DISCRETIZING THE INFINITE-DIMENSIONAL DISTRIBUTION SPACE TO APPROXIMATE MARKOV EQUILIBRIA WITH EX-POST HETEROGENEITY AND AGGREGATE RISK

Elisabeth Pröhl

University of Geneva and Swiss Finance Institute

November 1, 2016

Dynamic stochastic general equilibrium models with ex-post heterogeneity due to idiosyncratic risk typically have to be solved numerically. However, as the cross-sectional distribution of endogenous variables becomes an element of the state space due to the aggregate risk, this is a nontrivial task. Existing algorithms assume bounded rationality of the agents to reduce dimensionality meaning that they only keep track of a limited number of moments of the cross-sectional distribution. In this paper, we do not take that assumption and compute a recursive fully rational equilibrium which depends on the whole state distribution. Dimensionality is tackled by polynomial chaos expansions, a projection technique for square-integrable random variables. Hence, this algorithm is a global solution method relying solely on projection. Another important distinction w.r.t. the existing algorithms is that a theoretical convergence result is given ensuring that the solution is indeed close to a recursive rational equilibrium.

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

JEL Classification: C63, C68, D31, D52, D58, E21.

University of Geneva, Bd du Pont d'Arve 40, 1211 Geneva 4, Switzerland, elisabeth.proehl@unige.ch

I thank Johannes Brumm, René Carmona, David Childers, Jérôme Detemple, Rajna Gibson Brandon, Kenneth Judd, Greg Kaplan, Felix Kübler, Semyon Malamud, Benjamin Moll, Juan Sagredo, Yuliy Sannikov, Olivier Scaillet, Karl Schmedders, Harald Uhlig, Thomas Winberry, and participants at ESEM 2016, EcoMod 2016, CEF 2016, ESNASM 2016, the MFM Summer Camp 2016, the 12th Cowles Conference on General Equilibrium and its Applications, ISCEF 2016, the MFM Winter Meeting 2016, CEF 2015, the SFI Research Days 2015, PASC 2015, and seminar participants in Chicago, Princeton, Lausanne, Zurich and Geneva for fruitful comments on this work. Earlier versions of this work were disseminated under the name "Computing the Cross-Sectional Distribution to Approximate Stationary Markov Equilibria with Heterogeneous Agents and Incomplete Markets".

1. INTRODUCTION

Economies consist of heterogeneous agents who are exposed to 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. In general, idiosyncratic risks affect aggregate variables in the economy. Other examples of idiosyncratic risks are firm-specific productivity shocks in models of firm exit and entry as in Hopenhayn (1992) or county-specific productivity shocks in real business cycle models as in den Haan et al. (2011).

These models, however, do not feature aggregate risk because it makes the equilibrium problem difficult to solve. The challenge in the construction of solution algorithms lies in the cross-sectional distribution of the agents' characteristics which becomes an infinite-dimensional element of the state space. Krusell and Smith (1998) were the first to propose a numerical solution algorithm for the Aiyagari growth model with aggregate risk. They handle the dimensionality problem in assuming bounded rationality of the agents. It means that the agents can only observe a limited number of moments of the cross-sectional distribution to decide on their policy. The authors solve for the optimal policy and the law of motion of these moments in a two-step iterative procedure. In the first step, the policy function is computed by iterating the Euler equation for a fixed law of motion. In the second step, the exogenous shocks are simulated across time and agents and the policy is iteratively applied to compute individual capital. The new law of motion is inferred from this simulated data. Various more recent papers improve the original algorithm especially by eliminating the agent dimension in the simulation step. However, these works still rely on the bounded rationality assumption and a

two-step iterative procedure.

There are several drawbacks to the Krusell-Smith algorithm and its more recent substitutes of similar style. Firstly, it is not clear a priori how many moments are necessary for the equilibrium to exist. In fact, it is shown in Kubler and Schmedders (2002) 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. Moreover, Kubler and Schmedders (2002) argue that it is not feasible to compute more general recursive equilibria because they and especially their ergodic state distributions are too abstract. Secondly, there are no theoretical convergence results for Krusell-Smith-style algorithms. Therefore, it is not clear whether they approach the true equilibrium if they converge numerically.

To the best of my knowledge, there is only one algorithm which does not rely on the bounded rationality assumption, namely Reiter (2009, 2010b). This algorithm first solves for the optimal policy and stationary distribution of the model without aggregate shocks using projection methods and then perturbs this solution to accommodate aggregate shocks. There are two major drawbacks. Firstly, the perturbation in aggregate shocks often is only linear. Therefore, any higher-order nonlinear effects of aggregate shocks are not accounted for. Secondly, as for all perturbation methods the solutions are only accurate for small aggregate shocks. Crises scenarios in terms of a large aggregate shock or a long series of aggregate shocks in one direction cannot be analyzed with confidence.

The contribution of this paper is to construct a general solution algorithm for DSGE models with heterogeneous agents and aggregate risk for which convergence is proven. We do not assume bounded rationality. Instead, we compute recursive equilibria which depend on the whole cross-sectional distribution. Furthermore, the dependance on the aggregate risk can be nonlinear and crisis scenarios can be analyzed as we do not rely on perturbations around a model without aggregate risk. In this algorithm, the dimensionality of the state space is reduced by param-

eterizing the cross-sectional distribution using generalized polynomial chaos. This is a technique going back to Xiu and Karniadakis (2002) which allows to project any square-integrable random variable onto orthogonal polynomials taking basic random variables as arguments. Ernst et al. (2012) provides a convergence result for this projection technique. Secondly, this work ensures convergence of the full algorithm proposed herein by leveraging on the structure of the underlying optimization problem and utilizing the concept of proximal points as laid out in Rockafellar (1970, 1976a,b); Güler (1992); Salzo and Villa (2012).

This paper is related to several strands of literature. The generic existence of solutions to DSGE models has been shown by Duffie and Shafer (1985, 1986) and Duffie et al. (1994). However, these results only apply to models with ex-ante heterogeneity, i.e. where agents differ on finitely many model ingredients. Existence of a solution to the Aiyagari-Bewley growth model with aggregate risk which features ex-post heterogeneity due to idiosyncratic risk has long been an open research question. It has been examined by Miao (2006). A slight flaw in the theoretical argument has been discovered in Cheridito and Sagredo (2016b) and corrected in Cheridito and Sagredo (2016a).

Furthermore, the literature on numerical algorithms for DSGE models is related. 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 and compares them in den Haan (2010). They have in common that they use a small finite number of moments 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) 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). They use parametric and nonparametric procedures to get around this issue. However, the variation due to

simulating over common exogenous shocks remains. Reiter (2010a) also parameterizes the cross-sectional distribution whereas den Haan and Rendahl (2010) use direct aggregation to obtain the moments of the cross-sectional distribution. All of the mentioned algorithms use projection methods or a hybrid of projection and simulation methods. The algorithm in Kim et al. (2010) differs as it utilizes a perturbation method. However, their approach includes a very simplified law of motion. All of these algorithms still rely on bounded rationality and component-wise fixed-point iteration for which theoretical convergence results are outstanding. One approach which differs from the aforementioned is the one by Mertens and Judd (2013). Their algorithm relies on a perturbation method as the authors use an approximation of the law of motion which is asymptotically true. They do provide a convergence result for their algorithm which is, however, subject to the criticism of Cheridito and Sagredo (2016b).

Another strand of literature focuses on solutions to mean field games. This field has been established by P.-L. Lions in lectures at the Collège de France. The present paper is related to that research area because mean field games are essentially continuous-time versions of DSGE models with heterogeneity and incompleteness. There are two different approaches to solve these models which are compared in Carmona et al. (2013), a PDE approach and a probabilistic approach. The construction of our algorithm is closer in spirit to the latter. Recently, the research on mean field games turned to numerical solutions in Achdou et al. (2014). They use partial differential equations to solve heterogeneous agent models. However, they are not able to solve models which include aggregate exogenous shocks. Their models solely incorporate idiosyncratic shocks.

The paper proceeds as follows. In the next section, we present the Aiyagari-Bewley growth model with aggregate risk which illustrates our algorithm throughout the paper. In Section 3, we introduce the methodology behind the algorithm. It explains the concept of proximal point algorithms which underlies our conver-

gence result and the polynomial chaos expansion which is used to parameterize the cross-sectional distribution. Section 4 then proves the convergence result. In Section 5, our numerical results are compared to the ones from the Krusell-Smith algorithm. The last section concludes. The appendix contains all proofs.

2. THE MODEL

For illustration, we use the same growth model with aggregate shocks as in den Haan et al. (2010) which is used for a comparison of Krusell-Smith-style algorithms in the special issue of the Journal of Economic Dynamics and Control in January 2010. We consider a discrete-time infinite-horizon model 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. The idiosyncratic shock with outcomes in $\mathcal{Z}^{id} = \{0, 1\}$ indicates that an agent is unemployed or employed, respectively. It is i.i.d. across agents conditional on the aggregate shock. We 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 security market consists of a claim to aggregate capital $(K_t)_{t\geq 0}$. An agent's share of physical capital is denoted by $(k_t)_{t\geq 0}$. The aggregate endogenous variable K is hence defined by

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

where μ_t is the cross-sectional distribution of idiosyncratic endogenous variables at time t. It is simply the probability distribution of individual capital across the unemployed and the employed agents given the trajectory of aggregate shocks

$$\mu_t\left(z^{id}, k\right) = \mathbb{P}\left(\left\{z_t^{id} = z^{id}\right\} \cap \left\{k_t \leq k\right\} \middle| \, z_t^{ag}, \ldots, z_0^{ag}\right)$$

for all $t \geq 0$, $z^{id} \in \mathcal{Z}^{id}$ and $k \in \mathbb{R}$. The aggregate shocks cause the cross-sectional distribution to vary over time which is indicated by the time subscript of μ_t.

