# Solving discrete time heterogeneous agent models with aggregate risk and many idiosyncratic states by perturbation

Zhou Shangyao

December 14, 2020

1. Background

# Background

- Models of heterogeneous agents have become widespread in macroeconomics, at least since Krusell and Smith (1997, 1998) developed the first widely applicable algorithm to solve them in an environment of aggregate risk.
- One of the most popular and powerful of these methods was originally developed by Reiter (2002, 2009).
- This method extends perturbation methods to heterogeneous agent environments, that is, it builds on the methods often used to solve dynamic stochastic general equilibrium models with a representative agent.

# Background

- The extension of perturbation methods to heterogeneous agent models relies on writing the model in the form of a nonlinear difference equation that is function valued instead of vector-valued (as in representative agent models).
- This equation is then (linearly) approximated around the stationary equilibrium of the heterogeneous agent model without aggregate risk.
- The two functionals that enter the difference equation are the distribution of agents over idiosyncratic states (e.g., the wealth distribution) and the function (value or policy function) that describes the optimal individual behavior.

# Background

- The key practical issue is how to approximate the functionals involved because they need to be replaced by finite dimensional objects for the actual computation of the model's dynamics.
- In particular, when the individual planning problem is rich insofar as it has many idiosyncratic states, this issue is severe.
- The first idea to tackle this issue was to be as sparse as possible in the parametric approximation of functions when solving for the stationary equilibrium (see, e.g., Reiter (2009)), for example, through sparse grid methods in the dynamic planning problem.
- In other words, these methods rely on achieving dimensionality reduction ex ante, before solving for the stationary equilibrium, and hence impose a numerical constraint on this solution.

# Background

- An alternative attack, also suggested by Reiter (2010a), is to use singular value decomposition for dimensionality reduction of the Jacobian of the system after linearizing the difference equation but before solving it.
- However, the above methods will also lead to the calculation of a very large Jacobian.
- In fact, the solution of the stationary equilibrium provides us with a priori knowledge for the distribution function and the value/policy function of heterogeneous agents.
- Therefore, we propose a dimensionality reduction step after the stationary equilibrium of the economy (i.e., without aggregate risk) has been determined, but before perturbing the system.

# What this paper do

- In detail, we suggest using sparse expansions of value and distribution functions around their nonsparse stationary equilibrium counterparts.

- First, we write the value function in the stationary equilibrium as a sum of a full set of basis functions and determine the coefficients on these.

- Second, we split the high-dimensional distribution function into the histograms of its marginals and their (joint) copula.

- We show, both for an incomplete markets model with one asset and for a model with two assets, that the assumption of a fixed copula has little impact on the model dynamics but substantially speeds up the computation.

# 2. Method

# Prerequisites and notation

- Let $S_t \in \mathbb{R}^n$ denotes the aggregate states in the economy, $s_{i,t} \in \mathbb{R}^m$ for the idiosyncratic states of individual $i$ at time $t$.
- The heterogeneous agents distribution function denoted by $\mu_t$.
- In general, $S_t$ and $s_{it}$ can be partitioned into an exogenous stochastic and an endogenous deterministic component as follows:

$$S_t = \begin{bmatrix} X_t \\ D_t \end{bmatrix}, \quad s_{it} = \begin{bmatrix} x_{it} \\ d_{it} \end{bmatrix}$$

- where the dimension is respectively $n = n_x + n_d$ and $m = m_x + m_d$.

# Prerequisites and notation

- Assuming that all stochastic variables follows the stationay Markov chain, such that:

$$X_{t+1} = H^X(X_t) + \varepsilon_{t+1}, \quad x_{it+1} = h^x(x_{it}) + \epsilon_{it+1}$$

- The innovations $\varepsilon_{t+1}, \epsilon_{it+1}$ have variances $\omega\Omega$ and $\sigma\Sigma$ for the aggregate and idiosyncratic variables, respectively.

- The idiosyncratic endogenous state variables $d_{it}$ are chosen by households in order to maximize their utility, such as:

$$\nu(x_{it}, d_{it}, S_t, \mu_t) = \max_{d_{it+1}} u(x_{it}, d_{it}, d_{it+1}; P_t) + \beta\mathbb{E}\nu(x_{it+1}, d_{it+1}, S_{t+1}, \mu_{t+1})$$

- where $d_{it+1}$ subjects to the correspondence, i.e. $d_{it+1} \in \Gamma(x_{it}, d_{it}, P_t)$ and $P_t = P(X_t, D_t, \mu_t)$ represents the aggregate pricing kernel.

# Prerequisites and notation

- From the individual's point of view, aggregates and distributions only matter through prices;
- Thus we can simplify the notation for the above Bellman equation to be:

$$\nu_t(x, d) = \max_{d' \in \Gamma_t(x,d)} u_t(x, d, d') + \beta \mathbb{E} \nu_{t+1}(x', d')$$

- Based on the solution to the above operator (Bellman equation), we can directly obtain the individual policy function $h^d(\cdot)$ which will be further introduced thereafter.

