorithm. This is mainly due to the fact that it solves a full optimization problem in each iteration to ensure convergence for a model with unbounded utility function. Hence, if one can ensure that policy function iteration converges, then this should be the method of choice. If convergence of a faster ad hoc method is not clear, one can use the proximal point algorithm as a benchmark.

In the following, I investigate whether the algorithms using polynomial chaos expansions really yield higher precision than the existing methods. One way of comparing these sets of numerical solutions is to analyze their Euler equation errors. There have been two different Euler equation error tests put forward in

code can be found in the supplementary material. Instructions to reproduce the results are given in Section OA.8 of the Online Appendix.

Algorithm	# Grid Points for $z' \times k \times \mu$	# CPUs	Compute Time
Krusell-Smith (K-S)	$4 \times 80 \times 04$	4	53s
Reiter (R)	$4 \times 80 \times 04$	4	2m 59s
den Haan-Rendahl (D-R)	$4 \times 80 \times 16$	4	22s
Proximal Point Algorithm M=0 (PPA0)	$4 \times 80 \times 04$	20	24s
Proximal Point Algorithm M=1 (PPA1)	$4 \times 80 \times 12$	20	31s
Proximal Point Algorithm M=2 (PPA2)	$4 \times 80 \times 24$	20	1m
Proximal Point Algorithm M=3 (PPA3)	$4 \times 80 \times 48$	20	1m 49s
Policy Function Iteration M=0 (PFI0)	$4 \times 80 \times 04$	20	6s
Policy Function Iteration M=1 (PFI1)	$4 \times 80 \times 12$	20	8s
Policy Function Iteration M=2 (PFI2)	$4 \times 80 \times 24$	20	13s
Policy Function Iteration M=3 (PFI3)	$4 \times 80 \times 48$	20	18s

TABLE 1. **Summary of the algorithms to be compared.** In the first column, M denotes the order of truncation of the polynomial chaos expansion. The abbreviation in the parenthesis is the algorithm identifier used in the comparison analysis, which follows. The second column displays the total number of grid points to discretize the policy function. Note that the methods of discretizing the distribution μ vary across algorithms. If the distribution is discretized with several parameters, the number of grid points is the product of the number of grid points in each parameter.

the literature (see e.g., den Haan, 2010), the standard Euler equation error test and the dynamic Euler equation error test. The standard Euler equation errors are calculated by comparing the numerical solution for optimal consumption c against the explicitly calculated conditional expectation in the Euler equation denoted by $\tilde{c}$. It is the absolute percentage error[11]

$$\epsilon^{SE} = \frac{|c - \tilde{c}|}{\tilde{c}}.$$

In contrast to the standard Euler equation error, the dynamic equivalent denoted by ϵ^{DE} is computed for several consecutive periods. This test is more stringent as the numerical solution and the explicit conditional expectation usually diverge with more periods. I compute the standard and the dynamic Euler equation error for a random sample of aggregate shocks over 3000 periods for the different numerical solutions from Table 1. Note that I compute the standard Euler Equation error test also over multiple periods, but it is reset every period and hence, does not accumulate. The errors' summary statistics are given in Online Appendix OA.6.1.1. They indicate that the polynomial chaos based algorithms are advantageous. This advantage can be visualized better by displaying the full error distribution in terms of boxplots in Figure 3. One can see that the existing algorithms produce much

[11]It is also sensible to look at the error without taking the absolute value as systematic biases can be identified more easily this way. This analysis can be found in Online Appendix OA.6.1.2.

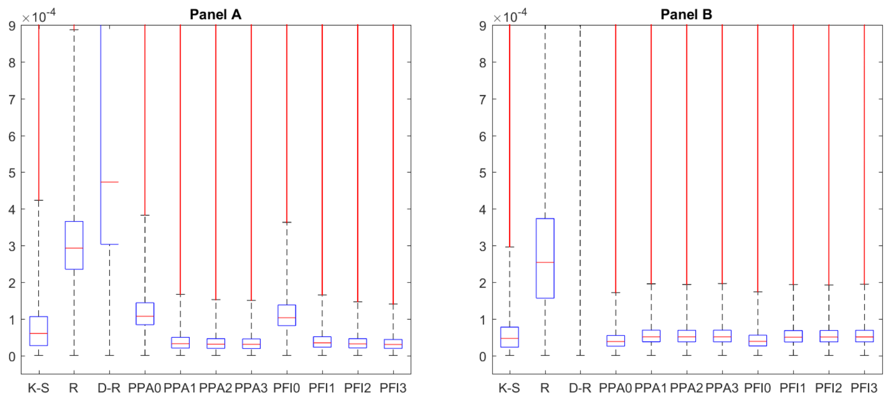

FIGURE 3. **Boxplots of the Euler equation error distributions of individual capital.** Panel A shows the standard Euler equation error and Panel B displays the dynamic Euler equation error for the numerical solutions from Table 1. The error is computed over a finer grid in the dimension of individual capital than the grid on which the solutions are computed. I compute the error for a random sample of aggregate shocks over 3000 periods. The initial cross-sectional distribution is the same for all algorithms. The red lines mark the medians, whereas, the blue boxes denote the 25^{th} to 75^{th} percentiles. The whiskers indicate the range of the distribution and the red dots outside are outliers.

wider error distributions for both the standard and the dynamic Euler equation error. It is interesting to observe that the Reiter algorithm, although improving on the extreme points of the error distribution, does not lead to any improvement w.r.t. the width of the error distribution compared to the Krusell-Smith algorithm. The same is true for the den Haan-Rendahl method, which performs much worse. It seems that the reason why they performed well in the comparison by den Haan (2010) is that they use considerably more grid points, whereas here, I deliberately run all methods on the same discretization. In comparison, all polynomial chaos based solutions produce much narrower error bands. This is mainly due to the better anticipation of the cross-sectional distribution's law of motion in the polynomial chaos algorithms.

It seems counter-intuitive that the error distributions for the algorithms based on polynomial chaos do not differ much. One would expect that they decrease with increasing order. To understand why, recall that the approximation error bound in (21) consists of two terms. A decrease in error due to the increase in the truncation order of the polynomial chaos might be offset by an increase in the interpolation error. Note that even though I use the same interpolation grid for all algorithms, the interpolation error may still differ as it depends on the curvature of the solution. The higher the curvature, the higher the interpolation error. Therefore, we need to disentangle the truncation error from the interpolation error. This is possible by comparing the prediction of the next-period aggregate capital by the algorithm K with the true next-period aggregate capital $\tilde{K}$. I display the law of motion error, computed by

$$\epsilon^{LoM} = \frac{K - \tilde{K}}{\tilde{K}},$$

in Figure 4. I exclude the den Haan-Rendahl algorithm because it already performs

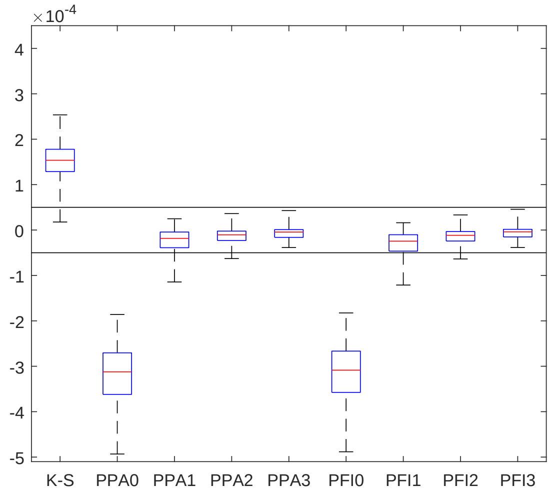

FIGURE 4. **Boxplots of the law of motion error distributions of aggregate capital.** This figure shows the low of motion error ϵ^{LoM} produced by the Krusell-Smith algorithm and the polynomial chaos algorithms from Table 1. The red lines mark the medians, whereas, the blue boxes denote the 25^{th} to 75^{th} percentiles. The whiskers indicate the range of the distribution and the red dots outside are outliers.

worse than the others in terms of the Euler equation errors. I also exclude the Reiter algorithm because it was not possible to extract the prediction of next-period aggregate capital from Reiter's Matlab implementation. The figure shows that, indeed, the law of motion error becomes smaller when the truncation order is increased. In particular, we observe an exponential decrease as predicted by the error bound. Furthermore, the law of motion error is within the region of the truncation criterion when truncating at second or higher order. This indicates that a polynomial chaos expansion up to second order yields sufficient precision for the standard calibration of the Aiyagari-Bewley economy described in Online Appendix OA.1.1.

