Solving Heterogeneous Agent Models with Aggregate Uncertainty and many Idiosyncratic States in Discrete Time by Perturbation Methods

AMCM, Lillehammer

C. Bayer, U Bonn, CEPR & IZA

R. Luetticke, UCL & CEPR

February 2019

Heterogeneous agents models with aggregate uncertainty

. . . have become of widespread use in macro

because heterogeneity can change the transmission mechanism

- McKay et al. (2016), Kaplan et al. (2017), Luetticke (2018)

because re-distributional policies have business cycle impact

- e.g. McKay and Reis (2016)

because business cycle policies have distributional impact

- e.g. Gornemann et al. (2012)

because heterogeneity gives rise to new drivers of the cycle

- e.g. Guerrieri and Lorenzoni (2017), Mitman (2016), Bayer et al. (2018)

Heterogeneous agents models with aggregate uncertainty

Yet, these models are computational demanding to solve

- The original Krusell and Smith (1997, 1998) algorithm is notoriously slow
- Therefore, many papers use MIT shocks
- or are restricted to relatively simple household decisions

Heterogeneous agents models with aggregate uncertainty

Yet, these models are computational demanding to solve

- The original Krusell and Smith (1997, 1998) algorithm is notoriously slow
- Therefore, many papers use MIT shocks
- or are restricted to relatively simple household decisions

- We depart from the Reiter (2002, 2009) perturbation method
- And (try to) provide an accessible algorithm that can deal with high-dimensional heterogeneity

Placing what we do in the literature

& Intuition

Reiter (2002): Solve by perturbation

The heterogeneous agent model:

- Can be written as a non-linear difference equation
- that is function valued and
- needs to be linearized around the stationary equilibrium (StE)

Reiter (2002): Solve by perturbation

The heterogeneous agent model:

- Can be written as a non-linear difference equation
- that is function valued and
- needs to be linearized around the stationary equilibrium (StE)

- Functions need to be approximated by finite dimensional objects (e.g. coefficients of polynomials, splines, etc.)

Reiter (2009): Reduce dimensionality ex ante

Problem:

- Dimensionality of the difference equation is large

Proposal:

- Reduce dimensionality ex ante (before solving the StE)
- e.g. (sparse) splines to represent policy functions
- Then linearize

Winberry (2016) extends this to the distribution functions

Ahn et al. (2017): Reduce dimensionality after linearization

Problem:

- Hard to determine where to be sparse
- StE is easy to solve
- Why not use all information/precision here?

Proposal:

- Reduce dimensionality after linearization
- Relies on SVD of the Jacobians of the difference equation

What we do

Problem:

- Jacobian is large before reduction
- Only sparse in continuous time

Proposal:

- Reduce dimensionality after StE, but before linearization
- Extract from the StE the *important* basis functions to represent individual policies (akin to image compression)
- Perturb only those basis functions but use the StE as "reference frame" for the policies (akin to video compression)
- Similarly for distributions (details later)

Video compression

Setup

Recursive Dynamic Planning Problem

Consider a household problem in presence of aggregate and idiosyncratic risk

- S_t is an (exogenous) aggregate state
- s_{it} is a partly endogenous idiosyncratic state
- μ_t is the distribution over s

- Bellman equation:

$$v(s_{it}, S_t, \mu_t) = \max_{x \in \Gamma(s_{it}, P_t)} u(s_{it}, x) + \beta \mathbb{E} v(s_{it+1}(x, s_{it}), S_{t+1}, \mu_{t+1})$$

Recursive Dynamic Planning Problem

Consider a household problem in presence of aggregate and idiosyncratic risk

- S_t is an (exogenous) aggregate state
- s_{it} is a partly endogenous idiosyncratic state
- μ_t is the distribution over s

- Euler equation:

$$u'[x(s_{it}, S_t, \mu_t)] = \beta R(S_t, \mu_t) \mathbb{E} u'[x(s_{it+1}, S_{t+1}, \mu_{t+1})],$$