# Prerequisites and notation

- To close the model, we need a description of market clearing.
- We define an excess demand function $\Phi_t(h_t^d, \mu_t)$ that maps the idiosyncratic policies and the distribution, as well as prices and aggregate states (captured by the time index), into a real vector.
- In a bond economy with only IOUs, we would have:

$$\Phi = \int h_s^d(s) d\mu_t(s)$$

- and in an economy with government bonds this would be:

$$\Phi_t = \int h_s^d d\mu_t(s) - B_t$$

- where $B_t$ is the amount of government bonds.

## 2.2 Stationary equilibrium and approximate solution

# Stationary equilibrium and approximate solution

- Inspired by Rerter (2009), we approximate the aggregate dynamics around the stationary equilibrium. Therefore, firstly consider an economy without aggregate risk, that is, $\omega = 0$. Then define a stationary equilibrium as follows:

- Definition 1. A stationary equilibrium is a value function $\bar{v}$, a distribution function $\bar{\mu}$, a policy function $\bar{h}^d(s)$, and prices $\bar{P}$ such that:

# Stationary equilibrium and approximate solution

1. The individual policy $\bar{h}^d(s)$ is the maximizer of the Bellman equation given $\bar{P}$,

$$\bar{h}^d(x,d) = \arg\max_{d' \in \Gamma_{\bar{P}}(x,d)} u(x,d,d') + \beta \mathbb{E} \bar{\nu}(x',d')$$

2. The value function solves the Bellman equation (3) given the individual policy $\bar{h}^d(s)$.

3. Markets clear, that is, $\Phi(\bar{h}^d, \bar{\mu}) = 0$.

4. The distribution $\bar{\mu}$ is the stationary distribution of the Markov chain induced by $\bar{h}(s,\epsilon) := \begin{bmatrix} h^x(s) + \epsilon \\ \bar{h}^d(s) \end{bmatrix}$

# Stationary equilibrium and approximate solution

- Some computational details:
- the model is solved for a full tensor grid of points in $\mathbb{R}^m$ replacing the functionals by some parametric approximation, e.g. replace the value functions with splines with the nodes of the spline being equal to the grid points.
- The distribution is approximated by a step function (the density being replaced by a point mass) on the grid points. (This paper used)
- Further, the distribution of household can by approximated by a histogram $d\mu$, replacing the density.

# Stationary equilibrium and approximate solution

- Under the above method of approximation, the dynamics of distribution can be determined by point-mass $d\bar{\mu}$ and a transition matrix $\Pi_{\bar{h}}$ induced by individual policy function $\bar{h}$, such that:

$$d\bar{\mu} = d\bar{\mu}\Pi_{\bar{h}}$$

- This is the discrete time analogue to the Kolmogorov forward/Fokker–Planck equation in continuous time systems.

- Similarly, if value function is replaced by a linear interpolant, we obtain the solution to the Bellman equation is given by a finite vector of values $\bar{v}$, which needs to satisfy:

$$\bar{v} = u_{\bar{h}^d} + \beta\Pi_{\bar{h}}\bar{v}$$

- where $u_{\bar{h}^d}$ is the utility under optimal policy.

2.3 Sequential equilibrium with recursive individual planning

# Sequential equilibrium with recursive individual planning

- Definition 2. A sequential competitive equilibrium with recursive individual planning is a sequence of value functions $v_t$, a sequence of distribution functions $\mu_t$, a sequence of policy functions $h_t^d(s)$, a sequence of aggregate states $S_t$, and a sequence of prices $P_t$ such that at each point in time t:

  1. The individual policy is the maximizer of the Bellman equation given the prices $P_t$:

  $$h_t^d(x, d) = \arg \max_{d' \in \Gamma(x, d; P_t)} u\left(x, d, d'; P_t\right) + \beta \mathbb{E} \nu_{t+1}\left(x', d'\right)$$

  2. The value function solves the Bellman equation given the individual policy $h_t^d(s)$ and the expected continuation value $v_{t+1}$.

# Sequential equilibrium with recursive individual planning

3. Markets clear, that is,

$$\Phi_t\left(h_t^d, \mu_t, P_t, S_t\right) = 0$$

4. The distribution $\mu_{t+1}$ is induced by,

$$h_t(s, \epsilon) := \begin{bmatrix} h^x(s) + \epsilon_t \\ h_t^d(s) \end{bmatrix}$$

and distribution $\mu_t$;

5. The sequence of aggregate states is induced by,

$$\begin{bmatrix} X_{t+1} \\ D_{t+1} \end{bmatrix} = \begin{bmatrix} H^X\left(X_t, D_t\right) + \varepsilon_{t+1} \\ H^D\left(X_t, D_t, \mu_t\right) \end{bmatrix}$$