Overall, the error analysis shows that the polynomial chaos based algorithms outperform the existing algorithms and that the polynomial chaos expansion up to second order suffices to approximate the growth model. Recall that order zero implies that the optimal policies depend solely on aggregate capital. Order one and higher, however, imply a dependance on the full approximated distribution. Hence, to approximate the rational expectations equilibrium of the growth model sufficiently, the agents need to know more than the aggregate capital. However, a crude approximation of the cross-sectional distribution seems to be enough.

5. APPROXIMATE AGGREGATION REVISITED

As shown in the previous section, using polynomial chaos to approximate the cross-sectional distributions seems to imply that the first moment is not enough to achieve a satisfactory level of precision in the numerical solution of the Aiyagari-Bewley model. Looking at this result superficially, it seems to contrast the approximate aggregation result by Krusell and Smith (1998). However, this result is simply due to the higher sensitivity of the polynomial chaos approximation w.r.t. changes in the cross-sectional distribution. Evaluating aggregation requires a more in-depth analysis of the model's stationary cross-sectional distribution.

I cannot compute the full stationary state distribution for this model though, as it is a distribution of distributions $\mathcal{P}(z', k, \mu)$. However, I can consider the expected conditional cross-sectional distribution $\mathbb{E}^{\mu}(\mathcal{P}(z', k|\mu))$, which essentially represents the average stationary cross-sectional distribution. It is computed as a fixed point of the cross-sectional distribution's law of motion (6) and displayed in Figure 5. As the Euler equation errors for the den Haan-Rendahl algorithm are large, I will

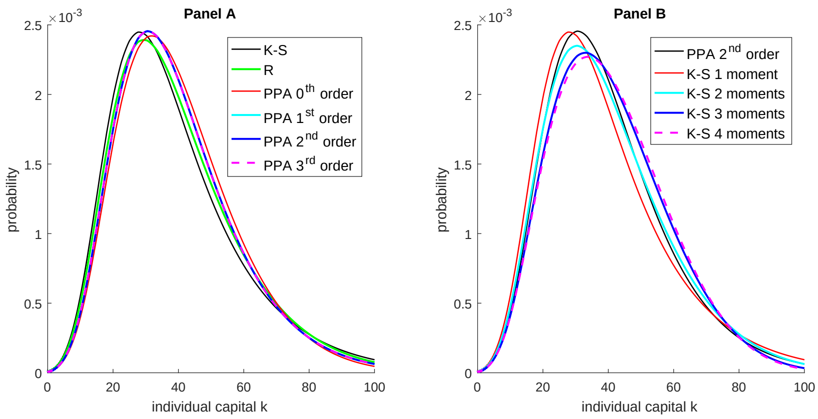

FIGURE 5. **Average cross-sectional distributions produced by the Krusell-Smith, Reiter and the polynomial chaos algorithms from Table 1.** This graph displays the average stationary cross-sectional distribution, i.e., the expectation of the stationary distribution of the state space (z', k, μ) conditional on μ. It displays histogram approximations of the distributions with bin size 0.1. Panel A compares the p.d.f.s of the Krusell-Smith algorithm using the first moment, the Reiter algorithm and the proximal point algorithm using polynomial chaos. Panel B compares the benchmark proximal point algorithm truncated at second order with the Krusell Smith algorithm using multiple moments.

compare the polynomial chaos algorithms[12] only to the Krusell-Smith and the

[12]Note that only the proximal point algorithms are considered because policy function iteration produces the same solutions.

Reiter algorithm. The largest conceptual difference between these algorithms is that the Krusell-Smith method assumes bounded rationality in terms of a rule of thumb, i.e., a parametric law of motion for the aggregate variables depending on a finite number of moments. The polynomial chaos based algorithms, however, use the nonparametric law of motion of the aggregate variables stemming from the cross-sectional distribution. The Reiter algorithm lies conceptually in between the former two. It maps the set of moments to a parameterized cross-sectional distribution rather than a rule of thumb to compute the prediction for aggregate variables. We can see that the distributions of the proximal point algorithm with different orders of truncation are almost indistinguishable. They are further to the right than the distribution of the Krusell-Smith algorithm. As expected, the distribution of the Reiter algorithm lies in between the former. Panel B serves as a robustness check. It displays the solutions of the Krusell-Smith algorithm with several moments. As the number of moments increases, the distributions also seem to converge. However, they do not converge to the distribution produced by the proximal point algorithm. The fact that the distributions of the different proximal point algorithms are so close explains the approximate aggregation result from Krusell and Smith (1998). In terms of stationary distributions, higher orders do not matter in this calibration of the growth model.[13]

An interesting question remains. When does approximate aggregation fail? To answer that question, I do not even have to consider a different model but only change one parameter. I keep the parameters of Online Appendix OA.1.1 except for the unemployment benefit rate ν which is adjusted from 15% to 65%.[14] Therefore, the higher unemployment insurance leads to more redistribution in the new calibration and hence, better risk sharing. The precision results for this calibration are qualitatively the same as for the benchmark calibration and can be found in Online Appendix OA.6.3. That approximate aggregation does not hold in this calibration becomes clear when considering the stationary distributions in Figure 6. The stationary distribution for polynomial chaos truncated at first order is not plotted as it did not converge. It shows that a rough approximation of the cross-sectional distribution does not suffice to solve the model. This anomaly is remedied by truncating at higher order where the expected stationary distribution stabilizes quickly. Given it's methodology, it is not surprising that the distribution

[13] The change in the average stationary distribution when truncating the polynomial chaos expansion at different orders can also serve as an evaluation criterion on when to stop increasing the order of the approximation. Online Appendix OA.6.1.3 describes the details.

[14] This figure does in fact correspond to the 2015 OECD median of the net replacement rate in the initial unemployment phase of an average-wage household with two children and one earner (see http://www.oecd.org/els/benefits-and-wages-statistics.htm).