Each agent chooses her share of physical capital and consumption such that they satisfy certain constraints. Firstly, individual consumption must be positive at all times $c_t > 0$, $t \geq 0$, and capital holdings are subject to a hard borrowing constraint $k_t \geq 0$, $t \geq 0$. Secondly, given an initial capital endowment $k_{-1} \geq 0$ and an initial cross-sectional distribution μ_{-1}[1] with non-negative support, each agent adheres to a budget constraint which equates individual consumption and current capital stock to productive income and saved capital stock

$$k_t + c_t = I\left(z_t, K_{t-1}\right) + \left[1 - \rho\right] k_{t-1} \,\forall\, t \geq 0. \tag{2}$$

The time line underlying this equation is clarified in Figure 1.[2] The parameters in

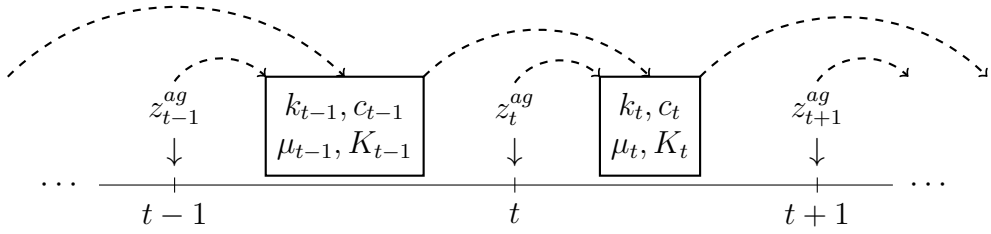

FIGURE 1. **Time line of events.** Before period t, the agent observes how much capital she saved in the previous period k_{t-1} and what the cross-sectional distribution of individual capital savings μ_{t-1} and hence, aggregate capital K_{t-1} is. At period t, the agent first observes the exogenous shocks z_t and then decides how much to consume c_t and how much capital k_t to save in that period.

this budget constraint are defined as follows. The capital stock brought forward

[1] The initial cross-sectional distribution μ_{-1} does not only imply the initial aggregate capital K_{-1}, but also the initial aggregate economic state due to $p^e_{-1} = (1/K_{-1}) \int_0^\infty k d\mu_{-1}(1,k)$ which is an exogenously given quantity.

[2] Note that I specify the time line slightly differently than den Haan et al. (2010) and Krusell and Smith (1998). These authors substitute k_t with k_{t+1} in the budget constraint (2) because this is the capital which is put forward as start capital to period $t+1$. In contrast to that notation, however, I want to emphasize the time period at which the agent optimally chooses the magnitude of her capital savings. Taking this view, the optimal consumption and capital savings choice have the same time subscript. My time line therefore indicates which filtration the endogenous variables are adapted to.

from period $t-1$ depreciates by a rate $\rho \in (0,1)$. The productive income is given by

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

It is composed of, firstly, the return on capital stock and, secondly, 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 $\tau_t = \nu(1-p_t^e)/p_t^e$ 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. It is reasonable to assume $\nu < 1-\tau_t \Leftrightarrow \nu < p_t^e$ 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

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

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 a time endowment factor $\pi > 0$.

We assume that all agents have time-separable CRRA utility with a risk aversion coefficient $\gamma > 1$ and time preference parameter $\beta \in (0,1)$. Then, given an agent's initial capital endowment $k_{-1} \geq 0$ and the initial cross-sectional distribution μ_{-1} with non-negative support, the individual optimization problem reads

$$
\begin{aligned}
(4) \qquad \max_{\{c_t, k_t\} \in \mathbb{R}^2} \; & \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\gamma}-1}{1-\gamma}\right] \\
\text{s.t. } & k_t + c_t = I\left(z_t, K_{t-1}\right) + \left[1-\rho\right] k_{t-1} \, \forall \, t \geq 0 \\
& c_t > 0, \; k_t \geq 0 \, \forall \, t \geq 0
\end{aligned}
$$

where the productive income I is defined as in (3).

In a competitive equilibrium, the individual problems are solved subject to the market condition (1) that aggregate capital equals the expected optimal individual capital holdings. The question of existence of a competitive equilibrium, in particular one which has a recursive form, has first been examined by Miao (2006) and has later been solved by Cheridito and Sagredo (2016a,b). 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) subject to the market condition (1) given an initial cross-sectional distribution of individual capital μ_{-1} with non-negative support is called recursive if there exist functions $h_i : \mathcal{Z} \times \mathbb{R} \times \mathcal{P}(\mathcal{Z}^{id} \times \mathbb{R}) \to \mathbb{R}$, $i \in \{1, 2\}$, such that, for any point in time, the current optimal consumption and capital savings choices equal*

$$c' = h_1(z', k, \mu)$$
$$k' = h_2(z', k, \mu)$$

for any agent with previous-period capital stock k who observes the previous-period cross-sectional distribution μ and the current-period exogenous shock $z' = (z^{ag'}, z^{id'})$.

Recursive equilibria of models with ex-post heterogeneity can rarely be computed in closed form such that they have to be numerically approximated. When designing a numerical solution algorithm, it is important to show theoretically that the algorithm converges to a true equilibrium. This is the goal of this work. To obtain the convergence result, however, we take a different point of view regarding

the functions for the optimal choices. In the existing literature, these functions are approximated point-wise w.r.t. the endogenous arguments whereas here, we approximate these functions in terms of distributions. Hence, I make the following assumption.

ASSUMPTION 2 (Square-integrability) (*i*) The initial distribution μ_{-1} and the cross-sectional distributions at any point in time $\{\mu_t\}_{t\geq 0}$ are square-integrable.

(*ii*) Given any cross-sectional distribution μ, the corresponding functions for the optimal consumption and savings choice $h_i(z^{ag'}, \zeta, \kappa, \mu)$, $i \in \{1, 2\}$, are square-integrable w.r.t. the idiosyncratic random variables $(\zeta, \kappa) \sim \mu$, i.e. in short-hand notation $h_1, h_2 \in L^2(\mathcal{Z}^{id} \times \mathbb{R}, \mathcal{B}(\mathcal{Z}^{id} \times \mathbb{R}), \mu)$.

In order to obtain a full description of equilibrium, we need to define the consistent law of motion of μ to μ'. Given a fixed distribution μ over the cross-section of individual capital at the end of the previous period and a recursive equilibrium, the distribution in the current period changes in two steps $\mu \to \tilde{\mu}' \to \mu'$. In the first step, the new shocks z' for all agents realize and shift the quantities of employed and unemployed agents depending on the outcome of the aggregate shock. Formerly employed agents either stay employed or become unemployed, the same holds for the formerly unemployed. Therefore, the distribution at the beginning of the current period $\tilde{\mu}'$ is given by

$$\begin{aligned}(5)\qquad \tilde{\mu}'\left(z^{id'}, k\right) &= \sum_{z^{id}\in\mathcal{Z}^{id}} \frac{p^{(z^{ag'}, z^{id'})|(z^{ag}, z^{id})}}{p^{z^{ag'}|z^{ag}}} \mu\left(z^{id}, k\right) \\ &= \sum_{z^{id}\in\mathcal{Z}^{id}} \frac{p^{(z^{ag'}, z^{id'})|(z^{ag}, z^{id})}}{p^{z^{ag'}|z^{ag}}} \mathbb{P}\left(\left\{\zeta = z^{id}\right\} \cap \left\{\kappa \leq k\right\} \middle| z^{ag}\right)\end{aligned}$$

for all $z^{id'} \in \mathcal{Z}^{id}$ and $k \in \mathbb{R}$. The multipliers in front of the previous distribution are the probabilities that the employment status changes from z^{id} to $z^{id'}$ given the observed trajectory of z^{ag} to $z^{ag'}$. In the second step, the agents implement their

optimal capital savings which leads to the new current-period distribution

$$(6)\qquad \mu'\left(z^{id'},k\right)=\sum_{z^{id}\in\mathcal{Z}^{id}}\frac{p^{(z^{ag'},z^{id'})|(z^{ag},z^{id})}}{p^{z^{ag'}|z^{ag}}}\mathbb{P}\left(\left\{\zeta=z^{id}\right\}\cap\left\{h_2\left(z^{ag'},z^{id'},\kappa,\mu\right)\leq k\right\}\middle|\, z^{ag'},z^{ag}\right).$$

From this definition of the new distribution, the new aggregate capital K' follows immediately due to (1). Now that all model ingredients are defined, the next section lays out the methodology to compute the recursive equilibrium.

3. THE METHODOLOGY

The methodology proposed herein builds on the observation that the optimal policy functions h_1 for consumption and h_2 for capital savings solve the Euler equation which is equivalent to the first-order condition of the following constrained optimization problem

$$(7)\qquad \min_{\{h_1,h_2\}}\ -u(h_1)-\sum_{z''\in\mathcal{Z}}p^{z''|z'}\beta u\left(I\left(z'',K'\right)+\left[1-\rho\right]h_2-h_2\circ h_2\right)$$
$$\text{s.t. } 0=I\left(z',K\right)+\left[1-\rho\right]k-h_1-h_2$$
$$0\geq -h_1,\ 0\geq -h_2.$$

The utility function $u:\mathbb{R}_{>0}\to\mathbb{R}$, $c\mapsto\frac{1}{1-\gamma}(c^{1-\gamma}-1)$ is defined as in (4), $p^{z''|z'}$ is the exogenously given transition probability that z' is followed by z'' and I is as in (3).

The standard way of solving a constrained optimization problem is to set up the corresponding Lagrangian and to find a saddle point by minimizing over the policies and maximizing over the Lagrange multipliers. This can be done using the proximal point algorithm as explained subsequently. However, complications arise due to the ex-post heterogeneity which introduces the cross-sectional distribution to the state space. We have to discretize the space of distributions which

is an infinite-dimensional object. A solution to this challenge is discussed after introducing the proximal point algorithm.

3.1. *The Proximal Point Algorithm*

Applying the proximal point algorithm to minimax problems on saddle functions like Lagrangians goes back to a series of papers by Rockafellar (1970, 1976a,b) who sets the theoretical base for the convergence of said algorithm. Let me first define what a saddle function is in this context.

DEFINITION 3 (Saddle function (see Rockafellar, 1970)) *(i) Let $\mathcal{C}$ and $\mathcal{D}$ be Hilbert spaces over $\mathbb{R}$. A saddle-function is an everywhere-defined function $L : \mathcal{C} \times \mathcal{D} \to [-\infty, \infty]$ such that $L(c, d)$ is a convex function of $c \in \mathcal{C}$ for any $d \in \mathcal{D}$ and a concave function of $d \in \mathcal{D}$ for any $c \in \mathcal{C}$.*

(ii) A saddle function is called proper if there exists a point $(c, d) \in \mathcal{C} \times \mathcal{D}$ with $L(c, \tilde{d}) < +\infty$ for any $\tilde{d} \in \mathcal{D}$ and $L(\tilde{c}, d) > -\infty$ for any $\tilde{c} \in \mathcal{C}$.