No aggregate risk

Recall how to solve for a StE

- Discretize the state space (vectorized)
- Optimal policy $\bar{h}(s_{it}; P)$ induces flow utility $\bar{u}_{\bar{h}}$ and transition probability matrix $\Pi_{\bar{h}}$
- Discretized Bellman equation

$$\bar{v} = \bar{u}_{\bar{h}} + \beta \Pi_{\bar{h}} \bar{v} \tag{1}$$

holds for optimal policy (assuming a linear interpolant for the continuation value)

No aggregate risk

- and for the law of motion for the distribution (histograms)

$$d\bar{\mu} = d\bar{\mu}\Pi_{\bar{h}} \tag{2}$$

No aggregate risk

Equilibrium requires

- $\bar{h}$ is the optimal policy given P and ν (being a linear interpolant)
- $\bar{\nu}$ and $d\bar{\mu}$ solve (1) and (2)
- Markets clear (some joint requirement on $\bar{h}, \mu, P$, denoted as $\Phi(\bar{h}, \mu, P) = 0$)

This can be solved for efficiently

- $d\bar{\mu}$ is vector corresponding to the unit-eigenvalue of $\Pi_{\bar{h}}$
- Using fast solution techniques for the DP, e.g. EGM
- Using a root-finder to solve for P

Introducing aggregate risk

With aggregate risk

- Prices and distributions change over time

Yet, for the household

- Only prices and continuation values matter
- Distributions do not influence the decisions directly

Redefining equilibrium (Reiter, 2002)

A **sequential equilibrium with recursive individual planning**

- A sequence of discretized Bellman equation, such that

$$\nu_t = \bar{u}_{P_t} + \beta \Pi_{h_t} \nu_{t+1} \tag{3}$$

  holds for optimal policy, h_t (which results from ν_{t+1} and P_t)
- and a sequence of histograms, such that

$$d\mu_{t+1} = d\mu_t \Pi_{h_t} \tag{4}$$

  holds given the optimal policy
- (Policy functions, h_t, that are optimal given P_t, ν_{t+1})
- Prices, distributions and policies lead to market clearing

Compact notation (Schmitt-Grohé and Uribe, 2004)

The equilibrium conditions as a non-linear difference equation

- Controls: $Y_t = \begin{bmatrix} \nu_t & P_t & Z_t^Y \end{bmatrix}$ and
- States: $X_t = \begin{bmatrix} \mu_t & S_t & Z_t^X \end{bmatrix}$
  where Z_t are purely aggregate states/controls
- Define

$$F\left(d\mu_t, S_t, d\mu_{t+1}, S_{t+1}, \nu_t, P_t, \nu_{t+1}, P_{t+1}, \varepsilon_{t+1}\right) \quad (5)$$

$$= \begin{bmatrix} d\mu_{t+1} - d\mu_t \Pi_{h_t} \\ \nu_t - \left(\bar{u}_{h_t} + \beta \Pi_{h_t} \nu_{t+1}\right) \\ S_{t+1} - H(S_t, d\mu_t, \varepsilon_{t+1}) \\ \Phi(h_t, d\mu_t, P_t, S_t) \\ \varepsilon_{t+1} \end{bmatrix}$$

s.t.

$$h_t(s) = \arg \max_{x \in \Gamma(s, P_t)} u(s, x) + \beta \mathbb{E} \nu_{t+1}(s') \quad (6)$$

Compact notation (Schmitt-Grohé and Uribe, 2004)

The equilibrium conditions as a non-linear difference equation

- Function-valued difference equation $\mathbb{E}F(X_t, X_{t+1}, Y_t, Y_{t+1}, \varepsilon_{t+1}) = 0$
- turns real-valued when we replace the functions by their discretized counterparts
- Standard techniques to solve linearized version

So, is all solved?