# Sequential equilibrium with recursive individual planning

- As in the stationary equilibrium, we replace the distribution function by a histogram and write the value function as a linear interpolant, then we have:

$$d\mu_{t+1} = d\mu_t \Pi_{h_t}$$

- Under the optimal policy induced $u_{h_t^d}$, the Bellman equation determined by:

$$\nu_t = u_{h_t^d} + \beta \Pi_{h_t} \nu_{t+1}$$

# Sequential equilibrium with recursive individual planning

- Combining these equilibrium conditions, we can summarize the sequential equilibrium conditions by the nonlinear difference equation given by:

$$F\left(d\mu_t, S_t, d\mu_{t+1}, S_{t+1}, \nu_t, P_t, \nu_{t+1}, P_{t+1}, \varepsilon_{t+1}\right) = \begin{bmatrix} d\mu_{t+1} - d\mu_t \Pi_{ht} \\ X_{t+1} - H^X\left(X_t, D_t\right) + \varepsilon_{t+1} \\ D_{t+1} - H^D\left(X_t, D_t, d\mu_t\right) \\ \nu_t - \left(u_{h_t^d} + \beta \Pi_{h_t} \nu_{t+1}\right) \\ \Phi_t\left(h_t^d, d\mu_t\right) \\ \varepsilon_{t+1} \end{bmatrix}$$

s.t.

$$h_t^d(s) = \arg \max_{d' \in \Gamma(x, d; P_t)} u\left(x, d, d'; P_t\right) + \beta \mathbb{E} \nu_{t+1}\left(x', d'\right)$$

- A sequential equilibrium now fulfills:

$$\mathbb{E}_t F\left(d\mu_t, S_t, d\mu_{t+1}, S_{t+1}, \nu_t, P_t, \nu_{t+1}, P_{t+1}, \varepsilon_{t+1}\right) = 0$$

# Sequential equilibrium with recursive individual planning

- In the above equilibirum, we can define the aggregate states and the control variables as:

$$\hat{S}_t := [d\mu_t, X_t, D_t]'$$

$$\hat{C}_t := [\nu_t, P_t]'$$

- If we are working with first-order conditions, value functions might be replaced with marginal utilities.

# Sequential equilibrium with recursive individual planning

- Based on the definition of two type of variables, i.e. predetermined and non-predetermined variables, we can obtain following Jacobian matrix:

$$\begin{bmatrix} F_{\hat{S}} & F_{\hat{S}'} & F_{\hat{C}} & F_{\hat{C}'} \end{bmatrix}$$

- Let $A := \left[F_{\hat{S}'}; F_{\hat{C}'}\right]$ and $B := \left[F_{\hat{S}}; F_{\hat{C}}\right]$.
- According to the generalized Schur form (QZ decomposition), the above matrix satisfies that:

$$QAZ = S \quad , \quad QBZ = T$$

- where $S$ and $T$ are both upper triangular, generalized eigenvalue $\lambda(A, B) = \{t_{ii}/s_{ii} : s_{ii} \neq 0\}$ and $Q, Z$ are unitary matrices;

# Sequential equilibrium with recursive individual planning

- Bayer and Luetticke (2020) solves linearized difference equations system by relating its solution to generalized eigenvalue problem:

$$\underbrace{\left[\begin{array}{cc} F_{\hat{S}'} & F_{\hat{C}'} \end{array}\right]}_{A:=} Z\Lambda = -\underbrace{\left[\begin{array}{cc} F_{\hat{S}} & F_{\hat{C}} \end{array}\right]}_{B:=} Z,$$

- Actually, we can easily obtain solution by solving transformed dynamic system from QZ decomposition:

$$Sx^*_{t+1} = Tx^*_t + QC\varepsilon_t$$

2.5 State-space reduction: Fixed copula, compressed value function

# Some computational issues

- In practice, however, solving the generalized eigenvalue problem (qz-decomposition of A, B) becomes easily numerically infeasible when the number of state variables (and controls) becomes very large;
- Consider, for example, a household planning problem with two assets and idiosyncratic income;
- Even if we use only 9 points for the income grid and 50 points for each of the two asset grids, then both $d\mu$ and $v$ are vectors with a length of 22,500 entries;
- As the function system $F(\cdot) = 0$ contains both the law of motion to $d\mu$ and Bellman equation for each household, thus there would be more than $45000 \times 45000$ entries.

# Dimensionality Reduction in Control Space

- To aviod the computational infeasibility induced by large dimension of state space, this paper propose to use the fixed copula and compressed value function;

- More specifically, we achieve dimensionality reduction of control space ( value $v$ ) by following parametrization:

$$\hat{v}_t(s) = g_v\left(s; \theta_t, \overline{v}\right)$$

- where $g_v$ is the parametric function we selected and the dimension of time-varying parameter vector $\theta_t$ is much smaller than the size of the tensor grid for $s$;

- How to select and construct $g_v$?

# Dimensionality Reduction in Control Space

- In order to reduce the dimensionality of coefficients that is used for parametrization, one particularly useful way is to select Chebyshev polynomial as functional basis.
- Let $\bar{v}$ be the array of the value function values at the nodes of the full tensor grid in the stationary equilibrium.
- Then let $\bar{\Theta} = \text{dct}(\bar{\nu})$ be its cosine transform which maps array $\bar{v}$ into the space of coefficients of fitted Chebyshev polynomial $\bar{\Theta}$;
- Moreover, the dimensionality of $\bar{\Theta}$ could still be large and we can further achieve reduction by preserving parts of the above coefficients.