(iii) The operator associated with the saddle function L is defined as the set-valued mapping

$$\begin{aligned}\mathbf{T}_L(c, d) = \{(v, w) | & L(\tilde{c}, d) - \langle \tilde{c}, v \rangle + \langle d, w \rangle \\ & \geq L(c, d) - \langle c, v \rangle + \langle d, w \rangle \\ & \geq L(c, \tilde{d}) - \langle c, v \rangle + \langle \tilde{d}, w \rangle \, \forall (\tilde{c}, \tilde{d}) \in \mathcal{C} \times \mathcal{D} \Big\}.\end{aligned}$$

A saddle point is a point $(c^*, d^*) \in \mathcal{C} \times \mathcal{D}$ such that $0 \in \mathbf{T}_L(c^*, d^*)$.[3]

According to this definition, finding a recursive equilibrium translates into the following task: Given that a zero of the operator $\mathbf{T}_L$ associated with the saddle function of (7) exists (see Cheridito and Sagredo, 2016a), we want to construct a saddle point corresponding to the root of $\mathbf{T}_L$. This saddle point construction

[3] The operator $\mathbf{T}_L$ is closely related to the subdifferential of the saddle function L as v equals the subgradient of $L(., d)$ at $c \in \mathcal{C}$ and w is the subgradient of $-L(c, .)$ at $d \in \mathcal{D}$.

relies on an important mathematical property called maximal monotonicity.[4] The operator $\mathbf{T}_L$ associated with a saddle function L possesses this property if the following corollary is satisfied.

COROLLARY 4 (Rockafellar (1970)) *Let $\mathcal{C}$ and $\mathcal{D}$ be Hilbert spaces over $\mathbb{R}$. If $L(c,d)$ is a proper saddle function on $\mathcal{C}\times\mathcal{D}$ which is lower semicontinuous in its convex element $c\in\mathcal{C}$ and upper semicontinuous in its concave element $d\in\mathcal{D}$, then the operator $\mathbf{T}_L$ associated with L is maximal monotone.*

The reason for the importance of this property is that the resolvent[5] of a maximal monotone operator $\mathbf{T}$ is firmly nonexpansive.[6] This fact is due to Minty (1962). It is well known that any firmly nonexpansive operator is equivalent to a mixture $(1/2)\mathbf{Id}+(1/2)\mathbf{R}$ of the identity operator $\mathbf{Id}$ and a nonexpansive operator $\mathbf{R}$ (see e.g. Bauschke and Combettes, 2011, Remark 4.24 (iii)). Weak convergence of the iteration of such a mixture to its fixed point is well established (see e.g. Zeidler, 1986a, Proposition 10.16). This procedure is also known as damped fixed-point iteration.

Iterating on the resolvent of a maximal monotone operator yields the proximal point algorithm. This algorithm hence results in the fixed point of the resolvent which is equivalent to a root of the operator $\mathbf{T}$ itself. Therefore, it leads to a recursive equilibrium when we consider the resolvent of the operator $\mathbf{T}_L$ associated with the saddle function of (7). To understand how such a resolvent is constructed, let us look at a simplified example first.

[4] Maximal monotonicity (see e.g. Phelps, 1997; Bauschke and Combettes, 2011): Let $\mathcal{E}$ be a Hilbert space. An operator $\mathbf{T}:\mathcal{E}\to\mathcal{E}$ is called a monotone operator if for any two elements of its graph $(e,f),(\tilde{e},\tilde{f})\in G(\mathbf{T})=\{(e,f)\in\mathcal{E}^2|f\in\mathbf{T}(e)\}$ it holds that $\langle e-\tilde{e},f-\tilde{f}\rangle\geq 0$. It is, additionally, called maximal monotone if any $(\tilde{e},\tilde{f})\in\mathcal{E}^2$ with $\langle e-\tilde{e},f-\tilde{f}\rangle\geq 0\,\forall\,(e,f)\in G(\mathbf{T})$ is necessarily also an element of the graph $(\tilde{e},\tilde{f})\in G(\mathbf{T})$.

[5] Resolvent (see e.g. Bauschke and Combettes, 2011): Let $\mathcal{E}$ be a Hilbert space. The resolvent of an operator $\mathbf{T}:\mathcal{E}\to\mathcal{E}$ is the operator $(\mathbf{Id}+\mathbf{T})^{-1}$ where $\mathbf{Id}$ is the identity operator.

[6] Nonexpansiveness (see e.g. Bauschke and Combettes, 2011): Let $\mathcal{E}$ be a Hilbert space. An operator $\mathbf{T}:\mathcal{E}\to\mathcal{E}$ is called nonexpansive if it is Lipschitz continuous with constant 1. It is called firmly nonexpansive if for all $e,\tilde{e}\in\mathcal{E}$ it holds that $\|\mathbf{T}(e)-\mathbf{T}(\tilde{e})\|^2\leq\langle e-\tilde{e},\mathbf{T}(e)-\mathbf{T}(\tilde{e})\rangle$.

EXAMPLE (Resolvent of a subdifferential) Let $\mathcal{E}$ be a Hilbert space. Consider a lower semicontinuous proper convex function $F : \mathcal{E} \to [-\infty, \infty]$. It is well known that its subdifferential ∂F is maximal monotone (see e.g. Bauschke and Combettes, 2011, Theorem 20.40). We are looking for a fixed point $e^* \in \mathcal{E}$ of the resolvent of F which can be computed by simple iteration with iteration count n

$$e_n \overset{n\to\infty}{\longrightarrow} e^* \text{ with } e_{n+1} = (\mathbf{Id} + \partial F)^{-1}(e_n).$$

The resolvent $(\mathbf{Id} + \partial F)^{-1}$ can be represented by

$$\begin{aligned} e_{n+1} = (\mathbf{Id} + \partial F)^{-1}(e_n) &\Leftrightarrow e_n = (\partial F + \mathbf{Id})(e_{n+1}) \\ &\Leftrightarrow 0 = (\partial F + \mathbf{Id})(e_{n+1}) - \mathbf{Id}(e_n) \\ &\Leftrightarrow e_{n+1} = \arg\min_{e\in\mathcal{E}} F(e) + \frac{1}{2}\|e - e_n\|^2. \end{aligned}$$

The latter is the update of the proximal point algorithm.[7]

This example shows that the proximal point algorithm in our case translates into an algorithm on augmented Lagrangians. To ensure convergence, a regularization term containing the previous iterate has to be added to the Lagrangian. We define the update of the proximal point algorithm for the Lagrangian of our agents in the growth model in the following.

3.1.1. *The Proximal Point Algorithm for the Growth Model*

We follow Rockafellar (1976b) for defining the proximal point algorithm's update. This algorithm iterates on the resolvent of the operator associated with the Lagrangian of (7). Hence, each iteration on the resolvent updates the agents' optimal choices for consumption h_1 and individual capital h_2 as well as the three Lagrange multipliers y_1 for the equality constraint and y_2 and y_3 for the inequality

[7] The proximal point update presented here is a simplified version. Rockafellar (1976a) proves convergence for a resolvent $(\mathbf{Id} + \lambda^n \mathbf{T})^{-1}$ where $\{\lambda^n\}_{n=1}^{\infty}$ is either constant and bounded away from zero or a series $0 < \lambda^n \nearrow \lambda^\infty \leq \infty$.

constraints of (7). Similarly to the simplified example in the previous section, the $[n+1]$-th iterate of the agent's optimal choices, i.e. h_1^{n+1} and h_2^{n+1}, is the minimizer of the Lagrangian which is augmented by terms featuring the n-th iterate. The augmented Lagrangian is a function $L^A : \prod_{i=1}^{5} L^2(\mathcal{Z}^{id} \times \mathbb{R}, \mathcal{B}(\mathcal{Z}^{id} \times \mathbb{R}), \mu) \to [-\infty, \infty]$ given by

$$
\begin{aligned}
(8) \qquad L^A\left(h_1, h_2, y_1, y_2, y_3; h^n\right) = & -u(h_1) \\
& - \sum_{z'' \in \mathcal{Z}} p^{z''|z'} \beta u\left(I\left(z'', K'\right) + [1-\rho]\, h_2 - h_2 \circ h_2\right) \\
& + \frac{1}{2\lambda}\|h_1 - h_1^n\|^2 + \frac{1}{2\lambda}\|h_2 - h_2^n\|^2 \\
& + y_1\left(I\left(z', K\right) + [1-\rho]\, k - h_1 - h_2\right) \\
& + \frac{\lambda}{2}\left(I\left(z', K\right) + [1-\rho]\, k - h_1 - h_2\right)^2 \\
& + \begin{cases} -y_2 h_1 + \frac{\lambda}{2}(h_1)^2 & , h_1 \leq \frac{y_2}{\lambda} \\ -\frac{1}{2\lambda}(y_2)^2 & , h_1 > \frac{y_2}{\lambda} \end{cases} \\
& + \begin{cases} -y_3 h_2 + \frac{\lambda}{2}(h_2)^2 & , h_2 \leq \frac{y_3}{\lambda} \\ -\frac{1}{2\lambda}(y_3)^2 & , h_2 > \frac{y_3}{\lambda} \end{cases},
\end{aligned}
$$

where $h^n = (h_1^n, h_2^n)$ and $\lambda > 0$ is the step size parameter of the proximal point algorithm. The first two lines of the Lagrangian features the objective of (7) whereas the fourth line contains its equality constraint with its Lagrange multiplier. The third and fifth line consist of the objective and the equality constraint's proximal point augmentations which transform the saddle-point operator into its resolvent. The last two lines correspond to the inequality constraints. They also consist of the Lagrange term and the augmentation but they are defined piecewise to account for the case of a binding inequality.

With the augmented Lagrangian as above, we now state the algorithm to approximate a recursive equilibrium of the growth model in Algorithm 1. Note that Rockafellar (1976a) shows that the proximal point algorithm converges to a sad-

Algorithm 1 Proximal point algorithm for the growth model

▷ *A Initialization*

Set $n = 0$. Initialize the agents' choices of consumption and individual capital and the Lagrange multipliers $H^n = (h_1^n, h_2^n, y_1^n, y_2^n, y_3^n)$.