The dimensionality of the system F is still an issue

- With high dimensional idiosyncratic states, discretized value functions and distributions become large objects
- For example:
  4 income states (grid points)
  $\times$ 100 illiquid asset states
  $\times$ 100 liquid asset states
  $\Longrightarrow \geq 40,000$ control variables in F
- Same number of state variables

Our Method

Our idea

1.) Apply compression techniques as in video encoding

- Apply a discrete cosine transformation to all value/policy functions (Chebycheff polynomials on roots grid) ▸ DCT
- Define as reference "frame": the StE value/policy function
- Write fluctuations as differences from this reference frame
- Assume all coefficients of the DCT from the StE close to zero do not change after shock

Our idea

2.) Neglect changes in the rank correlation structure of μ

- Calculate the Copula, $\bar{C}$ of μ in the StE
- Perturb only the marginal distributions
- Use fixed Copula to calculate an approximate joint distribution from marginals

- Idea follows Krusell and Smith (1998) in that some moments of the distribution do not matter for aggregate dynamics

Details

1.) Apply compression techniques as in video encoding

- Let $\bar{\Theta} = dct(\bar{v})$ be the coefficients obtained from the DCT of the value function in StE
- Define an index set $\mathcal{I}$ that contains the x percent largest (i.e. most important) elements from $\bar{\Theta}$
- Let θ be a sparse vector with non-zero entries only for elements $i \in \mathcal{I}$
- Define $\tilde{\Theta}(\theta_t) = \begin{cases} \bar{\Theta}(i) + \theta_t(i) & i \in \mathcal{I} \\ \bar{\Theta}(i) & else \end{cases}$

Details

Decoding

- Now we reconstruct $\nu_t = \nu(\theta_t) = idct(\tilde{\Theta}(\theta_t))$
- This means that in the StE the reduction step adds no additional approximation error as $\nu(0) = \bar{\nu}$ by construction
- Yet, it allows to reduce the number of derivatives that need to be calculated from the outset

Details

Analogously for the histogram

- We μ_t as $\bar{C}(\bar{\mu}_t^1, \ldots, \bar{\mu}_t^n)$ for n being the dimensionality of the idiosyncratic states
- The StE distribution is obtained when $\mu = \bar{C}(\bar{\mu}^1, \ldots, \bar{\mu}^n)$
- Typically prices are only influenced through the marginal distributions
- The approach ensures that changes in the mass of one dimension, say wealth, are distributed in a sensible way across the other dimensions
- The implied distributions look "similar" to the StE one (different in (Reiter, 2009))

Details

Too many equations

- The system

$$F\left(\{d\mu_t^1,\ldots,d\mu_t^n\},S_t,\{d\mu_{t+1}^1,\ldots,d\mu_{t+1}^n\},S_{t+1},\right. \tag{7}$$
$$\left.\theta_t,P_t,\theta_{t+1},P_{t+1}\right)=$$
$$\begin{bmatrix} d\bar{C}(\bar{\mu}_t^1,\ldots,\bar{\mu}_t^n)-d\bar{C}(\bar{\mu}_t^1,\ldots,\bar{\mu}_t^n)\Pi_{h_t} \\ dct\left[idct(\tilde{\Theta}(\theta_t))-\left(\bar{u}_{h_t}+\beta\Pi_{h_t}idct(\tilde{\Theta}(\theta_{t+1}))\right)\right] \\ S_{t+1}-H(S_t,d\mu_t) \\ \Phi(h_t,d\mu_t,P_t,S_t) \end{bmatrix}$$

  has too many equations

- Use only difference in marginals and the differences on $\mathcal{I}$

A first working example

A simple KS economy

Incomplete Markets and TFP

- Household productivity can be high or low
- No contingent claims
- Households save in capital goods (which they rent out)
- Households supply labor (disutility) and consume (utility)
- Aggregate productivity (TFP) follows a log AR-1 process
- Cobb-Douglas production function

A simple KS economy

Numerical setup

- Asset grid has 100 points ($\implies$ a total grid size of 200)
- Policies solved by EGM (instead of VFI)

Different levels of “compression”