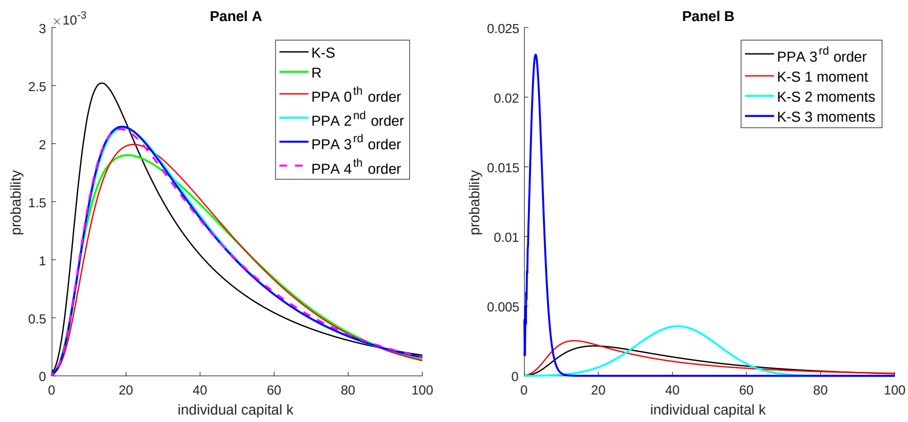

FIGURE 6. **Average cross-sectional distributions produced by the Krusell-Smith algorithm and the polynomial chaos algorithms from Table 1.** This graph displays the average stationary cross-sectional distribution, i.e., the expectation of the stationary distribution of the state space (z', k, μ) conditional on μ for 65 % unemployment benefit rate. It plots histogram approximations of the distributions with bin size 0.1. Panel A compares the p.d.f.s of the Krusell-Smith algorithm using the first moment, the Reiter algorithm and the proximal point algorithm using polynomial chaos. Panel B compares the benchmark proximal point algorithm truncated at third order with the Krusell Smith algorithm using multiple moments.

from Reiter's method is closer to ours than to Krusell-Smith's distribution which exhibits much fatter tails compared to the distributions induced by the polynomial chaos based algorithms. More importantly, Panel B shows that the Krusell-Smith method does not converge for this calibration when the number of moments is increased which exemplifies that approximate aggregation fails in this calibration.

Another example where approximate aggregation does not hold is the Huggett economy with aggregate risk. The expected stationary cross-sectional distributions of asset holdings are displayed in Figure 7. The truncation of the polynomial chaos at order zero is not plotted as this polynomial always equals to zero in equilibrium and thus, does not contain any information about the cross-sectional distribution. We can see that the largest change in the average stationary distribution occurs with the first and second-order approximations, but third and fourth-order approximations are very close. This is confirmed when zooming into the left tail, i.e., the area where mass points occur in the distribution. Note that the budget constraint in this model has a much larger effect than in the Aiyagari-Bewley economy as a larger share of the population is close to hitting the constraint.

6. ECONOMIC INSIGHTS

As the polynomial chaos based algorithm's convergence is theoretically founded, its solution represents indeed an approximation of the rational expectations equilibrium. Furthermore, the convergence was illustrated numerically even in models

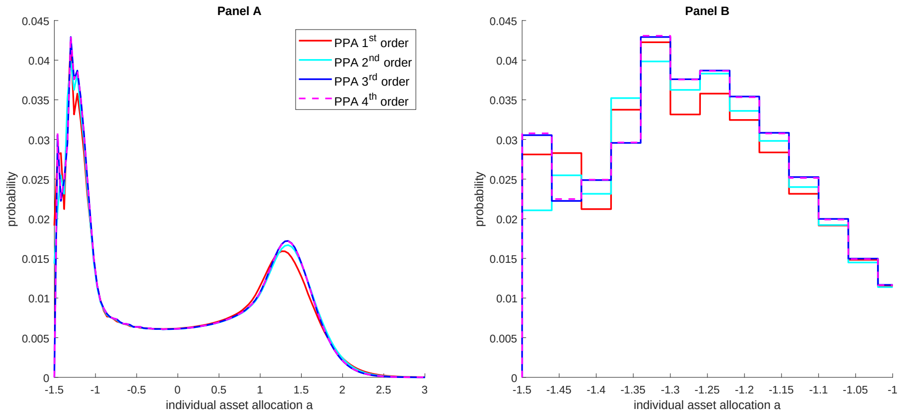

FIGURE 7. **Average cross-sectional distributions produced by the polynomial chaos algorithms from Table OA-5.** This graph displays the average stationary cross-sectional distribution, i.e., the expectation of the stationary distribution of the state space (z', a, μ) conditional on μ. It displays histogram approximations of the distributions. Panel A compares the p.d.f.s of the proximal point algorithm using polynomial chaos truncated at different orders. Panel B zooms into the left tail of the average cross-sectional distributions.

where approximate aggregation is lacking and the Krusell-Smith algorithm fails. Hence, we can be confident in deriving economic implications from these numerical solutions. Just considering the two standard models the Aiyagari-Bewley economy and the Huggett economy both with aggregate risk, two novel economic implications arise. First, the risk within the model increases when risk sharing increases and second, the amplification of the aggregate risk due to the heterogeneity is stronger when there is more risk sharing. The details of these two results are explained in the following. Both results can be found in both models but to keep the main text concise, I illustrate the first result within the Aiyagari-Bewley model and the second result within the Huggett economy. The respective missing results can be found in Online Appendix OA.7.1. Furthermore, I show that these results are robust to changes in the model parameters in Online Appendix OA.7.2.

6.1. *The Volatility Paradox*

In the Aiyagari-Bewley model, the amount of risk sharing varies with the unemployment insurance. Agent share the idiosyncratic risk more intensively if employed agents pay a higher contribution to increase the unemployment benefit rate. I compute the solution to the model for various unemployment benefit rates and report the corresponding expected stationary cross-sectional distributions in Figure 8.[15] One can see that the expected ergodic distribution has fatter tails the

[15]The supplementary material contains videos for the cases of $\nu = 15\%$ and $\nu = 65\%$ which show how the cross-sectional capital distribution evolves in a simulation over 3000 periods. It starts from the stationary distribution of the model without aggregate risk. One can observe

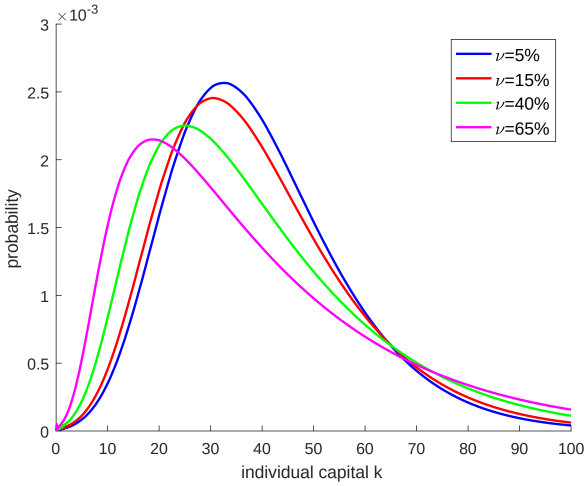

FIGURE 8. **Average cross-sectional distributions produced by the polynomial chaos algorithm for different levels of the unemployment benefit.** This graph displays the average stationary cross-sectional distribution, i.e., the expectation of the stationary distribution of the state space (z', k, μ) conditional on μ for different unemployment benefit rates ν. The proximal point algorithm for up to 15% unemployment benefit is truncated at first order, whereas, the proximal point algorithm for higher benefits is truncated at second order.

higher the unemployment benefit rate. Moreover, the mode of the distributions corresponding to higher unemployment benefit rates lie further to the right than for low benefits. Since the expected stationary cross-sectional distribution can be interpreted as the probability distribution of one agent's individual capital over the infinite time horizon, the higher volatility and fatter tails can be interpreted as higher systemic risk in the model where agents share idiosyncratic risk more intensively.