Set the parameter $\lambda > 0$.

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

▷ *B Iterative procedure*

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

Update H^{n+1} by

$$
\begin{aligned}
h^{n+1} &\approx \arg\min_{h_1,h_2} L^A\left(h_1, h_2, y_1^n, y_2^n, y_3^n; h^n\right) \\
y_1^{n+1} &= y_1^n + \lambda\left(I\left(z', K\right) + \left[1-\rho\right]k - h_1^{n+1} - h_2^{n+1}\right) \\
y_2^{n+1} &= \max\left(0, y_2^n - \lambda h_1^{n+1}\right) \\
y_3^{n+1} &= \max\left(0, y_3^n - \lambda h_2^{n+1}\right)
\end{aligned}
$$

where L^A is defined as in (8).

Compute the distance $d = \|H^{n+1} - H^n\|$.

Set $n = n + 1$.

end while

dle point of the Lagrangian even if the update of the optimal consumption and individual capital is only approximate. This is important as the minimizer of the Lagrangian is often not known in closed form, but it can be approximated with standard nonlinear solvers. Salzo and Villa (2012) extend this result to different concepts of approximation. Let me define which kind of approximation applies in this work.

Definition 5 (Resolvent approximation[8]) *Let $\mathcal{C}$ and $\mathcal{D}$ be Hilbert spaces over $\mathbb{R}$. Consider the resolvent $(\mathbf{Id} + \lambda \mathbf{T}_L)^{-1}(c, d)$ of an operator $\lambda \mathbf{T}_L$ associated with a saddle function L at $(c, d) \in \mathcal{C} \times \mathcal{D}$ with $\lambda > 0$. The approximation with ϵ-precision of this resolvent at $(c, d) \in \mathcal{C} \times \mathcal{D}$ is defined as $(\tilde{c}, \tilde{d}) \in \left(\mathbf{Id} + \lambda \mathbf{T}_L^{\epsilon^2/(2\lambda)}\right)^{-1}(c, d)$*

[8] This definition corresponds to the type 2 approximation with ϵ-precision in Salzo and Villa (2012).

where

$$\begin{aligned}\mathbf{T}_L^{\epsilon^2/(2\lambda)}(c,d) = \{(v,w)|L(\tilde{c},d) - \langle \tilde{c},v\rangle + \langle d,w\rangle \\ \geq L(c,d) - \langle c,v\rangle + \langle d,w\rangle - \frac{\epsilon^2}{2\lambda} \\ \geq L(c,\tilde{d}) - \langle c,v\rangle + \langle \tilde{d},w\rangle \,\forall(\tilde{c},\tilde{d}) \in \mathcal{C}\times\mathcal{D}\Big\}.\end{aligned}$$

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

The convergence rate of Algorithm 1 is $O(n^{-1})$ as is shown by Güler (1991). The 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). In the following, I explain the acceleration.

3.1.2. *The Accelerated Proximal Point Algorithm for the Growth Model*

The idea behind the acceleration is to approximate the highly nonlinear augmented Lagrangian with a sequence of simple convex quadratic functions $\{\phi^n\}_{n=1}^{\infty}$ such that the difference to the Lagrangian is reduced by a fraction $(1-\alpha^n) \in (0,1]$ in every iteration step

$$\phi^{n+1} - L^A \leq (1-\alpha^n)(\phi^n - L^A).$$

The update for the agents' optimal choices $h = (h_1, h_2)$ is then determined such that the following condition is satisfied

$$L^A\left(h^{n+1}, y_1^n, y_2^n, y_3^n; h^n\right) \leq \hat{\phi}^{n+1} = \min_h \phi^{n+1}(h),$$

where ϕ^{n+1} is of the form $\phi^{n+1}(h) = \hat{\phi}^{n+1} + (A^{n+1}/2)\|h - \nu^{n+1}\|^2$.

Salzo and Villa (2012) show that this is achieved by Algorithm 2. Furthermore, they show that this algorithm has a convergence rate of $O(n^{-2})$ if the resolvent approximation precision increases by $\epsilon^n = O(1/n^q)$ with $q > 3/2$.

Algorithm 2 Accelerated proximal point algorithm for the growth model

▷ *A Initialization*

Set $n = 0$. Initialize the agents' choices of consumption and individual capital and the Lagrange multipliers $H^n = (h_1^n, h_2^n, y_1^n, y_2^n, y_3^n)$. Set $\nu^n = h^n = (h_1^n, h_2^n)$.

Set the parameters $\lambda > 0$, $A^n > 0$ and $b \in [0, 2)$.

Set the resolvent approximation precision $\{\epsilon^n\}_{n=0}^{\infty}$.

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

▷ *B Iterative procedure*

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

Update $\alpha^n = \frac{1}{2}\left(\sqrt{(b\lambda A^n)^2 + 4b\lambda A^n} - b\lambda A^n\right)$.

Update $x^n = (1 - \alpha^n)h^n + \alpha^n \nu^n$.

Update H^{n+1} by

$$
\begin{aligned}
h^{n+1} &\approx \arg\min_{h_1, h_2} L^A\left(h_1, h_2, y_1^n, y_2^n, y_3^n; x^n\right) \\
y_1^{n+1} &= y_1^n + \lambda\left(I\left(z', K\right) + \left[1 - \rho\right]k - h_1^{n+1} - h_2^{n+1}\right) \\
y_2^{n+1} &= \max\left(0, y_2^n - \lambda h_1^{n+1}\right) \\
y_3^{n+1} &= \max\left(0, y_3^n - \lambda h_2^{n+1}\right)
\end{aligned}
$$

where L^A is defined as in (8).

Update $A^{n+1} = (1 - \alpha^n)A^n$.

Update $\nu^{n+1} = \nu^n - \frac{\alpha^n}{(1-\alpha^n)\lambda A^n}(x^n - h^{n+1})$.

Compute the distance $d = \|H^{n+1} - H^n\|$.

Set $n = n + 1$.

end while

3.2. *Discretizing the Space of Distributions*

So far, I just introduced the standard methodology of the proximal point algorithm. However, our model demands an extension. The recursive equilibrium, we want to solve for, depends on the cross-sectional distribution which is an element of the state space. Therefore, we need to discretize the space of distributions. If we simply use a spline interpolation on the distribution, the discretized state space becomes very large very quickly. Another option would be a projection on orthogonal polynomials but a prerequisite is a smooth distribution. Due to the borrowing constraint which is occasionally binding, however, the cross-sectional distribution exhibits mass points at the constraint and elsewhere as is proven in the following.

PROPOSITION 6 (A condition for mass points[9]) *Consider a recursive equilibrium as in Definition 1 with an explicit debt constraint $k \geq \delta$ with $\delta \in \mathbb{R}$. Suppose that there exists a $\hat{z} \in \mathcal{Z}$ with $p^{\hat{z}|\hat{z}} > 0$ and a $\hat{k} > \delta$ such that $h_2(\hat{z}, k, \mu) \leq k$ for all $k \in [\delta, \hat{k}]$. Furthermore, assume that the optimal capital savings function has a kink at $k^* := \max\{k \geq \delta \,|\, h_2(\hat{z}, k, \mu) = \delta\} > \delta$, i.e. the debt constraint is binding, and that h_2 is strictly increasing in $k \geq k^*$. Then, the cross-sectional distribution has a mass point at the constraint δ. If, additionally, there exists a $\bar{z} \in \mathcal{Z}$ with $p^{\bar{z}|\hat{z}} > 0$ and $h_2(\bar{z}, \delta, \mu) > \delta$, then the cross-sectional distribution has multiple mass points.*

This result implies jumps in the cross-sectional distribution μ. Hence, standard orthogonal polynomial projection methods do not work here. There is, however, an efficient way of approximating distributions called polynomial chaos. This is a technique which projects the distribution on orthogonal polynomials of random variables rather than the real line. As such, it can also handle discontinuous distributions. In the following, I will summarize polynomial chaos in general and subsequently, I this technique is used for discrete distributions.

3.2.1. *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. uni-

[9]The proof can be found in Appendix A.

form, 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 $\kappa \in L^2$ with distribution $\kappa \sim \mu$ by

$$\kappa = \sum_{i=0}^{\infty} \varphi_i \Phi_i(\xi), \tag{9}$$

where φ_i are constant projection coefficients.

It is important to note that there is a specific 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

$$\langle \Phi_i, \Phi_j \rangle = \int_{-\infty}^{\infty} \Phi_i(\xi)\, \Phi_j(\xi)\, dF(\xi) = \frac{\delta_{ij}}{a_i^2}, \tag{10}$$

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)

$$\Phi_{i+1}(\xi) = (\xi - \theta_i)\, \Phi_i(\xi) - \omega_i \Phi_{i-1}(\xi), \; i \in \{0, 1, \ldots\}, \tag{11}$$

where the starting polynomials are defined as $\Phi_{-1}(\xi) = 0$ and $\Phi_0(\xi) = 1$ and the constant parameters are given by $\theta_i, \omega_i \in \mathbb{R}$ with $\omega_i > 0$.