# Dimensionality Reduction in Control Space

- Define $\mathcal{I}$ as the index set of some $\alpha\%$ largest elements of $\bar{\Theta}$;
- Then the usual sparse coefficient vector can be calculated by:

$$\tilde{\Theta} = \begin{cases} \bar{\Theta}(i) & \forall i \in \mathcal{I} \\ 0 & \text{else} \end{cases}$$

- However, this paper do not use sparse coefficients in stationary equilibrium, but use it when calculating the dynamics, that is:

$$\hat{\Theta}\left(\theta_t\right) = \begin{cases} \bar{\Theta}(i) + \theta_t(i) & \forall i \in \mathcal{I} \\ \bar{\Theta}(i) & \text{else} \end{cases}$$

- Therefore, in the steady state i.e. $\theta_t = 0$, the method used here can fully recovers the results in stationary equilibrium value function.

# Dimensionality Reduction in distribution function

- This paper split the determination of distribution $d\mu$ into a copula $\Xi_t$ and marginal distribution $\{\mu_{1t}(s), ..., \mu_{mt}(s)\}$:

$$\mu_t(s) = \Xi_t \{\mu_{1t}(s), \ldots, \mu_{mt}(s)\}$$

- An shortage in this paper is that we treat state contingent coputa $\Xi_t$ as fixed one which obtained from stationary equilibrium $\bar{\Xi}$.

- Under the above method of dimensionality reduction, the dynamic system F replaces value functions and distributions by the parameters $\{\theta_t, \mu_{1t}, ..., \mu_{mt}\}$.

- Nevertheless, the dimensionality reduction also lead to the number of equations in dynamic system being larger than the number of variables.

# Dimensionality Reduction in distribution function

- How to reduce the dimension of equations system?
- For control variables ( the value $v_t$ of each households ), we can replace error term of the Bellman equations such that

$$\Delta_\nu \left(v_t, v_{t+1}, P_t\right) := \nu_t - \left(u_{h_t^d} + \beta \Pi_{h_t} \nu_{t+1}\right)$$

to be:

$$\Delta_\nu \left(\theta_t, \theta_{t+1}, P_t\right) := \theta_t - \text{dct}\left\{T\left[\hat{\mathcal{V}}\left(\theta_{t+1}\right) \mid P_t\right]\right\}$$

- where $\hat{\mathcal{V}}\left(\theta_{t+1}\right)$ is the parametric map $\theta \to v$, and $T(\cdot)$ denotes the operator in Bellman equation.
- We can constructs a similar representation for $d\mu_t$ and fixed copula $\bar{\Xi}$ to replace the law of motions in dynamic system.

# 2.6 The algorithm in a nutshell

# The algorithm in a nutshell

- For the algorithm, define grid $s^j = \{d_1^j, ..., d_{n_j}^j\}$ for each $j = 1, ..., m_d$ of the idiosyncratic endogenous state variables $d^j$, with $n_j$ being the number of grid points for variable $j$.
- define grid $s^0 = \{x_1, ..., x_{n_0}\}$ for the exogenous stochastic one, $x$, which evolves with the transition matrix $\Pi_x$;
- Let $\mathcal{S} = \otimes_{j=0...m_d} s^j$ be the tensor product of $m_d + 1$ grids, and let $\mathcal{IS}$ be the corresponding tensor product of the indexes.
- Define $\hat{\nu}\left[\left(x, d^1 \ldots d^{m_d}\right) \mid \Pi_x \mathcal{V}\right]$ as the linear interpolant defined by the mesh $\mathcal{S}$ and node values $\Pi_x \mathcal{V}$

# The algorithm in a nutshell

- Prerequisites 1. 1. Define for a given price system $P$ a mapping $T(\mathcal{V} \mid P) : \mathbb{R}^J \to \mathbb{R}^J$ such that

$$\forall s = \left(x, d^1 \dots d^{m_d}\right) \in \mathcal{S}$$
$$T(\mathcal{V} \mid P)(s) := \max_{\left(d^{1'} \dots d^{m_{d'}}\right) \in \Gamma(s,P)} u\left(s, d^{1'} \dots d^{m_{d'}}\right)$$
$$+\beta \hat{\nu}\left[\left(x, d^{1'} \dots d^{m_{d'}}\right) \mid \Pi_x \mathcal{V}\right]$$

- 2. Define a mapping $\Pi = \Pi\left(\mathcal{V}_P\right) : \mathbb{R}^J \to \mathbb{R}^{J \times J}$ such that

$$\forall k = \left(k^0 \dots k^{m_d}\right), l = \left(l^0 \dots l^{m_d}\right) \in \mathcal{IS} : \Pi\left(\mathcal{V}_P\right)(k, l) = \Pi_x\left(k^0, l^0\right) \prod_{j=1}^{m_d} \Pi_{dj}(k, l)$$

# The algorithm in a nutshell

- where $\Pi_{d^j}$ are the coefficients to represent the policy $h_P^d(x) = \left(h_1^d(x) \ldots h_{m_d}^d(\right.$ as convex combinations of the nearest neighbors on the index mesh $\mathcal{IS}$ , that is,

$$\Pi_{d^j}(k,l) = \begin{cases} 0 & \text{if } h_j^d(k) \notin \left[d_{l-1}^j, d_{l+1}^j\right] \\ 1 - \frac{h_j^d(k) - d_l^j}{d_{l+1}^j - d_l^j} & \text{if } d_{l+1}^j \geq h_j^d > d_l^j \\ \frac{h_j^d - d_{l-1}^j}{d_l^j - d_{l-1}^j} & \text{if } d_l^j \geq h_j^d \geq d_{l-1}^j \end{cases}$$

- 3. The discrete cosine transformation of array $\mathcal{V}$.