This result is puzzling at first but can be explained when considering the different reasons for holding capital. First, it can be used for precautionary savings and second, it can be used purely as an investment object. The precautionary savings motive can be satisfied by either building up one's own capital savings or contributing to the unemployment insurance which is funded by a redistributive tax on employed agents. In the model with a high unemployment insurance, the poor employed agents will satisfy the precautionary savings motive mainly through the insurance as this is a mandatory payment. Due to the higher tax, they have a lower savings rate as the same agent in the model with a low insurance level. This explains the shift to the left of the cross-sectional distribution's mode. The fat tail on the other hand can be explained by the investment motive. Recall that, due to the Cobb-Douglas production function, equilibrium returns are good when

how the distribution does converge in expectation towards the respective stationary distributions displayed in Figure 8.

aggregate capital savings are low or moderate. The lower savings rate of the poor employed agents affects the aggregate capital negatively which, in turn, affects the equilibrium rental rate on capital positively. Hence, capital is an attractive investment object for rich agents. With a decent rental rate they prefer saving capital for later to consuming today resulting in a fatter right tail. It may be tempting to conclude that the redistributive social policy only has the negative effect of increasing inequality in capital holdings. However, it does also increase consumption especially on the lower end of the distribution. [16]

This result on the shifted tails of the cross-sectional capital distribution mirrors the volatility paradox in Brunnermeier and Sannikov (2014). In their paper, they analyze a model with two types of agents who face aggregate risk and find that low-risk environments where idiosyncratic risk is better hedged are conductive of greater systemic risk. The reason that the setup herein is comparable is the following. The cross-sectional distribution in our model at any particular point in time is just a screenshot of the economic state across agents at that point. However, the expected stationary cross-sectional distribution over the infinite time horizon represents all states an agent may reach over the infinite time horizon and their respective probabilities. The volatility of this distribution can, hence, be interpreted as the volatility of the individual capital holdings of each agent over time. This interpretation does confirm the volatility paradox of Brunnermeier and Sannikov (2014) in the Aiyagari-Bewley economy with aggregate risk

6.2. *Amplification of Aggregate Risk due to Heterogeneity*

Due to the computational issues, the literature on heterogeneous-agent models has so far often limited itself to models which solely feature idiosyncratic risk. Therefore, there is scope to further investigate the interaction between idiosyncratic and aggregate risk. This section serves as an illustration that the algorithm introduced in this paper provides the possibility to do so. In the following, I show that idiosyncratic and aggregate risk do indeed interact and amplify each other in the Huggett economy with aggregate risk.[17] I compare bond prices corresponding to the expected stationary distribution across different borrowing constraints

[16] A similar observation has been made in Krueger et al. (2016), Section 6.3, who show that savings of poor agents are lower and consumption is higher when there is higher unemployment insurance. However, in their work, they only document the shifts within the left side of the distribution. They do not reach any conclusions on the right side of the distribution.

[17] Due to the computational difficulty, the question of how the interaction between idiosyncratic risk and aggregate risk affects asset prices has not been answered yet. Huggett (1993) investigates the effect of idiosyncratic risk on asset prices but does not look at aggregate risk. Krusell et al. (2011) consider aggregate risk in addition to idiosyncratic risk but limit themselves to a borrowing constraint at $\bar{a} = 0$ and, therefore, an autarkic economy without trade.

and model configurations in Table 2. First consider the Huggett economy without

Borrowing Constraint $\bar{a}$	No Aggregate Risk bad	No Aggregate Risk good	Aggregate Risk average	Aggregate Risk bad	Aggregate Risk good
-0.5	1.0440	1.0161	1.0855	0.9787	1.1922
-1.0	1.0247	1.0111	1.0782	0.9786	1.1778
-1.5	1.0149	1.0080	1.0637	0.9703	1.1570
-2.0	1.0099	1.0057	1.0505	0.9612	1.1397
-2.5	1.0063	1.0039	1.0404	0.9514	1.1294

Table 2. **Equilibrium prices corresponding to the average stationary cross-sectional distributions for various levels of the borrowing constraint.** This graph displays the prices corresponding to the average stationary cross-sectional distribution, i.e., the expectation of the stationary distribution of the state space (z', a, μ) conditional on μ. It compares the model without aggregate risk which is either fixed to the bad or good state, respectively, and the model with aggregate risk where the prices are conditional on the aggregate state.

aggregate risk. In this case, the agent's endowment is shifted up and down solely by the idiosyncratic risk but not by aggregate risk. I compute two versions of the model without aggregate risk fixing the aggregate shock at the bad and the good state, respectively. This produces two equilibria with different stationary cross-sectional distributions and corresponding prices. One can see that the equilibrium prices are higher in the bad state. The reason for this is that the average level of the income process is lower in the bad state and thus, the average marginal propensity to consume is higher. This increases the precautionary savings motive for the agents and results in higher demand. Furthermore, tightening the borrowing constraint from $\bar{a} = -2.5$ to $\bar{a} = -0.5$ increases the price level in both cases which is in line with the existing literature.[18] With these two different steady states, one could then try to approximate aggregate shocks by looking at an MIT shock and computing the transition between the two steady states. As it arises deterministically, this transition is typically rather smooth.

The question which remains is how these different steady states compare to the model with aggregate risk. To investigate this question, I compute the stationary state distribution of the model with aggregate risk. As before, this dis-

[18]A further interesting observation is that the prices are above one for all model configurations without aggregate risk. This implies that agents with high income are willing to pay a fee to store wealth for future consumption due to their precautionary savings motive. On the opposite side, agents with low income are willing to sell the bond to increase their current consumption. However, they ask for a premium to pay for the risk of hitting the borrowing constraint in the next period which would limit their future consumption. This result is not entirely surprising as it somewhat mirrors the willingness of investors to accept negative interest rates on safe government bonds like Germany or Switzerland in recent years. The price level is heavily influenced by the discount factor in the model. A lower discount factor leads to lower price levels as impatient agents have a lower precautionary savings motive and thus, demand for the bond is reduced.

tribution is a distribution of distributions. Hence, I compute the expected conditional state distribution given that the economy is in a particular aggregate state $\mathbb{E}^{\mu}(\mathcal{P}(z', a|\mu, z^{ag'}))$. I also compute the equilibrium price corresponding to this conditional distribution. In the bad aggregate state, the equilibrium price is lower than the one of both model configurations without aggregate risk. In the good state, on the contrary, it is higher. Hence, the price levels corresponding to a particular aggregate outcome in the model with aggregate risk are reversed compared to the model without aggregate risk. Furthermore, the average price level over both states is higher than in the model without aggregate risk. The increased average price level can be explained by considering the distribution of the income process of the model with aggregate risk. Even though the income level in the model with aggregate risk is the average of the income levels of the model without aggregate risk, the variance shifts nonlinearly. It is higher than the average of the variances of the models without aggregate risk. This leads to an on average stronger precautionary savings motive in the model with aggregate risk. There is more demand for the bond on average which results in the higher equilibrium price level. However, this does not explain the different price levels when conditioning on the aggregate shock outcome. This is due to the borrowing constraint amplifying the impact of aggregate risk. Given that the economy is in a bad state, the impact of the borrowing constraint on agents with low endowment is less severe because their ability to pay back debt in the next period is increased due to the higher expected next-period income compared to the model without aggregate risk fixed at the bad state. Therefore, these agents have a higher ability to borrow which increases the supply of the bond and lowers the price compared to the model without aggregate risk. This effect is so strong that the price is even lower than in the model without aggregate risk fixed at the good economic state. The mechanism is mirrored for the equilibrium price conditional on the economy being in the good state. In that case, the borrowing constraint has a much stronger grip on the agents with low income as their ability to pay back debt is decreased due to the lower expected next-period income. This leads to a decrease in supply and an increase in the price. The effect is so strong that the agents who want to buy the bond are willing to pay a significant fee for storing their wealth. The disparity between the prices in the two aggregate states illustrates the forces at play. In the good state, the demand for storing wealth from agents with high income is much stronger than for consuming future income when agents have low endowment. Thus, the bond buyers pay a fee to invest to give an incentive to other agents to sell. On the contrary, the supply in the bad state is much higher than the demand which results in a price lower than one. The agents with low endowment who want to consume future wealth today are willing to pay an interest rate to incentivize other agents to buy the bond. Overall, aggregate risk is clearly amplified through the combination of ex-post heterogeneity and borrowing constraints. The amplification effect is stronger in the Huggett economy than in the Aiyagari-Bewley economy as the borrowing constraint affects a larger share of the population.[19] Accordingly, amplification is observed across all computed levels of the borrowing constraint in the Huggett economy, whereas, it is observed only for sufficiently high levels of risk sharing in the Aiyagari-Bewley economy.