The projection coefficients in the polynomial chaos expansion of a random variable $\kappa \in L^2$ with distribution $\kappa \sim \mu$ are defined as usual by $\varphi_i = \langle \kappa, \Phi_i \rangle / \langle \Phi_i, \Phi_i \rangle$ for all $i \in \{0, 1, \ldots\}$. If κ is not a direct function of the basic random variable ξ, one uses the fact that both $\mu, F \sim \mathcal{U}[0,1]$ are uniform to compute the coefficients

$$\varphi_i = \frac{1}{\langle \Phi_i, \Phi_i \rangle} \int_{-\infty}^{\infty} \mu^{-1}\left(F\left(\xi\right)\right) \Phi_i\left(\xi\right) dF\left(\xi\right) \; \forall\, i \in \{0, 1, \ldots\}, \tag{12}$$

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

3.2.2. *Applying Polynomial Chaos to our Growth Model*

We apply the polynomial chaos expansion to discretize the cross-sectional distribution of our growth model. Let us first condition the distribution on the employment status, i.e. $\mu^0(k) = \mu(0,k)/p^{z^{id}=0|z^{ag}}$ denotes the cross-sectional distribution of the unemployed. Accordingly, we denote the basic random variable, the projection coefficients and the polynomials of the unemployed with superscript 0 and of the employed with superscript 1, respectively. The optimal consumption and capital savings choices can then be rewritten as $c' = h_1(z', k, \{\varphi_i^0\}_{i=0}^{\infty}, \{\varphi_i^1\}_{i=0}^{\infty})$ and $k' = h_2(z', k, \{\varphi_i^0\}_{i=0}^{\infty}, \{\varphi_i^1\}_{i=0}^{\infty})$. Similarly ,we can derive the law of motion of the projection coefficients from (12), (9) and (6). W.l.o.g. we will write down formulas only w.r.t. the distribution of the unemployed in the following. The law of motion for the projection coefficients of the unemployed reads

$$\varphi_i^{0\prime} = \frac{1}{\langle \Phi_i^0, \Phi_i^0 \rangle} \int_{-\infty}^{\infty} \left[\mu^{0\prime}\right]^{-1} \left(F^0\left(\xi^0\right)\right) \Phi_i^0\left(\xi^0\right) dF^0(\xi^0), \; \forall\, i \in \{0, 1, \ldots\}$$

with

$$\mu^{0\prime}(k) = \frac{1}{p^{z^{id\prime}=0|z^{ag\prime}}} \sum_{j=0}^{1} \frac{p^{(z^{ag\prime},0)|(z^{ag},j)}}{p^{z^{ag\prime}|z^{ag}}} \mathbb{P}\left(\{\zeta = j\} \cap \right.$$
$$\left. \left\{ h_2\left(z^{ag\prime}, 0, \sum_{i=0}^{\infty} \varphi_i^j \Phi_i^j\left(\xi^j\right), \{\varphi_i^0\}_{i=0}^{\infty}, \{\varphi_i^1\}_{i=0}^{\infty}\right) \leq k \right\} \middle| z^{ag\prime}, z^{ag} \right).$$

3.2.3. *A Specific Choice of the Basic Random Variables*

It was illustrated in Xiu and Karniadakis (2002) that the speed of convergence significantly improves if the distribution of the basic random variable is not too far from the distribution we want to approximate. Since the cross-sectional distribution in our growth model is an endogenous object, we do not know its shape a priori. We do know, however, that it will have mass points according to Proposition 6. Also, we know that the same growth model without aggregate shocks, i.e. where z^{ag} is fixed at either 0 or 1, has an endogenous cross-sectional distribution which stays constant as time goes on. This case is easy to compute because $K' = K$ in the agents' optimization problem (7). Therefore, one just has to solve the individual optimization problem at different values of aggregate capital K. In a second step, given these optimal responses, one can compute the stationary distribution as a fixed point of the distribution's law of motion (6). Naturally, this cross-sectional distribution will have features similar to the distribution of the model with aggregate shocks. Hence, we fix ξ^0 as the cross-sectional distribution of the unemployed in the model without aggregate shocks averaged over the two cases of keeping z^{ag} fixed as 0 or 1. Analogously, ξ^0 is fixed as the distribution of the employed. I compute these stationary distributions using the histogram approach by Young (2010). They are displayed in Figure 2. The distributions exhibit several mass points measured as the local extrema in the histogram representation. One can observe that the distribution of the unemployed is shifted to the left compared to the distribution of the employed as they generally have lower capital savings. Furthermore, the capital constraint is binding solely for a fraction of

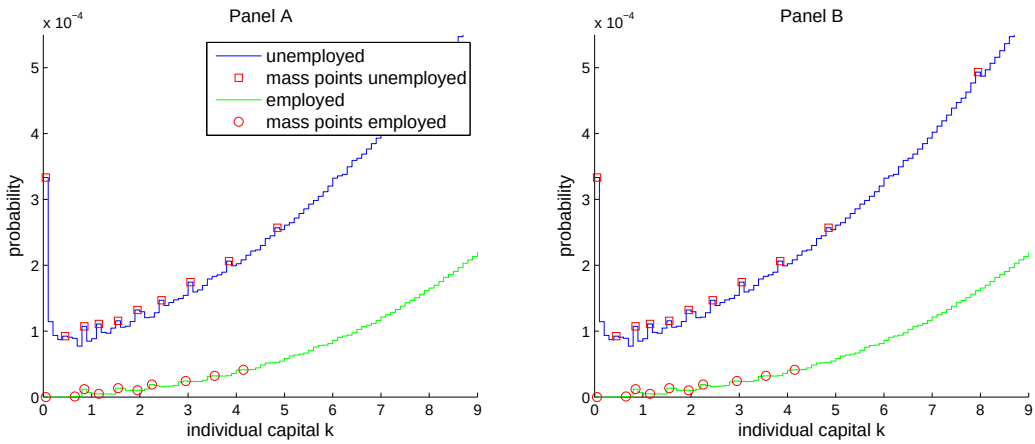

FIGURE 2. **Stationary cross-sectional capital distribution for the growth model without aggregate shocks.** Panel A shows the histogram representation with bin size 0.1 of the stationary distribution conditional on the employment status of the agents. A mass point is identified as a bin whose probability is higher than the ones of its direct neighbors, but the global maxima are excluded. Panel B zooms into the left tail of the distribution.

the unemployed agents. To obtain an accurate approximation of these stationary distributions, one should choose a reasonably small bin size for the histogram.

3.2.4. *Generation of the Corresponding Orthogonal Polynomials*

As our basic random variables are represented by histograms, they are essentially discrete distributions where the midpoints of the bins $\{\xi_n^0\}_{n=1}^N$ have probability $\{p_n^0\}_{n=1}^N$. Generally, orthogonal polynomials w.r.t. a discrete distribution with finite support are considered discrete as well in the sense that their maximal degree is N. Furthermore, the highest-order polynomial Φ_N has the midpoints $\{\xi_n^0\}_{n=1}^N$ as roots. In Zheng et al. (2015), different methods for generating polynomials corresponding to discrete distributions are compared. Of their suggested methods, we use the Stieltjes method which performs well in terms of precision. It directly computes the parameters θ_i and ω_i in (11) using the standard inner product of L^2 and is

explained in detail in Gautschi (1982). The constant parameters are given by

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

with $\langle ., . \rangle$ denoting the standard inner product of L^2 w.r.t. the corresponding basic distribution. As these distributions are represented as discrete distributions, the inner product is a sum rather than an integral. The definitions of these parameters follow from inserting the three-term recurrence relation (11) into the orthogonality condition (10). With the parameters defined as above, the orthogonal polynomials are easily constructed using (11). They are displayed in Figure 3. As usual, the

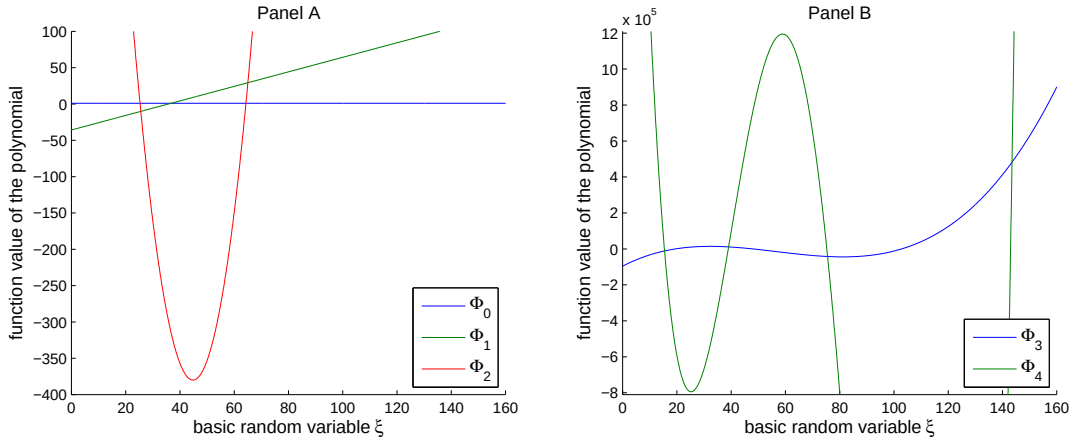

FIGURE 3. **Orthogonal polynomials corresponding to the stationary distribution for the model without aggregate shocks.** Panel A shows the polynomials up to order 2 and Panel B displays the polynomials of order 3 and 4.

number of roots of each polynomial corresponds to its degree. Note that the first-order polynomial has its root at the mean of the distribution F^0 of the basic random variable.

With the basic random variables defined and the corresponding polynomials generated, the polynomial chaos expansion is fully defined up to order N. 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 4. In this figure, we consider a polynomial

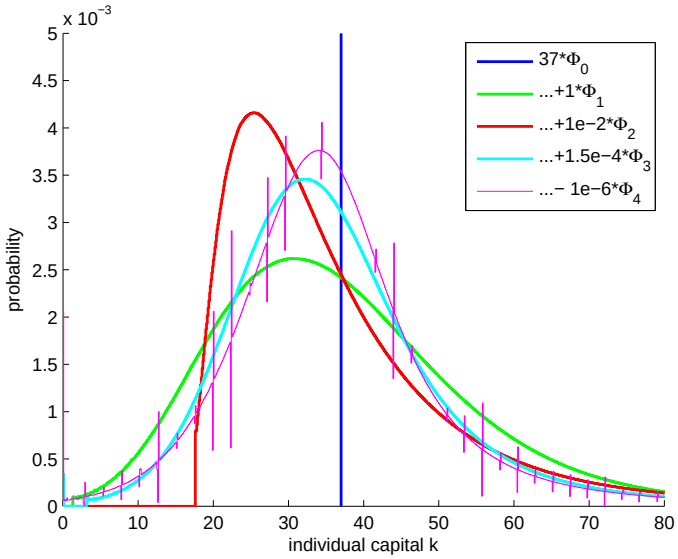

FIGURE 4. **Example distributions resulting from truncated polynomial chaos expansions.** The graph displays the histogram representations with bin size 0.1 of distributions which result from the polynomial chaos series truncated at different orders ranging from order 0 to 4. The basic random variable used is the stationary cross-sectional distribution of the unemployed in the growth model without aggregate shocks as in Figure 2. The projection coefficients for this example are fixed as $[\varphi_0, \ldots, \varphi_4] = [38, 1, 0.01, 0.0002, -0.000001]$.

chaos expansion with fixed projection coefficients $\{\varphi_i\}_{i=0}^{\infty}$ which is truncated at different orders. The zeroth-order polynomial results in a mass point at φ_0 which due to its definition is the mean of the projected distribution. Adding the second-order polynomial simply shifts the distribution of the basic random variable to the left or right depending on the projection coefficients. Adding the third-order polynomial modifies the skewness of the basic random variable whereas the fourth-order polynomial adjusts the kurtosis. In the example projection, the fifth-order does not seem to modify the tails much further but rather shapes the distribution in the high-probability region.