# The algorithm in a nutshell

- Algorithm 1. 1. Finding the stationary equilibrium

- (a) For a given price system $P$, iterate $T^{(n)} = \underbrace{T\Big(T\Big(\dots T\Big(\mathcal{V}^{(0)} \mid P\Big) \mid P\Big) \mid}_{n \text{ times}}$ until convergency to obtain an equilibrium value function $\mathcal{V}_P$ as the limit $n \to \infty$.

- (b) Calculate the equilibrium distribution $d\mu_P$ by solving $d\mu_P = d\mu_P \Pi(\mathcal{V}_P)$.

- (c) Calculate excess demand $\Psi$ as a function $\Psi(h_P^d, d\mu_P)$.

- (d) Search over prices, repeating (a) to (c) until $\Psi(h_P^d, d\mu_P) = 0$.

# The algorithm in a nutshell

- 2. Dimensionality reduction

- (a) Define the joint distribution function $\bar{\mu}(s) = \sum_{x \leq s} \overline{d\mu}(x)$. Define $\bar{\mu}^j \in [0,1]$, $j = 0, ..., m_d$ as the $m_d + 1$ vectors of the marginal distribution w.r.t the $n_j$ points on the $s_j$ - grids.
Generate the fixed copula $\bar{\bar{\Xi}}\left(\mu^0, \ldots, \mu^{m_d} \mid \bar{\mu}\right) : [0,1]^{m_d+1} \to [0,1]$ as the interpolant of $\bar{\mu}$ on the tensor product $\otimes_{j=0}^{m_d} \bar{\mu}^j$.

- (b) Calculate the discrete cosine transformation of $\bar{\mathcal{V}}$ along all $m_d+1$ dimensions. This yields coefficients $\overline{\Theta}$. Find the minimal indxe set $\mathcal{I}$ such that

$$\frac{\sum_{i \in \mathcal{I}} \overline{\Theta}(i)^2}{\sum_i \overline{\Theta}(i)^2} > 1 - \epsilon$$

# The algorithm in a nutshell

- 2. Dimensionality reduction
- (c) Define a sparse vector that has $\#\mathcal{I}$ nonzero entries, , and hence is effectively much shorter than $\Theta \in \mathbb{R}^J$.
- In the following, when we speak of perturbing $\theta_t$, we mean perturbing its nonzero entries, which is given by:

$$\hat{\Theta} = \begin{cases} \bar{\Theta}(i) + \theta(i) & \text{if } i \in \mathcal{I} \\ \bar{\Theta}(i) & \text{if } i \notin \mathcal{I} \end{cases}$$

- 2. Linearization
  - (a) Define the following objects:
  - We apply the discrete cosine transformation to the value functions and evaluate on all points in $\mathcal{S}$:

$$\Delta_{\nu}\left(\theta_t, \theta_{t+1}, P_t\right) := \theta_t - \text{dct}\left\{T\left[\hat{\mathcal{V}}\left(\theta_{t+1}\right) \mid P_t\right]\right\} \in \mathbb{R}^J$$

  - for all variables $j = 0, ..., m_d$ the difference between the marginal distribution for time $t+1$ obtained from iterating forward once the distribution implied by $(\mu_t^j)_{j=0,\dots,m_d}$ and the fixed copula $\bar{\Xi}$,

$$\Delta_{\mu}^{*}\left[\left\{\mu_t^j\right\}_{j=0\dots m_d}, \left\{\mu_{t+1}^j\right\}_{j=0\dots m_d}, P_t, \theta_{t+1}\right] \in \mathbb{R}^{\sum_{(m_d+1)} n_j}$$

  - the excess demand function

# The algorithm in a nutshell

- 2. Linearization

- (b) Use these differences equations to define a function:

$$F\left(\left\{\mu_t^j\right\}_{j=0\ldots m_d}, S_t, S_{t+1}, \left\{\mu_{t+1}^j\right\}_{j=0\ldots m_d}, \theta_t, P_t, \theta_{t+1}, P_{t+1} \mid \Xi, \overline{\mathcal{V}}, \mathcal{I}\right)$$

that describes the economy as a system of nonlinear difference equations

$$F = \left[\begin{array}{c} \Delta_\nu^*\left(\theta_t, \theta_{t+1}, P_t\right) \\ \Delta_\mu^*\left[\left\{\mu_t^j\right\}_{j=0\ldots m_d}, \left\{\mu_{t+1}^j\right\}_{j=0\ldots m_d}, P_t, \theta_{t+1}\right] \\ S_{t+1} - H(S_t) \\ \Phi\left(\left\{\mu_t^j\right\}_{j=0\ldots m_d}, \theta_{t+1}, P_t, S_t, S_{t+1}\right) \end{array}\right]$$

# The algorithm in a nutshell

- 2. Linearization
- (c) Calculate the Jacobian of F. Define A, B as defined in the text before and as in Schmitt-Grohé and Uribe (2004).
- (d) Calculate the qz decomposition and solve for the linearized dynamics using the algorithm provided by Schmitt-Grohé and Uribe (2004).