7. CONCLUSIONS

The contribution of this paper is a theoretically founded numerical method to globally solve general equilibrium models with both idiosyncratic as well as aggregate risk. Solving this type of models is challenging because the cross-sectional distribution of agent-specific variables becomes a time-varying element of the state space. What sets my algorithm apart from existing methods is that, rather than approximating the law of motion of aggregate variables with a more or less parametric formula, it approximates the cross-sectional distribution of individual variables. I use a projection technique which extends orthogonal polynomial projection from spaces of smooth functions to the space of square-integrable random variables. This technique is known as generalized polynomial chaos and can be interpreted as a probabilistic polynomial projection method. Most importantly, I show that the algorithm using this discretization converges to the theoretical rational expectations equilibrium.

From a practical perspective, the algorithm developed in this paper increases the precision of the solution by a significant amount when compared to existing approaches, especially in terms of the law of motion of aggregate variables. Furthermore, I illustrate that it can be readily applied to models where approximate aggregation does not hold. The first example is rather surprising. I show that approximate aggregation and thus, the widely used Krusell-Smith algorithm fails when high levels of risk sharing are introduced in the Aiyagari-Bewley model with aggregate risk. The second example model I choose for illustrative purposes is the Huggett economy with aggregate risk. Even in these two cases where approximate aggregation does not hold, my algorithm performs well according to the theoretical

[19] The supplementary material contains videos which show how the cross-sectional bond holdings distribution evolves in a simulation over 3000 periods for the cases $\bar{a} = -1.5$ and $\bar{a} = -0.5$. It starts from the stationary distribution of the model without aggregate risk. One can see that the bond holding distribution contracts in good periods and expands in bad periods. In bad periods, there is a sizable fraction of agents for whom the constraint binds.

convergence result. Furthermore, the more precise numerical solutions of these two standard models yield two novel economic insights. First, systemic risk in equilibrium increases with higher levels of idiosyncratic risk sharing. This result emerges in the Aiyagari-Bewley economy through fatter tails in the expected stationary capital distribution and, in the Huggett economy, through higher volatility in the stationary distribution of equilibrium bond prices. Second, sufficiently high levels of idiosyncratic risk sharing among agents facing borrowing constraints amplifies aggregate risk. This result follows from the large discrepancy between the global solution of the model with idiosyncratic and aggregate risk and the solution of the model without aggregate risk at various steady states.

APPENDIX A: PROOFS

A.1. *Proof of Lemma 3*

It is shown in Pröhl (2018) that Algorithm 1 converges to the optimal policy when the policy update in line 5 is exact. However,Rockafellar (1976a) shows that the proximal point algorithm converges to an optimum of the Lagrangian even if the update of the optimal policy is only approximate. Salzo and Villa (2012) extend this result to different concepts of approximation. Before I define which kind of approximation applies in this work, let me recall some important terminology first. Starting with the generalized Euler equation $\mathbf{T}(h)$, it is shown in Pröhl (2018) that iterating on the Euler equation's resolvent $(\mathbf{T}+\mathbf{Id})^{-1}(h)$, where $\mathbf{Id}$ denotes the identity operator, converges to the optimal policy. To construct the resolvent, we need the Langrangian associated with the Euler equation meaning that this Lagrangian's first-order condition coincides with the Euler equation $\mathbf{T}$. This Lagrangian turns out to be

$$L\left(h, y\right) = \left\langle \mathbf{T}(h), h \right\rangle + \left\langle y, \bar{x} - h \right\rangle , \tag{22}$$

where y is the Lagrange multiplier enforcing the borrowing constraint. According to Rockafellar (1976b) and Pröhl (2018), the resolvent of the Euler equation is equal to minimizing the planner's objective L^A in (12).

Definition 5 (Resolvent approximation[20]) *Let $\mathcal{C}$ be a Hilbert space over $\mathbb{R}$. Consider the resolvent $(\mathbf{Id} + \lambda\mathbf{T}_L)^{-1}(c)$ of an operator $\lambda\mathbf{T}_L$ associated with a*

[20] This definition corresponds to the type 2 approximation with ϵ-precision in Salzo and Villa (2012). Note that the approximation operator is an approximate subdifferential operator. This is the case because I minimize the controls for fixed Lagrange multipliers rather than computing a minimax problem immediately in Algorithm 1.

saddle function L at $c \in \mathcal{C}$ with $\lambda > 0$. The approximation with ϵ-precision of this resolvent at $c \in \mathcal{C}$ is defined as $\tilde{c} \in \left(\mathbf{Id} + \lambda \mathbf{T}_L^{\epsilon^2/(2\lambda)}\right)^{-1}(c)$ where

$$\mathbf{T}_L^{\epsilon^2/(2\lambda)}(c) = \left\{ v \,\middle|\, L(c) - L(\tilde{c}) + \langle \tilde{c} - c, v \rangle \leq \frac{\epsilon^2}{2\lambda} \, \forall \tilde{c} \in \mathcal{C} \right\}.$$

It is denoted by $\tilde{c} \approx (\mathbf{Id} + \lambda \mathbf{T}_L)^{-1}(c)$.

Proof of Lemma 3: Generally, convergence of the proximal point algorithm is well established as explained above. It remains to show that the approximate policy update in Algorithm 1 for the models in this paper satisfies Definition 5. As $\max_{h \in \mathcal{H}}(h^{n+1} - h) = h^{n+1} - \bar{x} \leq B(z', x, X') - \bar{x}$, equation (13) implies that

$$\langle h^{n+1} - h, \nabla \tilde{L}\left(h^{n+1}, y^n; z', k, h^n\right) - v \rangle \leq \frac{\epsilon^2}{2\lambda}, \; \forall h \in \mathcal{H},$$

where

$$\begin{aligned} \tilde{L}\left(h^{n+1}, y^n; h^n\right) =& L^A\left(h^{n+1}, y^n; h^n\right) - \frac{1}{2\lambda}\|h - h^n\|^2 \\ v =& \frac{1}{\lambda}\left(h^n - h^{n+1}\right) \end{aligned}$$

with L^A as in (12) and ∇ denoting the gradient w.r.t. h. Adding a zero and applying the definition of the gradient (or more generally, an element of the sub-differential) then implies

$$\begin{aligned} &\left[\tilde{L}\left(h^{n+1}, y^n; h^n\right) - \tilde{L}\left(h, y^n; h^n\right)\right] - \left[\tilde{L}\left(h^{n+1}, y^n; h^n\right) - \tilde{L}\left(h, y^n; h^n\right)\right] \\ &+ \langle h - h^{n+1}, v - \nabla \tilde{L}\left(h^{n+1}, y^n; h^n\right) \rangle \leq \frac{\epsilon^2}{2\lambda} \end{aligned}$$