4. CONVERGENCE RESULTS

With the methodology laid out in the previous section, we now show convergence of the proposed accelerated proximal point algorithm which uses polyno-

mial chaos to discretize the cross-sectional distribution. This is done in two steps. Firstly, the convergence for the theoretical resolvent, i.e. the iteration on the policy functions $c' = h_1(z', k, \{\varphi_i^0\}_{i=0}^{\infty}, \{\varphi_i^1\}_{i=0}^{\infty})$ and $k' = h_2(z', k, \{\varphi_i^0\}_{i=0}^{\infty}, \{\varphi_i^1\}_{i=0}^{\infty})$ is shown. In the second step, we show that also the discretized resolvent, i.e. the iteration on the discretized policy functions $c' = \hat{h}_1(z', k, \{\varphi_i^0\}_{i=0}^{M}, \{\varphi_i^1\}_{i=0}^{M})$ and $k' = \hat{h}_2(z', k, \{\varphi_i^0\}_{i=0}^{M}, \{\varphi_i^1\}_{i=0}^{M})$, converges. The hat denotes the finite-element discretization w.r.t. the second argument k, whereas $M \leq N$ denotes the order at which the polynomial chaos is truncated.

4.1. *Convergence of the Theoretical Solution Operator to the Recursive Equilibrium*

According to the theory on the proximal point algorithm (i.e. Rockafellar, 1970, 1976a,b; Güler, 1992; Salzo and Villa, 2012) summarized in Section 3.1, it suffices to show that the Lagrangian of the agents' optimization problem (7) satisfies the conditions of Corollary 4. This guarantees the nonexpansiveness of the corresponding resolvent and hence, convergence of the proximal point algorithm.

THEOREM 7 (Convergence) *Consider the growth model from Section 2. Consider the function space $\mathcal{H}$ defined in Proposition 8 for the consumption and capital savings choice. The Lagrangian $L : \mathcal{H} \times \prod_{i=1}^{3} L^2(\mathcal{Z}^{id} \times \mathbb{R}, \mathcal{B}(\mathcal{Z}^{id} \times \mathbb{R}), \mu) \to [-\infty, \infty]$ of the agents' optimization problem (7) in the growth model given by*

$$
\begin{aligned}
(13) \qquad L\left(h_1, h_2, y_1, y_2, y_3\right) = & -u(h_1) \\
& - \sum_{z'' \in \mathcal{Z}} p^{z''|z'} \beta u\left(I\left(z'', K'\right) + \left[1-\rho\right] h_2 - h_2 \circ h_2\right) \\
& + y_1\left(I\left(z', K\right) + \left[1-\rho\right] k - h_1 - h_2\right) - y_2 h_1 - y_3 h_2
\end{aligned}
$$

satisfies the conditions of Corollary 4 and therefore, Algorithm 1 and Algorithm 2 converge to a recursive equilibrium of the growth model.

5. COMPARISON TO THE KRUSELL-SMITH ALGORITHM

We compute the recursive equilibrium solution of Algorithm 2 using Matlab.[10] As an example, we analyze the worst-case scenario for 25 periods, i.e. we set the 25 aggregate shocks to the bad economic state. As we consider months as the time period in our calibration of the model, this corresponds to a two year long downturn after initial time. Furthermore, we run the Krusell-Smith algorithm implemented by Maliar et al. (2010) in Matlab for the same grid on individual capital and the same policy function initialization to obtain a Krusell-Smith equilibrium solution. Since the Krusell-Smith does not explicitly compute the stationary distribution, we solve for it by finding a fixed point of the law-of-motion operator

$$\left(\mathbf{D}^{h_{KS}}\mu_{KS}^{k|z}\right)(v|u) = \sum_{z\in\mathcal{Z}} p^{*z}\frac{p^{z'=u|z}}{p^{*z'=u}}\mu_{KS}^{k|z}\left(\tilde{\mathcal{K}}_u^v\right)$$

where

$$\tilde{\mathcal{K}}_u^v = \left\{ \left[0, \max_{\tilde{k}\in\mathbb{R}_{\geq 0}} \tilde{k}\right] \,\middle|\, h_{KS}\left(u, \tilde{k}, \mathbb{E}^{\mu_{KS}^k}(k)\right) = v \right\}.$$

We then use the corresponding law of motion to compute the distributions and their first moments produced by the Krusell-Smith policy at every time point in the worst-case scenario.

One way of comparing these two sets of numerical solutions is to analyze their errors. There have been two error tests put forward in 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 $c(z';k;\mu^k)$ against the explicitly calculated conditional expectation in the Euler equation denoted by

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

$\tilde{c}(z';k,\mu^k)$. It is the percentage error

$$\epsilon^{SEE}(z';k,\mu^k) = \frac{c(z';k;\mu^k) - \tilde{c}(z';k,\mu^k)}{\tilde{c}(z';k,\mu^k)}.$$

In contrast to the standard Euler equation error, the dynamic equivalent denoted by ϵ^{DEE} 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.

Table 1 compares the Euler equation errors for our numerical solution and the Krusell-Smith solution. We see that when we consider the grid on which the algo-

	$\mathbb{E}\left(\lvert\epsilon^{SEE}\rvert\right)$	$\max\left\{\lvert\epsilon^{SEE}\rvert\right\}$	$\mathbb{E}\left(\lvert\epsilon^{DEE}\rvert\right)$	$\max\left\{\lvert\epsilon^{DEE}\rvert\right\}$
		Algorithm Grid		
PPA with gPC	2.4741e-09	5.9722e-08	1.4035e-05	2.0232e-04
Krusell-Smith	1.7323e-02	2.6728e-01	1.7323e-02	2.6728e-01
		Finer Grid		
PPA with gPC	5.1879e-05	1.4417e-01	4.9455e-05	1.4418e-01
Krusell-Smith	3.5920e-02	3.5823e-01	3.5924e-02	3.5824e-01

TABLE 1. Average and maximum standard (SEE) and dynamic (DEE) Euler equation errors in % for the numerical solution of the growth model's recursive equilibrium from the proximal point algorithm with polynomial chaos and the Krusell-Smith solution. The first panel shows the errors w.r.t. the grid which has been used in the algorithms to obtain the numerical solutions. The second panel shows the errors w.r.t. a finer grid than the algorithm grid and therefore incorporates the approximation error.

rithms were computed, the average standard Euler equation errors for our solution algorithm are in the region of the termination criterion $1e-8$. This is expected because the algorithm is designed to correctly match the Euler equation for all time steps by keeping track of the exact law of motion. This holds even for the more stringent dynamic test. The Krusell-Smith algorithm, however, exhibits a considerably larger error for both tests which comes from the parametric approximation of the law of motion whose estimation relies on an error-prone simulation. The picture is not as clear when we compute the errors w.r.t. a finer grid. These

errors include the actual approximation failure. The errors for our solution algorithm worsen considerably. However, when we consider the whole error distributions, the outperformance of our solution algorithm is confirmed. Let us look at the whole error distributions in form of boxplots in Figure 5. The boxplots show

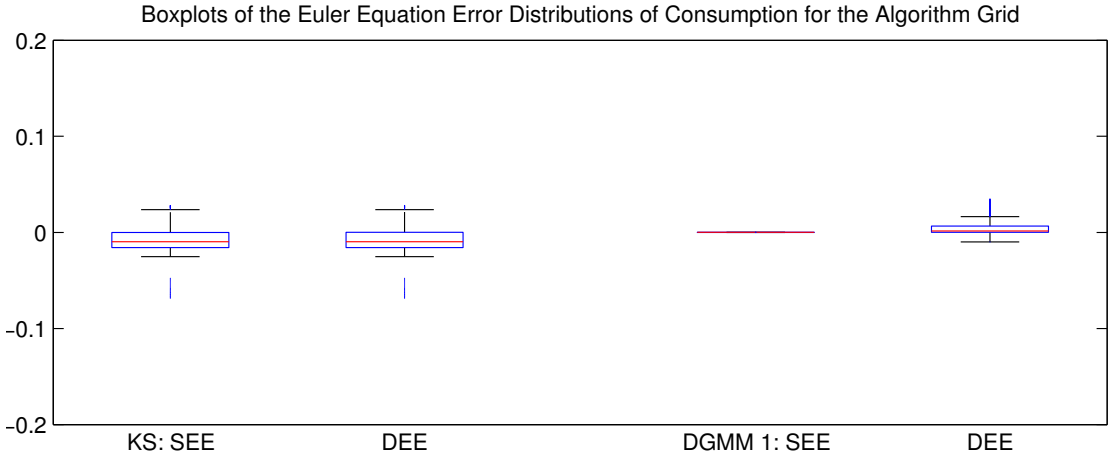

FIGURE 5. Boxplots of the distributions of the standard (SEE) and dynamic (DEE) Euler equation errors for the numerical solution of the growth model's recursive equilibrium from the proximal point algorithm with polynomial chaos and the Krusell-Smith solution. The central marker corresponds to the median whereas the edges of each box indicate the 25^{th} and 75^{th} percentile. The whiskers indicate the extreme values not considered as outliers. Outliers are shown as dots outside the whiskers, they are the data points outside the 25^{th} and 75^{th} percentile ± 1.5 times the difference of these percentiles.

very clearly that the error distributions of our solution algorithm are clustered very closely around zero except for outliers whereas the error distributions of the Krusell-Smith algorithm are significantly more dispersed. The analysis of the Euler equation errors indicates that our solution algorithm is superior in terms of a precise computation of the law of motion of individual capital.