# 3. Examples

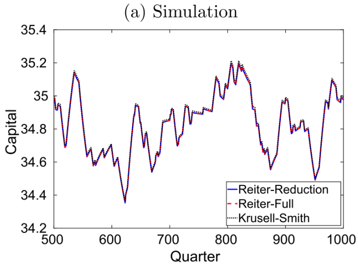
# Examples

- In the following, we discuss two examples to illustrate our modification of Reiter's method to solve general equilibrium models with heterogeneous agents and aggregate risk.
- Both examples share the same model of consumption-savings choice in which households face uninsurable income risk and use assets to self-insure.
- We then specify two variants of the model:
  - The first one without nominal frictions and only one asset, that is, the setup of Krusell and Smith (1998);
  - second, a setup with two assets of different liquidity and a nominal rigidity.

# 4. Numerical Performance

# Comparision to Krusell and Smith(1998)

- Figure 1 shows simulations of the K-S model for three different solution methods: (1) perturbation with state-space reduction via the fixed copula assumption and policy function compression; (2) perturbation with a full policy function and histogram; (3) the original Krusell and Smith algorithm.



# Comparision to Krusell and Smith(1998)

- Table 1 confirms this. The mean absolute error between the time series from the two linearization methods and the K-S algorithm is 0.03%. What is more, the linearization methods with and without state and control space reduction yield basically the same simulation for the aggregate stock of capital with a maximum absolute error of 0.001%.

TABLE 1. Simulation errors relative to Krusell and Smith algorithm.

| | Absolute difference (in %) of log capital stocks $K_t$ between simulations | | |
|---|---|---|---|
| | Reiter-Reduction vs. K-S | Reiter-Full vs. K-S | R.-Reduction vs. R.-Full |
| Mean | 0.0324 | 0.0324 | 0.0003 |
| Max | 0.0670 | 0.0662 | 0.0012 |

# Comparision to Krusell and Smith(1998)

- To further evaluate the accuracy of our solution method,we use the error metrics suggested by Den Haan (2010a), comparing the simulation from the linearized solution of the model to one in which we solve for the equilibrium interest rate every period and track the full histogram over time.

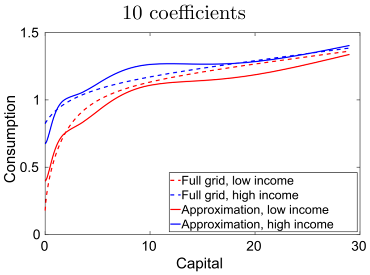
TABLE 2. Den Haan errors.

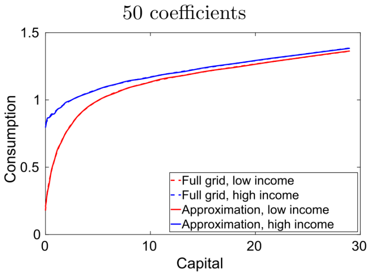
| | Absolute error (in %) for log capital $K_t$ | | |
|---|---|---|---|
| | Reiter-Reduction | Reiter-Full | K-S |
| Mean | 0.0100 | 0.0102 | 0.0051 |
| Max | 0.0191 | 0.0193 | 0.0131 |

*Note*: Differences in percent between the simulation of the linearized solutions of the model and simulations in which we solve for the intratemporal equilibrium prices in every period and track the full histogram over time for $t = \{1, \ldots, 1000\}$; see Den Haan (2010a).

# Details on using the DCT for dimensionality reduction

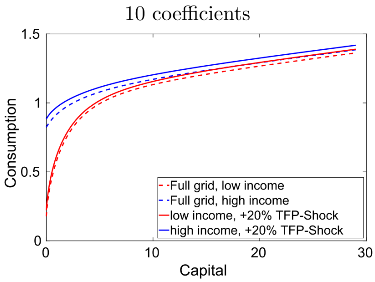
- First, we compare the policy function in the stationary equilibrium with the policy function that would have been obtained by solving the stationary equilibrium with the sparse Chebyshev polynomial.
- The approximation with 10 coefficients is fairly rough and unsatisfactory. It shows excessive fluctuation and oscillation. With 50 out of 200 coefficients, the approximation becomes much better.



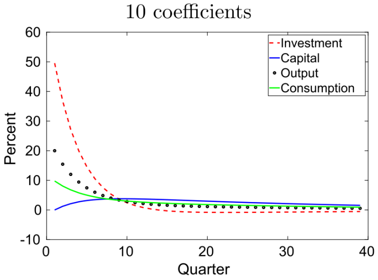
# Details on using the DCT for dimensionality reduction

- However, low number of coefficients, however, has hardly any impact on the response of individual policies to a TFP shock.
- Figure 3 shows how consumption policies change for different levels of sparseness of $\theta$.
- The reason for this is that the shock mostly produces a level shift for consumption together with a small change in the steepness of the consumption policy in wealth and income. Changes in the large coefficients of the discrete cosine transform of the consumption policy can capture these shifts well.