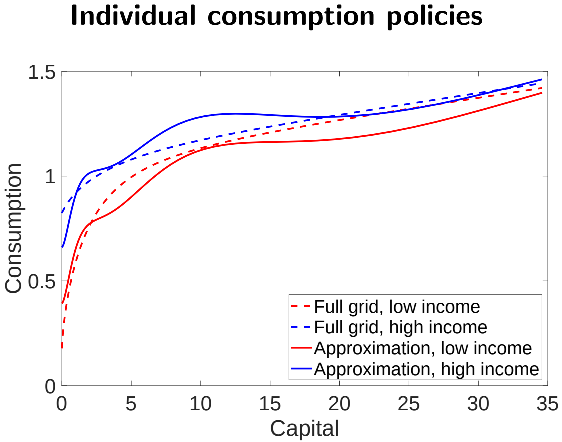


Figure: Policy and **10** most important basis functions

Different levels of "compression"

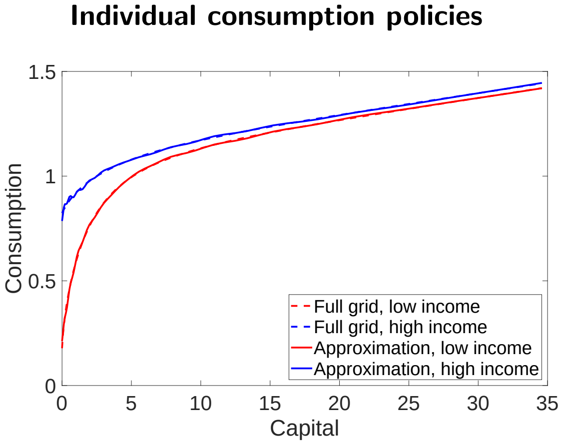


Figure: Policy and **50** most important basis functions

Different levels of "compression"

Individual policy response to a 20%TFP shock

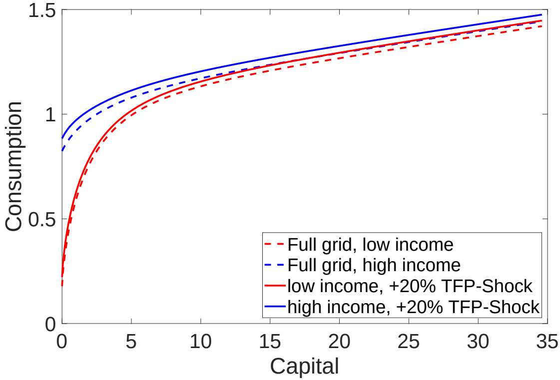


Figure: Perturbing **10** most important basis functions

Different levels of "compression"

Individual policy response to a 20%TFP shock

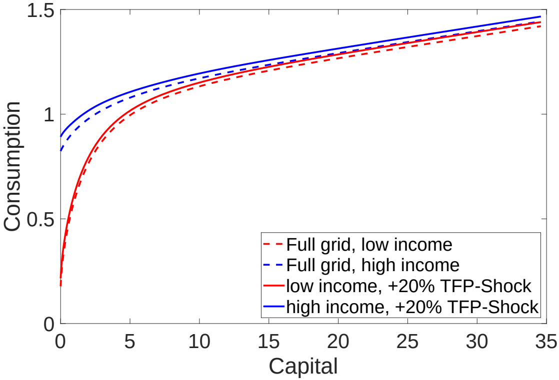


Figure: Perturbing all **200** basis functions

Different levels of "compression"

Aggregate response to a 20%TFP shock

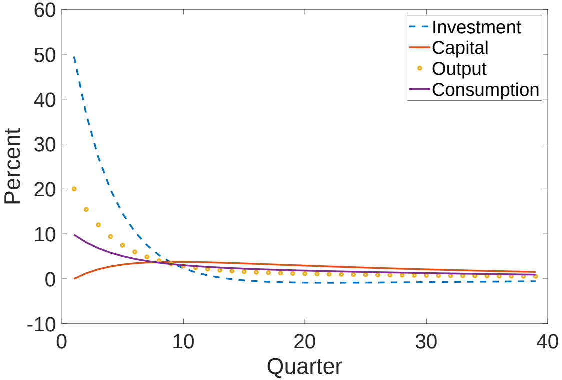


Figure: Perturbing **10** most important basis functions

Different levels of "compression"