Let us now analyze which benefits the abandonment of the assumption of bounded rationality has. Theoretically, our solution should be closer to an exact equilibrium in terms of the agent's utility, i.e. it should result in higher utility at all dates and states. Let us compare the utility summed over all dates of our test

scenario and averaged over the cross-sectional distribution of optimal consumption

$$U = \sum_{t=0}^{24} \beta^t \mathbb{E}^{\mu^c} \left(\frac{c^{1-\gamma} - 1}{1-\gamma} \right).$$

The aggregate utility is given in Table 2. We see that it is marginally higher for

U^{DGMM}	U^{KS}	K_0
32.1461	31.9289	23.0276

TABLE 2. Aggregate utility for the worst-case scenario of 25 bad aggregate shocks in the growth model computed by the numerical solution of the recursive equilibrium from the proximal point algorithm with polynomial chaos compared to the aggregate utility given by the Krusell-Smith policy at the same beginning-of-period aggregate capital K_t as the former two.

our solution algorithm for any order of Taylor approximation when comparing to the Krusell-Smith algorithm. Also, the utility is higher for the first-order Taylor approximation compared to the second-order approximation. The reason for that is that the conditional stationary distributions for these two solutions differ which can be seen in the stationary initial beginning-of-period capital K_0.

Let us compare these two stationary distributions in detail. We compute the stationary conditional distribution for the Krusell-Smith policy by solving for a fixed point of the law-of-motion operator. The stationary conditional distributions $\mu^{k|z}$ are compared in Figure 6. The differences in distributions are especially visible in the right tail. The Krusell-Smith policy will lead to higher capital than the recursive equilibrium for the richest agents in the economy. Furthermore, the capital of the richest agents shrinks with higher order of approximation of the recursive equilibrium.

These significant differences in the cross-sectional distributions, in particular in the aggregate capital, imply that we cannot compare utility in nominal terms. We have to look at it in relative terms. I suggest to consider the aggregate consumption

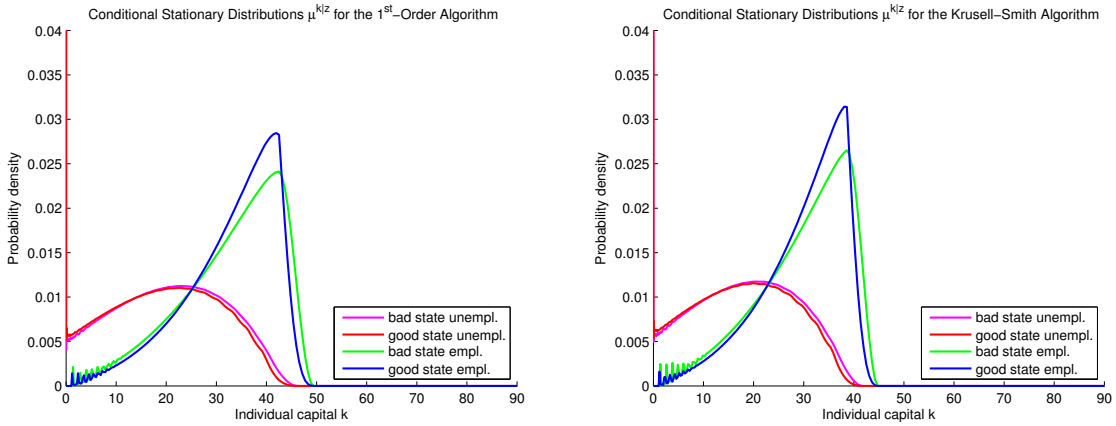

FIGURE 6. The stationary conditional distributions of the growth model $\mu^{k|z}$ by the recursive equilibrium from the proximal point algorithm with polynomial chaos and the Krusell-Smith solution.

to aggregate wealth ratio

$$\frac{C_t}{\Omega_{t+1}} = \frac{C_t}{p^e_{t+1} W\left(z^{ag}_{t+1}; \mu^k_t\right) + \left[1 + R\left(z^{ag}_{t+1}; \mu^k_t\right) - \delta\right] K_t},$$

where aggregate wealth Ω results from the aggregated budget constraint. The ratio is displayed in Table 3. We see that the aggregate consumption-to-wealth

	initial ratio	after 1 year	after 2 years
PPA with gPC	16.1600	16.2437	16.3126
Krusell-Smith	15.0945	15.1594	15.2147

TABLE 3. Aggregate consumption-to-wealth ratio (in %) at $t = 0, 12, 24$ for the worst-case scenario of 25 bad aggregate shocks in the growth model computed by the numerical solution of the recursive equilibrium from the proximal point algorithm with polynomial chaos and the Krusell-Smith solution.

ratio is higher for our solution algorithm. The reason for the higher portion of wealth being consumed by the agents in our algorithm lies in the abandonment of the assumption of bounded rationality. By letting the response depend on the whole distribution, the agents can evaluate their position w.r.t. the tails of the distribution. It leads to higher curvature in the policy for medium-sized individual capital, i.e. these agents consume more. This leads to a decrease in aggregate capital for higher relative consumption.

Alternatives to the Krusell-Smith algorithm were developed in Algan et al. (2008, 2010); den Haan and Rendahl (2010); Reiter (2010a) and Young (2010). They avoid the simulation across agents in using different techniques to approximate the distribution over a continuum of agents. Most of them use a parametric law of motion of aggregate capital using higher-order moments of the cross-sectional distribution rather than the correct law of motion. Even though this is certainly an improvement regarding the approximation of the law of motion, the assumption of bounded rationality remains. More importantly, these algorithms still rely on component-wise fixed-point iteration. They use a finite number of simulation steps for the aggregate shock to compute the evolution of the stationary capital distribution from which they then derive the law of motion.

In contrast, Reiter (2009, 2010b) does not assume bounded rationality. This algorithm deals with aggregate shocks by approximating the model without aggregate shocks first and then perturbing the solution around the volatility of the aggregate shock. In the first step, they also use a component-wise iterative procedure. Without aggregate shocks, this procedure does not cause problems because the equations for the policy and the distribution are sufficiently decoupled. However, due to the linear perturbation in the aggregate risk dimension, this algorithm cannot handle large and nonlinear aggregate shocks.

6. CONCLUSIONS

In this paper, I develop a novel solution algorithm to solve a wide group of DSGE models with ex-post heterogeneity and aggregate risk. There are two major differences to the existing algorithms, most prominently the Krusell and Smith (1998) algorithm. Firstly, the algorithm herein does not require bounded rationality of the agents and hence, it does not rely on an additional model assumption. Instead I solve for the original recursive equilibrium where agents observe the full cross-sectional distribution. Because I do not abstract from the cross-sectional distribution, the whole state distribution is an integral part of my algorithm. I pa-

rameterize it using polynomial chaos expansions which projects the distribution onto an orthogonal basis spanning the whole space of square-integrable random variables. Economically, this leads to more curvature in the optimal policy and thus higher levels of consumption relative to wealth, especially at the core of the wealth distribution.

Another advantage of the proposed algorithm is that it is designed using the theory of monotone operators and proximal points such that a theoretical convergence result carries over. This ensures that any solution resulting from the algorithm is indeed close to a recursive rational equilibrium.

Overall, my approach provides a new tool to analyze numerical solutions of DSGE models for which convergence is ensured. It provides insights into the tails of the cross-sectional distribution after large or persistent aggregate shocks which is important for risk analysis. This is interesting for instance for macro-finance models which investigate systemic risk in financial markets and its effect on the real economy.

APPENDIX A: PROOFS

A.1. *Proof of Proposition 6*

PROOF: Let us denote the support of the marginal distribution w.r.t. k of the cross-sectional distribution by $\operatorname{supp} \mu^k$. The minimum value of k which has positive probability is denoted by $\underline{k} = \min_k \operatorname{supp} \mu^k$. First, let us show that the constraint has positive probability $\delta \in \operatorname{supp} \mu^k$. Because of $p^{\hat{z}|\hat{z}} > 0$, eventually we have $\hat{z}^{ag}$ in the previous and the current period. Suppose that the start capital at which the constraint starts binding is not in the support $k^* < \underline{k} \leq \hat{k}$. Applying the optimal capital savings function, we obtain that $\underline{k}' = h_2(\hat{z}, \underline{k}, \mu) \leq \underline{k}$. By induction, this contradicts $k^* \notin \operatorname{supp} \mu^k$. Now let us show that there is a mas point at δ. Assume that $\delta < k^* = \underline{k}$. Because h_2 is continuous and strictly increasing to the right of its kink, there exists an interval $[k^*, \bar{k}]$ with $\bar{k} := \max\{k \geq \delta \,|\, h_2(\hat{z}, k, \mu) = k^*\} > k^*$

and positive measure $\mu^k([k^*, \bar{k}]) > 0$. Due to $p^{\hat{z}|\hat{z}} > 0$, a strictly positive part of this mass will stay at $\hat{z}$ and have future value δ. Hence, $\mu^k(\delta) > 0$ and $\underline{k} = \delta$. This yields the mass point at the constraint for the cross-sectional distribution. Using the same reasoning, one can easily see that this mass point at zero propagates to higher levels of individual capital at $\bar{z} \in \mathcal{Z}$. *Q.E.D.*

A.2. *Proof of Theorem 7*

In order to proof Theorem 7, we need the following preliminary result.

PROPOSITION 8 (Admissible Set of the Growth Model) *Consider the growth model from Section 2. Define a subspace $\mathcal{H}$ of the intersection of the square-integrable functions w.r.t. $\mu \in L^2$, i.e. $L^2\left(\mathcal{Z}^{id} \times \mathbb{R}, \mathcal{B}(\mathcal{Z}^{id} \times \mathbb{R}), \mu\right)$, and the functions with bounded first and second variation[11] such that for any element $h = (h_1, h_2) \in \mathcal{H}$, the following inequalities are satisfied almost surely for any $z, z' \in \mathcal{Z}$ and $(z^{id}, \kappa) \sim \mu$*

(*i*) *Nonnegative consumption:* $h_1(z', \kappa, \mu) \geq 0$

(*ii*) *Limited capital savings:* $h_2(z', \kappa, \mu) \leq \mathbf{I}(z', \kappa) + [1 - \rho]\kappa$

(*iii*) *Lower bound on the average second variation of capital savings:*

$$\sum_{z' \in \mathcal{Z}} \mathbf{P}(z', z, \kappa)\, \delta^2 h_2(z', \kappa, \mu; \tilde{\kappa}) \geq \sum_{z' \in \mathcal{Z}} \mathbf{P}(z', z, \kappa) \left[\delta^2 \mathbf{I}(z', \kappa; \tilde{\kappa}) - \gamma \frac{[\delta \mathbf{C}(z', \kappa; \tilde{\kappa})]^2}{\mathbf{C}(z', \kappa)}\right]$$

where $\mathbf{P}$ denotes the probability operator, $\mathbf{C}$ denotes the consumption operator and

[11] n^{th} variation (see e.g. Zeidler, 1986b): Let $\mathcal{E}$ be a Hilbert space. The n^{th} variation of an operator $\mathbf{T} : \mathcal{E} \to \mathcal{E}$ at a point $e \in \mathcal{E}$ in the direction $\tilde{e} \in \mathcal{E}$ is defined by $\delta^n \mathbf{T}(e; \tilde{e}) = d^n / dt^n \mathbf{T}(e + t\tilde{e})|_{t=0}$.