$$\Rightarrow \begin{cases} \tilde{L}\left(h^{n+1}, y^n; h^n\right) - \tilde{L}\left(h, y^n; h^n\right) + \langle h - h^{n+1}, v \rangle \leq \frac{\epsilon^2}{2\lambda} \\ \tilde{L}\left(h^{n+1}, y^n; h^n\right) - \tilde{L}\left(h, y^n; h^n\right) + \langle h - h^{n+1}, \nabla \tilde{L}\left(h^{n+1}, y^n; h^n\right) \rangle \leq 0 \end{cases} \tag{23}$$

for all $h \in \mathcal{H}$. Using the update of y^{n+1} in Algorithm 1, it follows that $\tilde{L}(h, y^n; h^n) = L(h, y^{n+1}) - \lambda/2\|(\bar{x} - h)\mathbb{1}_{\{\lambda(\bar{x}-h) \geq y^n\}}\|^2$ and $\nabla \tilde{L}(., y^n; h^n) = \nabla L(., y^{n+1})$ with L as in (22). Inserting this into (23) and applying the definition of the gradient of L leads to

$$L\left(h^{n+1}, y^{n+1}\right) - L\left(h, y^{n+1}\right) + \langle h - h^{n+1}, v \rangle \leq \frac{\epsilon^2}{2\lambda}. \tag{24}$$

Hence, we have that $v \in \mathbf{T}_L^{\epsilon^2/(2\lambda)}(h^{n+1})$, which concludes the proof as $h^{n+1} \in$

$\left(\mathbf{Id} + \lambda \mathbf{T}_L^{\epsilon^2/(2\lambda)}\right)^{-1}(h^n)$. *Q.E.D.*

A.2. *Proof of Theorem 4*

In order to prove Theorem 4, I first need to establish that any iterate of the optimal policy h^{n+1}, $n \geq 0$, as computed in the proximal point algorithm is analytic in the basic random variables $\xi^1, \dots, \xi^J$, i.e. there exist constants $\mathbf{c}_{h^{n+1},j}$ such that

$$\left\| D_j^p h^{n+1} \right\| \leq \mathbf{c}_{h^{n+1},j}^p p!,\ p \in \{1, 2, \dots\},\ j \in \{1, \dots, J\},$$

where the p-th derivative is denoted by $D_j^p = \partial^p / \partial^p \xi^j$. Before I show analyticity, let me specify how the policy depends on the basic random variables. In the discretization of the optimal security holdings policy, I impose a grid on the exogenous shocks, the initial individual holdings and the projection coefficients. Hence, at each grid point of $h(z', x, \{\varphi_i\}_{i=0}^M)$, the states z', the individual holdings and the projection coefficients are fixed. Due to the definition of the projection coefficient φ_0, this means that the current aggregate security holdings $X' = \varphi_0$ is also fixed. However, the policy implicitly depends on the basic random variables through the Euler equation because it contains the next-period aggregate security holdings

$$X'' = \sum_{\xi^1=1}^{3} \int_{-\infty}^{\infty} h\left(z^{ag'}, \mathbb{1}_{\{\xi^1 > 2 - z^{ag'}\}}, \chi, \{\varphi_i\}_{i=0}^M\right) dF\left(\xi^1, \dots, \xi^J\right),$$

where

$$\chi = \sum_{i=0}^{M} \varphi_i \Phi_i\left(\xi^1, \dots, \xi^J\right)$$

is a function of the basic random variables. Therefore, the savings policy is a function of $(\xi^1, \dots, \xi^J)$ as well.

Proposition 6 (Analytic policies) *Under the assumptions of Theorem 4, all iterates of the the savings policy and the Lagrange multipliers as functions of $\xi^1, \dots, \xi^J$ admit analytic extensions in the complex plane, namely in the region $\Sigma\left(\tau_{n+1}^j, \Gamma^j\right)$, $j \in \{1, \dots, J\}$, given in Theorem 4. Furthermore, it holds that the $(n+1)$-th policy iterate*

$$\max_{\tilde{\xi} \in \Sigma(\tau_{n+1}^j, \Gamma^j)} \left| h^{n+1}\left(\tilde{\xi}\right) \right| \leq \frac{\min(1, \mathbf{L}_{n+1})}{\min(1, \mathbf{L}_{n+1}) - 2\tau_{n+1}^j A_{n+1,j}^1}.$$

is bounded in the region $\Sigma\left(\tau_{n+1}^j, \Gamma^j\right)$, $j \in \{1, \dots, J\}$.

PROOF: The proof now proceeds in two steps. First, I establish that all iterates of the policy and hence, Lagrange multipliers are real analytic functions of the basic random variables. Second, I construct the complex analytic extension.

Real analytic: Equation (13) implies that the $(n+1)$-th iterate of the savings policy h^{n+1} in the proximal point algorithm solves the following first-order condition

$$(B(z', x, \varphi_0) - \bar{x}) \frac{\partial}{\partial h^{n+1}} L^A \left(h^{n+1}, y^n; h^n\right) = \mathbf{e} \tag{25}$$

for a.e. (z', x) with constant $\|\mathbf{e}\| \leq \frac{\epsilon^2}{2\lambda}$. Now, let us take the derivatives of the first-order condition (25) w.r.t. ξ^j, $j \in \{1, \ldots, J\}$. It is obvious that B and $\mathbf{e}$ do not depend on the basic random variables. The partial derivative of the augmented Lagrangian, however, does due to its dependence on X'' and because the optimal policies and hence, also the Lagrange multipliers depend on $(\xi^1, \ldots, \xi^J)$ as can be seen in

$$\begin{aligned} & \frac{\partial}{\partial h^{n+1}} L^A \left(h^{n+1}, y^n; h^n\right) \\ = \; & \frac{\partial}{\partial c} u \left(B(z', x, \varphi_0) - h^{n+1}\right) \\ & -\beta \sum_{z'' \in \mathcal{Z}} \left\{ p^{z''|z'} \frac{\partial}{\partial x} B(z'', h^{n+1}, X'') \frac{\partial}{\partial c} u \left(B(z'', h^{n+1}, X'') - h^{(n+1)'} \right) \right\} \\ & + \frac{1}{\lambda} \left(h^{n+1} - h^n\right) - \left(y^n + \lambda \left[\bar{x} - h^{n+1}\right]\right) \mathbb{1}_{\{\lambda(\bar{x} - h^{n+1}) \geq y^n\}}. \end{aligned} \tag{26}$$

I now investigate the derivatives of (25) w.r.t. the basic random variable ξ^j, $j \in \{1, \ldots, J\}$. Trivially,

$$D_j^p \left(\frac{\partial}{\partial h^{n+1}} L^A \right) = 0, \; p \in \{1, 2, \ldots\}.$$

It follows for $p = 1$ that

$$\frac{\partial^2 L^A}{\partial h^{n+1} \partial X''} D_j^1 X'' + \frac{\partial^2 L^A}{[\partial h^{n+1}]^2} D_j^1 h^{n+1} - \frac{1}{\lambda} D_j^1 h^n - \mathbb{1}_{\{\lambda(\bar{x} - h^{n+1}) \geq y^n\}} D_j^1 y^n = 0$$

and thus,

$$D_j^{p-1} \left(\frac{\partial^2 L^A}{\partial h^{n+1} \partial X''} D_j^1 X'' + \frac{\partial^2 L^A}{[\partial h^{n+1}]^2} D_j^1 h^{n+1} \right) = \frac{1}{\lambda} D_j^p h^n + \mathbb{1}_{\{\lambda(\bar{x} - h^{n+1}) \geq y^n\}} D_j^p y^n. \tag{27}$$

Let us first analyze the derivative of X''. It is easy to see from $D_j^1 X'' = \partial / \partial \chi X'' D_j^1 \chi$

that all derivatives of X'' are composed of the derivatives of the optimal policy w.r.t. start capital and the derivatives of χ w.r.t. the basic random variables. The latter component is obviously analytic. Hence, X'' is analytic in the basic random variables if the optimal policy is analytic in start capital. This fact is easily established by induction when taking derivatives of the first-order condition w.r.t. x and taking into account that h^0 is analytic in x. I exploit the fact that products, sums and compositions of analytic functions are analytic. Hence, there is a $\mathbf{c}_{X,j}$ such that

$$\left\|D_j^p X''\right\| \leq \mathbf{c}_{X,j}^p p!,\ p \in \{1, 2, \ldots\}.$$

Furthermore, note that it follows from (26) that $\partial/\partial h^{n+1} L^A$ is analytic in X'' and h^{n+1}.