# Details on using the DCT for dimensionality reduction

- Not very surprisingly, with these small differences in individual policies, the aggregate responses look also indistinguishable; see Figure 4.



# Details on using the DCT for dimensionality reduction

- we also investigate the numerical error induced by DCT-selection. Table 4 show the result of comparision and we can find that, despite the complete polynomial choice has somewhat stronger theoretical underpinning, it still performs substanially worse.
- The reason for the superior performance of the adaptive DCT-based method is that across different income states, the policy functions are relatively similar in the stationary equilibrium; the DCT method detects this, and this remains true even when prices change after a shock.

TABLE 4. Comparison of DCT-based coefficient selection to a nonadaptive rule.

| Degree of polynomial, $N$ | 50 | 40 | 30 | 20 |
|---|---|---|---|---|
| Number of coefficients | 101 | 81 | 61 | 41 |
| Selection of coefficients | Max absolute difference of log capital stocks ($\times 1e^{-8}$) | | | |
| (a) Complete polynomial | 0.08 | 0.80 | 6.24 | 37.37 |
| (b) DCT | 0.10 | 0.43 | 0.07 | 0.46 |
| | Mean absolute difference of log capital stocks ($\times 1e^{-8}$) | | | |
| (a) Complete polynomial | 0.02 | 0.25 | 1.95 | 11.59 |
| (b) DCT | 0.03 | 0.13 | 0.02 | 0.13 |

# Details on using the copula for dimensionality reduction

- To understand how restrictive the assumption of a fixed copula is, we compare the model-implied distributions over time for the solution that does no reduction (Reiter–Full) and our method, which fixes the copula.

- We simulate and compute the metric of Jensen–Shannon distance(JSD) for distribution functions. The JSD is defined as follows,

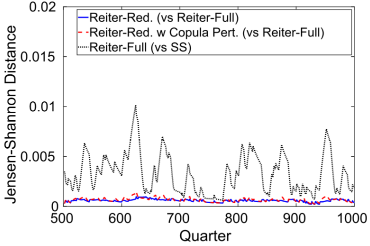
$$\text{JSD}\,(f_1, f_2) = \sqrt{\frac{1}{2} \sum_{x \in X} f_1(x) \log \left[ \frac{2 f_1(x)}{f_1(x) + f_2(x)} \right] + f_2(x) \left[ \log \frac{2 f_2(x)}{f_1(x) + f_2(x)} \right]}$$

- In particular, we consider two different cases of comparison.

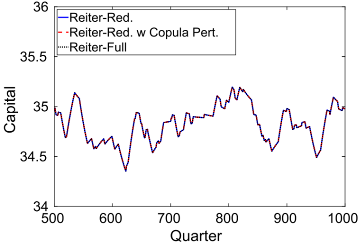
# Details on using the DCT for dimensionality reduction

- For the first case, we simulate the model using TFP shock the driving force.
- Indeed, we find that the approximation error measured in terms of the Jensen–Shannon distance (left column) between the joint distribution (of assets and income) in the Reiter solution with and without the fixed copula assumption is an order of magnitude smaller than the distance between either solution and the stationary equilibrium distribution. The distance between the distributions is, at 0.0005, negligibly small.

Distance of asset and income distributions

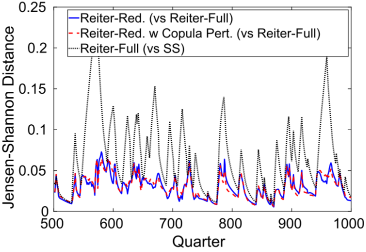


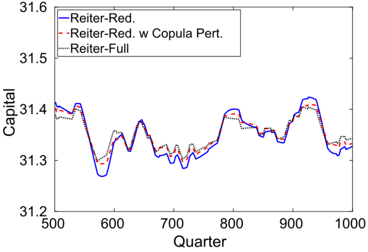

Aggregate capital stocks



# Details on using the DCT for dimensionality reduction

- For the second case, , we simulate the model with shocks to idiosyncratic income uncertainty as a driver of the cycle.
- Now, the distance of the simulated distributions to the steady-state one is much larger and the difference between the distribution from the full Reiter solution and the one with a fixed copula attains a significant order of magnitude. We also find some difference in the fluctuations of the capital stock that the model implies—a model where the fluctuations in capital are small, as there is little aggregate feedback.



# Extension : Two-asset model

- The true advantage of the state and control space reduction through separating marginals and copula and compressing the latter alongside the value functions lies in tackling the curse of dimensionality.
- Making it possible to solve models with high dimensional heterogeneity.
- In the following, we provide accuracy statistics and computational time for our model with a portfolio choice between liquid and illiquid assets. This model features heterogeneity with respect to three dimensions:
  - (a) liquid asset holdings,
  - (b) illiquid asset holdings,
  - (c) idiosyncratic productivity.