Aggregate response to a 20%TFP shock

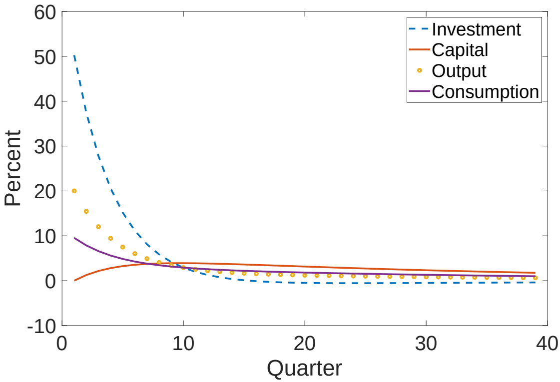


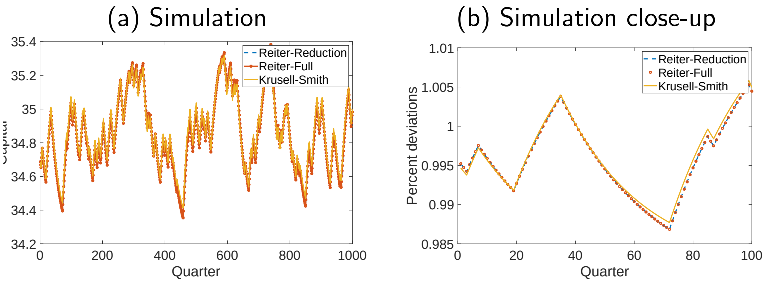
Figure: Perturbing all **200** basis functions

Taking stock

- When looking only at the StE policy function one concludes that roughly 50% of the information is needed to reconstruct the policies well
- This is roughly level of state reduction a Reiter (2009) approach would achieve
- Using the StE as reference one can achieve much higher reduction
- For the aggregate dynamics maintaining only 3-6% of the basis functions suffices

Simulation performance

Figure: Simulations of Krusell & Smith model



Notes: Both panels show simulations of the Krusell & Smith (1998) model with TFP shocks solved with (1) the Reiter method with our proposed state-space reduction, (2) the original Reiter method without state-space reduction, (3) the original Krusell & Smith algorithm

Error Statistics

Table: Den Haan errors

	Absolute error (in %) for capital K_t		
	Reiter-Reduction	Reiter-Full	K-S
Mean	0.0119	0.0119	0.1237
Max	0.0152	0.0152	0.3491

Notes: 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 $t = \{1, ..., 1000\}$; see Den Haan (2010)

Computing time

Table: Run time for Krusell & Smith model

	StE	K & S	Reiter-Reduction	Reiter-Full
in seconds	6.28	49.85	0.43	0.91

Notes: Run time in seconds on a Dell laptop with an Intel i7-7500U CPU @ 2.70GHz 4. Model calibration and number of grid points as in Den Haan et al. (2010). Code in Matlab.

What is it good for?
(an economically interesting example)

A two asset economy with nominal friction (Bayer et al., 2018)

K&S model plus

- a trading cost for capital
- a perfectly liquid bond
- a price setting friction for firms (Phillips curve)
- a central bank and a fiscal authority

A two asset economy with nominal friction (Bayer et al., 2018)

Analyze an increase in income shock variance

- requires both assets (100 points each)
- Here only 4 income states (allows to use only MATLAB's build in routines and compute on a laptop in reasonable time)
- uses both value functions and consumption policies as controls
- Full set would have $> 120,000$ variables
- Reduction to 204 distribution states and 635 controls for value functions and policies