I can now show analyticity of the optimal policy by induction in two dimensions: First increasing the iterate of the policy $n \to n+1$, then increasing the order of the derivative $p-1 \to p$. Assume that all iterates h^j, $j \leq n$, are analytic and that the derivatives of h^{n+1} w.r.t. the basic random variables up to order $p-1$ are bounded as required for analyticity. This implies that the derivatives w.r.t. X'' and h^{n+1} of the first-order condition also satisfy the analyticity condition up to order $p-1$ with coefficients $\mathbf{c}_{L_X^A,j}$ and $\mathbf{c}_{L_h^A,j}$. W.l.o.g., choose $\mathbf{c}_{L_X^A,j} \geq \mathbf{c}_{L_h^A,j}$. Applying the product rule, I rewrite (27) as

$$\begin{aligned}
\frac{\partial^2 L^A}{[\partial h^{n+1}]^2} D_j^p h^{n+1} =& \frac{1}{\lambda} D_j^p h^n + \mathbb{1}_{\{\lambda(\bar{x}-h^{n+1}) \geq y^n\}} D_j^p y^n \\
&- \sum_{l=0}^{p-1} \binom{p-1}{l} D_j^{p-l} X'' D_j^l \left(\frac{\partial^2 L^A}{\partial h^{n+1} \partial X''} \right) \\
&- \sum_{l=1}^{p-1} \binom{p-1}{l} D_j^{p-l} h^{n+1} D_j^l \left(\frac{\partial^2 L^A}{[\partial h^{n+1}]^2} \right).
\end{aligned}$$

Dividing by $p!$, taking norms and denoting $R_{n+1,j}^p = \|D_j^p h^{n+1}\|/p!$ leads to

$$\begin{aligned}
\underbrace{\frac{\partial^2 L^A}{[\partial h^{n+1}]^2}}_{=:\mathbf{L}_{n+1}} R_{n+1,j}^p \leq& \frac{1}{\lambda} \mathbf{c}_{h^n,j}^p + \mathbb{1}_{\{\lambda(\bar{x}-h^{n+1}) \geq y^n\}} \mathbf{c}_{y^n,j}^p + \sum_{l=0}^{p-1} \mathbf{c}_{X,j}^{p-l} \mathbf{c}_{L_X^A,j}^l \\
&+ \sum_{l=1}^{p-1} R_{n+1,j}^{p-l} \mathbf{c}_{L_h^A,j}^l
\end{aligned}$$

$$\leq \underbrace{\frac{1}{\lambda}\mathbf{c}^p_{h^n,j} + \mathbb{1}_{\{\lambda(\bar{x}-h^{n+1})\geq y^n\}}\mathbf{c}^p_{y^n,j} + \max\left(2\mathbf{c}_{X,j}, 2\mathbf{c}_{L^A_X,j}\right)^p}_{=:A^p_{n+1,j}} \tag{28}$$
$$+ \sum_{l=1}^{p-1} R^{p-l}_{n+1,j}\mathbf{c}^l_{L^A_h,j}.$$

Note that, due to convexity, $\mathbf{L}_{n+1} > 0$. Solving this recursion yields

$$R^p_{n+1,j} \leq \frac{A^p_{n+1,j}}{\mathbf{L}_{n+1}} + \sum_{l=0}^{p-1} \frac{A^{p-1-l}_{n+1,j}}{\mathbf{L}_{n+1}} 2^l \mathbf{c}^{l+1}_{L^A_h} \leq 2^p \frac{A^p_{n+1,j}}{\mathbf{L}_{n+1}} \leq \left(\frac{2A^1_{n+1,j}}{\min(1, \mathbf{L}_{n+1})}\right)^p,$$

where $A_0 = \|\mathbf{e}\|/X$. Hence, I obtain a uniform bound for all derivatives of the optimal policy. Analyticity follows by induction.

Complex continuation: I define the following power series for the $(n+1)$-th iterate in terms of the basic random variable ξ^j, $j \in \{1, \ldots, J\}$, on the complex plane

$$h^{n+1}\left(\tilde{\xi}\right) = \sum_{p=0}^{\infty} \frac{\left(\tilde{\xi} - \xi^j\right)^p}{p!} D^p_j h^{n+1}.$$

Taking norms leads to

$$\left|h^{n+1}\left(\tilde{\xi}\right)\right| = \sum_{p=0}^{\infty} \left(\left|\tilde{\xi} - \xi^j\right| \frac{2A^1_{n+1,j}}{\min(1, \mathbf{L}_{n+1})}\right)^p.$$

This series converges for all $|\tilde{\xi} - \xi^j| \leq \tau^j_{n+1} < \frac{\min(1, \mathbf{L}_{n+1})}{2A^1_{n+1,j}}$ such that

$$\left|h^{n+1}\left(\tilde{\xi}\right)\right| \leq \frac{\min(1, \mathbf{L}_{n+1})}{\min(1, \mathbf{L}_{n+1}) - 2\tau^j_{n+1}A^1_{n+1,j}}.$$

Therefore, by continuation the iterates can be extended analytically in the whole region $\Sigma(\tau^j_{n+1}, \Gamma^j)$, which concludes the proof. *Q.E.D.*

Remark Note that I follow the proof of Theorem 4.1 in Babuška et al. (2007) with bounded range of the basic random variables for the proof of Theorem 4 below. I use bounded range since I choose a histogram approximation of the basic random variables. For other types of approximation, one might need to modify the error bound estimates to accommodate an unbounded range. I refer to Babuška et al. (2007) for that case.

Proof of Theorem 4: The last term of the bound is the interpolation error from tensor-product finite elements of order 1 on a rectangular discretization D. It is well established (see e.g., Brenner and Scott, 2007, Theorem 4.6.14). The error bound due to truncation of the polynomial chaos expansion is a little more involved. Due to the fact that the continuous functions of the basic random variables are a subset of the square-integrable functions, i.e., $C^0(\Gamma^j) \subset L^2(\Gamma^j)$, we have that the truncation error is bounded by the best approximation error (see Babuška et al., 2007, Lemma 4.3)

$$\left\| h - h^M \right\|_{L^2} \leq \mathbf{b} \inf_{w \in \mathcal{H}^M} \| h - w \|_{C^0} ,$$

where constant $\mathbf{b}$ is independent of the order of truncation M. Given that h admits an analytic extension on the complex plane $\Sigma(\tau^j_{n+1}, \Gamma^j)$, the best approximation error is bounded by (see Babuška et al., 2007, Lemma 4.4)

$$\inf_{w \in \mathcal{H}^M} \| h - w \|_{C^0(\Gamma^j)} \leq \frac{2}{\eta^j - 1} e^{-M \log(\eta^j)} \max_{\tilde{\xi} \in \Sigma(\tau^j_{n+1}, \Gamma^j)} \left| h^{n+1} \left(\tilde{\xi} \right) \right| , \; j \in \{1, \ldots, J\},$$

where

$$\eta^j = \frac{2\tau^j_{n+1}}{|\Gamma^j|} + \sqrt{1 + \frac{4(\tau^j_{n+1})^2}{|\Gamma^j|^2}} > 1.$$

Combining this with Proposition 6 and keeping in mind that I truncate once for each ξ^j leads to the truncation error bound. *Q.E.D.*

REFERENCES

Yves Achdou, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. Income and wealth distribution in macroeconomics: A continuous-time approach. *Working Paper*, 2017.