# Extension : Two-asset model

- We solve the household problem on 100 grid points for both asset choices and 12 grid points for productivity. With 120,000 states and 240,000 controls (for the two value functions), it is infeasible to solve for the aggregate dynamics of the model on the full histogram.
- The fixed copula approximation reduces the number of states to 236.
- Maintaining only the coefficients of the discrete cosine transform of the value functions with the cumulative highest 99.9999% energy reduces the number of controls to 1427.

TABLE 5. Run times for two-asset model.

| | Running times* | |
|---|---|---|
| | Stationary equilibrium | Reiter-Reduction |
| In seconds | 1311 | 326 |

*On a Dell laptop with an Intel i7-7500U CPU @ 2.70 GHz, 4 cores. Code in Julia.

# Extension : Two-asset model

- Table 6 shows the error metric suggested by Den Haan (2010a) for the capital stock implied by the two-asset model in response to TFP shocks.
- The maximum absolute error is 0.12% and the mean absolute error is 0.05%, which are comparable to the errors in the case of single-asset model.
- To assess how sensitive this result is, we also consider a specification that perturbs the copula, too, and a specification that retains more coefficients of the DCT.

TABLE 6. Accuracy for two-asset model.

| | Absolute error (in %)* | | | |
|---|---|---|---|---|
| | Mean, for ... | | Max, for ... | |
| | capital $K_t$ | bonds $B_t$ | capital $K_t$ | bonds $B_t$ |
| Baseline | 0.033 | 0.081 | 0.092 | 0.617 |
| Retain *more* DCT coefficients | 0.031 | 0.080 | 0.087 | 0.610 |
| Baseline + perturb copula | 0.043 | 0.080 | 0.151 | 0.777 |

# Extension : Two-asset model

- Table 7 shows that for this business cycle calibration with TFP, monetary, and uncertainty shocks also the business cycle statistics do vary relatively little, when we change the numerical specification.

TABLE 7. Business cycle statistics for the two-asset model.

| | $\sigma(Y_t)$ | $\sigma(C_t)$ | $\sigma(I_t)$ |
|---|---|---|---|
| Baseline | 1.40 | 1.39 | 4.49 |
| Retain *more* DCT coefficients | 1.42 | 1.42 | 4.49 |
| Baseline + perturb copula | 1.31 | 1.28 | 4.56 |

# Extension : Two-asset model

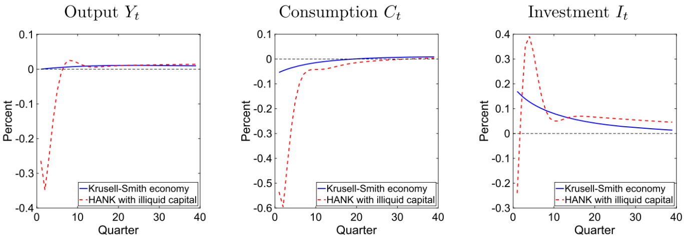
- In addition to the business cycle statistics, we also look at how the model performs in terms of asset prices; see Table 8.
- We report the average and maximum absolute deviation from asset market clearing that the linearized solution produces, that is, we evaluate the difference in asset supply and demand for both $B$ and $K$ given prices and the wealth distribution that we get from the simulation of the linearized solution.
- Since the model produces a steady state return difference, an illiquidity premium for capital, the model gets a long way in terms of being close to the observed Sharpe-ratios of Jordà et.al (2019) that range between 0.6 for housing and 0.25 for equities.

TABLE 8. Asset market clearing for the two-asset model.

| | Deviation from Market Clearing in % | | | | |
|---|---|---|---|---|---|
| | Mean-Absolute | | Max-Absolute | | |
| | on $K_t$ | on $B_t$ | on $K_t$ | on $B_t$ | Sharpe-Ratio |
| Baseline | 0.03% | 0.05% | 0.29% | 0.65% | 0.70 |
| Retain *more* DCT coefficients | 0.03% | 0.05% | 0.28% | 0.64% | 0.67 |
| Baseline + perturb copula | 0.01% | 0.02% | 0.22% | 0.59% | 0.68 |

# Extension : Two-asset model

- Figure 7 shows the effect of higher uncertainty about idiosyncratic productivity in the Krusell and Smith model and the two-asset HANK model.
- Consumption falls in both models as households increase their precautionary savings in response to higher uncertainty.
- In the Krusell and Smith model, higher savings translate one-for-one into capital, which leads to an economic expansion.
- In the two-asset model, by contrast, households prefer to hold more liquid portfolios. They sell illiquid capital to save more in liquid assets. Higher uncertainty therefore causes a simultaneous fall in consumption, investment, and output.



# Conclusion

- In this paper, we propose an extension of Reiter's method to solve heterogeneous agent models with aggregate risk by perturbation.
- The state-space reduction is achieved by "lossy compression" of the value functions, which are control variables of the system, and by approximating the dynamics of the multidimensional distribution of individual characteristics by a distribution with an (almost) fixed copula and varying marginals.
- Both steps effectively reduce the problem that high-dimensional idiosyncratic state spaces pose and allow us to efficiently and precisely solve for the equilibrium dynamics of heterogeneous agent economies as we have shown in two examples.

**Thanks !**