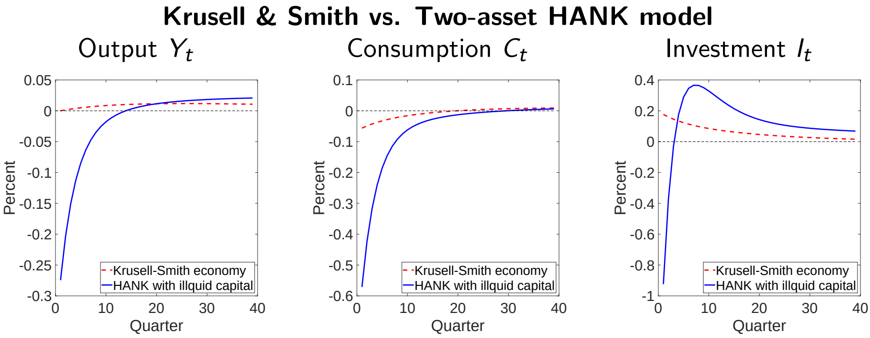
Error statistics

Table: Run times and accuracy for two-asset model

Running times	Stationary Equilibrium	Reiter-Reduction
In seconds	388.14	80.38
Absolute error (in %)	For capital K_t	For bonds B_t
Mean	0.0418	0.0828
Max	0.1425	0.4706

An uncertainty shock

Figure: Aggregate response to idiosyncratic uncertainty shock



Conclusion

No excuse!

- Even when heterogeneity is high dimensional,
- our algorithm is an easy approach to these models
- It is a fast and simple to code

Conclusion

No excuse!

- It requires knowledge of only two standard tools of macro:
  1. Solving a recursive het. agent model for a StE
  2. Linearizing a rep. agent model
  3. (and a little twist in between)
- The fixed design for dimensionality reduction allows to employ the method to estimate models with standard techniques

Discrete Cosine Transform

- A DCT expresses a finite sequence of data points in terms of sum of cosine functions at different frequencies
- Linear, invertible function $f = \Re^N - > \Re^N$ (equivalently: an invertible NxN matrix)
- x_n is transformed to X_n according to:

$$X_k = \sum_{n=0}^{N-1} x_n \cos\left[\pi / N(n + 1/2)k\right], k = 0, ..., N-1$$

- ▸ back

Ahn, S., Kaplan, G., Moll, B., Winberry, T., and Wolf, C. (2017). When inequality matters for macro and macro matters for inequality. *NBER Macroeconomics Annual*, Volume 32.

Bayer, C., Luetticke, R., Pham-Dao, L., and Tjaden, V. (2018). Precautionary savings, illiquid assets, and the aggregate consequences of shocks to household income risk. *Econometrica*.

Den Haan, W. J. (2010). Assessing the accuracy of the aggregate law of motion in models with heterogeneous agents. *Journal of Economic Dynamics and Control*, 34(1):79–99.

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

Gornemann, N., Kuester, K., and Nakajima, M. (2012). Monetary policy with heterogeneous agents. *Federal Reserve Bank of Philadelphia, WP No. 12-21*.

Guerrieri, V. and Lorenzoni, G. (2017). Credit crises, precautionary savings, and the liquidity trap. *Quarterly Journal of Economics*, 132(3):1427–1467.

Kaplan, G., Moll, B., and Violante, G. L. (2017). Monetary policy according to HANK. *American Economic Review*, forthcoming.

Krusell, P. and Smith, A. A. (1997). Income and wealth heterogeneity, portfolio choice, and equilibrium asset returns. *Macroeconomic Dynamics*, 1(02):387–422.

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

Luetticke, R. (2018). Transmission of monetary policy with heterogeneity in household portfolios. *CFM Discussion Paper Series*, (CFM-DP2018-19).

McKay, A., Nakamura, E., and Steinsson, J. (2016). The power of forward guidance revisited. *American Economic Review*, 106(10):3133–3158.

McKay, A. and Reis, R. (2016). The role of automatic stabilizers in the US business cycle. *Econometrica*, 84(1):141–194.

Mitman, K. (2016). Macroeconomic effects of bankruptcy and foreclosure policies. *American Economic Review*, 106(8):2219–55.

Reiter, M. (2002). Recursive computation of heterogeneous agent models. *mimeo, Universitat Pompeu Fabra*.

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

Schmitt-Grohé, S. and Uribe, M. (2004). Solving dynamic general equilibrium models using a second-order approximation to the policy function. *Journal of Economic Dynamics and Control*, 28(4):755–775.

Winberry, T. (2016). A toolbox for solving and estimating heterogeneous agent macro models. *mimeo, Chicago Booth*.