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

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

Yann Algan, Olivier Allais, and Wouter J. Den Haan. Solving the incomplete markets model with aggregate uncertainty using parameterized cross-sectional distributions. *Journal of Economic Dynamics and Control*, 34(1):59–68, 2010.

Ivo Babuška, Fabio Nobile, and Raúl Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 45(3):1005–1034, 2007.

Truman Bewley. The permanent income hypothesis: A theoretical formulation. *Journal of Economic Theory*, 16(2):252–292, 1977.

Timo Boppart, Per Krusell, and Kurt Mitman. Exploiting MIT shocks in heterogeneous-agent economies: the impulse response as a numerical derivative. *Journal of Economic Dynamics and Control*, 89:68–92, 2018.

Susanne Brenner and Ridgway Scott. *The Mathematical Theory of Finite Element Methods*, volume 15 of *Texts in Applied Mathematics*. Springer, 2007.

Markus K. Brunnermeier and Yuliy Sannikov. A macroeconomic model with a financial sector. *The American Economic Review*, 104(2):379–421, 2014.

David Childers. On the solution and application of rational expectations models with function-valued states. *Working Paper*, 2015.

Wilbur John Coleman. Solving the stochastic growth model by policy-function iteration. *Journal of Business and Economic Statistics*, 8(1):27–29, 1990.

George M. Constantinides and Darrell Duffie. Asset pricing with heterogeneous consumers. *Journal of Political Economy*, 104(2):219–240, 1996.

Wouter J. den Haan. Solving dynamic models with aggregate shocks and heterogenous agents. *Macroeconomic Dynamics*, 1(02):355–386, 1997.

Wouter J. den Haan. Comparison of solutions to the incomplete markets model with aggregate uncertainty. *Journal of Economic Dynamics and Control*, 34(1):4–27, 2010.

Wouter J. den Haan and Pontus Rendahl. Solving the incomplete markets model with aggregate uncertainty using explicit aggregation. *Journal of Economic Dynamics and Control*, 34(1): 69–78, 2010.

Wouter J. den Haan, Kenneth L. Judd, and Michel Juillard. Computational suite of models with heterogeneous agents: Incomplete markets and aggregate uncertainty. *Journal of Economic Dynamics and Control*, 34(1):1–3, 2010.

Oliver G. Ernst, Antje Mugler, Hans-Jörg Starkloff, and Elisabeth Ullmann. On the convergence of generalized polynomial chaos expansions. *ESAIM: Mathematical Modelling and Numerical Analysis*, 46(02):317–339, 2012.

Walter Gautschi. On generating orthogonal polynomials. *SIAM Journal on Scientific and Statistical Computing*, 3(3):289–317, 1982.

Osman Güler. On the convergence of the proximal point algorithm for convex minimization. *SIAM Journal on Control and Optimization*, 29(2):403–419, 1991.

Osman Güler. New proximal point algorithms for convex minimization. *SIAM Journal on Optimization*, 2(4):649–664, 1992.

Hugo A. Hopenhayn. Entry, exit, and firm dynamics in long run equilibrium. *Econometrica*, 60 (5):1127–1150, 1992.

Mark Huggett. The risk-free rate in heterogeneous-agent incomplete-insurance economies. *Journal of Economic Dynamics and Control*, 17(5):953–969, 1993.

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

Aubhik Khan and Julia K. Thomas. Idiosyncratic shocks and the role of nonconvexities in plant and aggregate investment dynamics. *Econometrica*, 76(2):395–436, 2008.

Dirk Krueger, Kurt Mitman, and Fabrizio Perri. *Macroeconomics and Household Heterogeneity*, pages 843–921. Handbook of macroeconomics. Elsevier, 2016.

Per Krusell and Anthony A. Smith. Income and wealth heterogeneity in the macroeconomy.

Journal of Political Economy, 106(5):867–896, 1998.

Per Krusell, Toshihiko Mukoyama, and Anthony A. Smith. Asset prices in a Huggett economy. *Journal of Economic Theory*, 146(3):812–844, 2011.

Felix Kubler and Karl Schmedders. Recursive equilibria in economies with incomplete markets. *Macroeconomic dynamics*, 6(02):284–306, 2002.

Huiyu Li and John Stachurski. Solving the income fluctuation problem with unbounded rewards. *Journal of Economic Dynamics and Control*, 45:353–365, August 2014 2014.

Lilia Maliar, Serguei Maliar, and Fernando Valli. Solving the incomplete markets model with aggregate uncertainty using the Krusell-Smith algorithm. *Journal of Economic Dynamics and Control*, 34(1):42–49, 2010.

Thomas Mertens and Kenneth Judd. Equilibrium existence and approximation for incomplete market models with substantial heterogeneity. *Working Paper*, 2013.

Elisabeth Pröhl. Existence and uniqueness of recursive equilibria with aggregate and idiosyncratic risk. *Working Paper, Available at SSRN 3250651*, 2018.

Michael Reiter. Solving heterogeneous-agent models by projection and perturbation. *Journal of Economic Dynamics and Control*, 33(3):649–665, 2009.

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

Michael Reiter. Approximate and almost-exact aggregation in dynamic stochastic heterogeneous-agent models. *Economics Series Working Paper 258*, 2010b.

José-Víctor Ríos-Rull. Computation of equilibria in heterogeneous agent models. *Federal Reserve Bank of Minneapolis Staff Report 231*, pages 238–264, 1997.

R. Tyrrell Rockafellar. Monotone operators and the proximal point algorithm. *SIAM Journal on Control and Optimization*, 14(5):877–898, 1976a.

R. Tyrrell Rockafellar. Augmented Lagrangians and applications of the proximal point algorithm in convex programming. *Mathematics of Operations Research*, 1(2):97–116, 1976b.

Saverio Salzo and Silvia Villa. Inexact and accelerated proximal point algorithms. *Journal of Convex Analysis*, 19(4):1167–1192, 2012.

Kjetil Storesletten, Christopher I. Telmer, and Amir Yaron. Asset pricing with idiosyncratic risk and overlapping generations. *Review of Economic Dynamics*, 10(4):519–548, October 2007.

Yeneng Sun. The exact law of large numbers via Fubini extension and characterization of insurable risks. *Journal of Economic Theory*, 126(1):31–69, January 2006.

Thomas Winberry. A method for solving and estimating heterogeneous agent macro models. *forthcoming in Quantitative Economics*, 2018.

Dongbin Xiu and George E. Karniadakis. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM Journal on Scientific Computing*, 24(2):619–644, 2002.

Eric R. Young. 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, 2010.

Mengdi Zheng, Xiaoliang Wan, and George E. Karniadakis. Adaptive multi-element polynomial chaos with discrete measure: Algorithms and application to SPDEs. *Applied Numerical Mathematics*, 90:91–110, 2015.