$\mathbf{I}$ *denotes the income operator defined as follows*

$$\begin{aligned}
\mathbf{P}\left(z', z, \kappa\right) &= \frac{p^{z'|z}\mathbf{C}(z', \kappa)^{-\gamma}}{\sum_{z' \in \mathcal{Z}} p^{z'|z}\mathbf{C}(z', \kappa)^{-\gamma}} \\
\mathbf{C}\left(z', \kappa\right) &= \mathbf{I}\left(z', \kappa\right) + \left[1 - \rho\right]\kappa - h_2\left(z', \kappa, \mu\right) \\
\mathbf{I}\left(z', \kappa\right) &= \alpha\tilde{a}(z^{ag'})\left[\frac{\langle \kappa, \mathbf{1}\rangle}{\pi p^{e'}}\right]^{\alpha-1}\kappa + [1-\alpha]\tilde{a}(z^{ag'})\tilde{\nu}\left(z'\right)\left[\frac{\langle \kappa, \mathbf{1}\rangle}{\pi p^{e'}}\right]^{\alpha}
\end{aligned}$$

with

$$\begin{aligned}
\tilde{a}(z^{ag'}) &= 1 + z^{ag'}a - [1 - z^{ag'}]a \\
\tilde{\nu}\left(z'\right) &= \nu + z^{id'}\left[1 - \frac{\nu}{p^{e'}}\right].
\end{aligned}$$

Then, $\mathcal{H}$ is a Hilbert space.

PROOF: It is well known that the subspace of functions with bounded variation within L^2 is complete and hence, a Hilbert space itself. With conditions $(i) - (iii)$, we take yet another subset of functions with bounded variations. It is easy to see that any limiting element h^* of a Cauchy sequence $h^n \in \mathcal{H}$, $n \in \{1, 2, \ldots\}$, satisfies conditions $(i) - (iii)$ as well. The subspace $\mathcal{H}$ is therefore complete and a Hilbert space itself. *Q.E.D.*

REMARK Note that condition (iii) implies that the expected second variation of capital savings in a nonnegative direction $\tilde{\kappa} \geq 0$ is greater equal a nonpositive threshold if $K \geq ([2-\alpha]\pi p^{e'})/([1-\alpha]\nu)$. In our calibration, this translates into $K \geq 18.2222$ which seems to be a reasonable minimum value for aggregate capital. Due to this fact, additionally to convex capital savings functions, the subspace also allows for capital savings functions with concave sections.

PROOF OF THEOREM 7: It suffices to show that the Lagrangian (13) of the agents' optimization problem (7) satisfies the conditions of Corollary 4. This guarantees the nonexpansiveness of the corresponding resolvent and therefore implies

convergence of the (accelerated) proximal point algorithm.

Saddle function: Let us start by specifying the Hilbert spaces $\mathcal{C} \times \mathcal{D}$ the Lagrangian is defined on. L depends on the optimal controls $h = (h_1, h_2)$. Hence, we define the first Hilbert space by $\mathcal{H}$ as given in Proposition 8. The Lagrange multipliers lie in the corresponding dual space which implies $\mathcal{D} \subseteq L^2(\mathcal{Z}^{id} \times \mathbb{R}, \mathcal{B}(\mathcal{Z}^{id} \times \mathbb{R}), \mu)$ such that $y_2, y_3 \geq 0$ for any exogenous shock and start capital.

Now it remains to show that the Lagrangian is convex in the optimal controls and concave in the Lagrange multipliers. The latter is trivial as the Lagrangian is linear in the multipliers. The former means that the Hessian (in terms of second variations) w.r.t. h_1 and h_2 is positive semidefinite (see e.g. Zeidler, 1986b, Corollary 42.8). As the cross-variation is zero and the second variation of the Lagrangian w.r.t. h_1 is nonnegative, we need to show that the second variation of the operator

$$\mathbf{U}(h_2) = \sum_{z'' \in \mathcal{Z}} p^{z''|z'} \beta u \left(I\left(z'', h_2\right) + \left[1 - \rho\right] h_2 - h_2 \circ h_2 \right)$$

is nonpositive. Defining $\mathbf{C}\left(z'', h_2\right) = \mathbf{I}\left(z'', h_2\right) + [1-\rho]h_2 - h_2 \circ h_2$, we have

$$\begin{aligned} \delta^2 \mathbf{U}\left(h_2; \tilde{h}\right) = \sum_{z'' \in \mathcal{Z}} p^{z''|z'} \beta \Big\{ & u_{cc}\left(\mathbf{C}\left(z'', h_2\right)\right) \left[\delta \mathbf{C}\left(z'', h_2; \tilde{h}\right)\right]^2 \\ & + u_c\left(\mathbf{C}\left(z'', h_2\right)\right) \delta^2 \mathbf{C}\left(z'', h_2; \tilde{h}\right) \Big\} \end{aligned}$$

and therefore convexity is achieved by a next-period consumption choice with second variation bounded from above

$$\sum_{z'' \in \mathcal{Z}} p^{z''|z'} \beta \mathbf{C}\left(z'', h_2\right)^{-\gamma} \delta^2 \mathbf{C}\left(z'', h_2; \tilde{h}\right) \leq \gamma \sum_{z'' \in \mathcal{Z}} p^{z''|z'} \beta \mathbf{C}\left(z'', h_2\right)^{-\gamma} \frac{\left[\delta \mathbf{C}\left(z'', h_2; \tilde{h}\right)\right]^2}{\mathbf{C}\left(z'', h_2\right)}.$$

This condition is equivalent to our condition (iii) of bounded second variation for $h \in \mathcal{H}$ and is therefore satisfied. Applying the resolvent corresponding to the Lagrangian amounts to finding the minimum of the augmented Lagrangian (8). Any

such minimum stays in $\mathcal{H}$ (see e.g. Bauschke and Combettes, 2011, Proposition 23.2) which ensures convexity throughout the algorithm.

Properness: We proceed in two steps. First we show that there exist Lagrange multipliers $y = (y_1, y_2, y_3) \in \mathcal{D}$ such that $L(h, y) > -\infty$ for all $h \in \mathcal{H}$. Let $y = 0$, then $L(h, y) > -\infty$ by definition because any $h \in \mathcal{H}$ results in a number on the real line excluding ∞. Secondly, there exists a capital savings and consumption choice $h = (h_1, h_2) \in \mathcal{H}$ such that $L(h, y) < \infty$ for all $y \in \mathcal{D}$. For any μ with aggregate capital $K > 0$, the productive income for any agent is positive $I > 0$. Fix a constant $0 < \epsilon < I$ and set $h_1 = \epsilon$ for any exogenous shock and start capital. Set capital savings according to the budget constraint. Then, the value of y_1 does not change the value of the Lagrangian. The nonnegativity of the other two multipliers ensures $L(h, y) < \infty$.

Semicontinuity: What is missing to conclude, is the continuity property of the Lagrangian in the policies $h \in \mathcal{H}$ as well as in the Lagrange multipliers $y \in \mathcal{D}$ which simply follows from the definition of the Lagrangian. *Q.E.D.*

REFERENCES

Yves Achdou, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. Heterogeneous agent models in continuous time. *Working Paper*, 2014.

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.

Heinz H. Bauschke and Patrick L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. CMS Books in Mathematics. Springer, 2011.

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

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

Patrick Cheridito and Juan Sagredo. Existence of sequential competitive equilibrium in Krusell-Smith type economies. *Working Paper, Available at SSRN 2794629*, 2016a.

Patrick Cheridito and Juan Sagredo. Comment on 'Competitive equilibria of economies with a continuum of consumers and aggregate shocks' by J. Miao. *Working Paper, Available at SSRN 2815971*, 2016b.

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.

Wouter J. den Haan, Kenneth L. Judd, and Michel Juillard. Computational suite of models with heterogeneous agents ii: Multi-country real business cycle models. *Journal of Economic Dynamics and Control*, 35(2):175–177, 2011.

Darrell Duffie and Wayne Shafer. Equilibrium in incomplete markets: I: A basic model of generic existence. *Journal of Mathematical Economics*, 14(3):285–300, 1985.

Darrell Duffie and Wayne Shafer. Equilibrium in incomplete markets: Ii: Generic existence in stochastic economies. *Journal of Mathematical Economics*, 15(3):199–216, 1986.

Darrell Duffie, John Geanakoplos, Andreu Mas-Colell, and Andrew McLennan. Stationary markov equilibria. *Econometrica*, pages 745–781, 1994.

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.

Sunghyun Henry Kim, Robert Kollmann, and Jinill Kim. Solving the incomplete market model with aggregate uncertainty using a perturbation method. *Journal of Economic Dynamics and Control*, 34(1):50–58, 2010.

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

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

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.

Jianjun Miao. Competitive equilibria of economies with a continuum of consumers and aggregate shocks. *Journal of Economic Theory*, 128(1):274–298, 2006.

George J. Minty. Monotone (nonlinear) operators in Hilbert space. *Duke Mathematical Journal*, 29(3):341–346, 1962.

Robert R. Phelps. Lectures on maximal monotone operators. *Extracta Mathematicae*, 12(3): 193–230, 1997.

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 associated with saddle-functions and minimax problems*, volume 18 of *Proceedings of Symposia in Pure Mathematics: Nonlinear Functional Analysis, Part 1*, pages 248–257. American Mathematical Society, 1970.

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.

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.

Eberhard Zeidler. *Nonlinear Functional Analysis and its Applications I. (Fixed Point Theorems)*. Springer, 1986a.

Eberhard Zeidler. *Nonlinear Functional Analysis and its Applications III. (Variational Methods and Optimization)*. Springer, 1986b.

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.