Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models

Adrien Auclert, Bence Bardóczy, Matt Rognlie, Ludwig Straub

Computational Economics and Finance Remote Brown Bag, May 2020

This paper

Q: How to solve heterogeneous-agent (HA) models in GE with aggregate shocks?

This paper

Q: How to solve heterogeneous-agent (HA) models in GE with aggregate shocks?

- When idiosyncratic risk $\gg$ aggregate risk, two leading options:
  1. Linearize wrt aggregate shocks, solve **linear state space system** [Reiter method]
  2. Assume perfect foresight wrt aggregate shocks, solve **nonlinear system** in space of aggregate sequences (**sequence space**) [MIT shock method]

This paper

Q: How to solve heterogeneous-agent (HA) models in GE with aggregate shocks?

- When idiosyncratic risk $\gg$ aggregate risk, two leading options:

  1. Linearize wrt aggregate shocks, solve **linear state space system** [Reiter method]
  2. Assume perfect foresight wrt aggregate shocks, solve **nonlinear system** in space of aggregate sequences (**sequence space**) [MIT shock method]

  For small shocks, 1 $\Leftrightarrow$ 2 by certainty equivalence [Boppart, Krusell, Mitman]

This paper

Q: How to solve heterogeneous-agent (HA) models in GE with aggregate shocks?

- When idiosyncratic risk $\gg$ aggregate risk, two leading options:
  1. Linearize wrt aggregate shocks, solve **linear state space system** [Reiter method]
  2. Assume perfect foresight wrt aggregate shocks, solve **nonlinear system** in space of aggregate sequences (**sequence space**) [MIT shock method]

  For small shocks, 1 $\Leftrightarrow$ 2 by certainty equivalence [Boppart, Krusell, Mitman]

- **Here**: directly solve **linear system** in the **sequence space**: same, but faster!

This paper

Q: How to solve heterogeneous-agent (HA) models in GE with aggregate shocks?

- When idiosyncratic risk $\gg$ aggregate risk, two leading options:
  1. Linearize wrt aggregate shocks, solve **linear state space system** [Reiter method]
  2. Assume perfect foresight wrt aggregate shocks, solve **nonlinear system** in space of aggregate sequences (**sequence space**) [MIT shock method]

  For small shocks, 1 $\Leftrightarrow$ 2 by certainty equivalence [Boppart, Krusell, Mitman]
- **Here**: directly solve **linear system** in the **sequence space**: same, but faster!
- **Our method:** three steps

This paper

Q: How to solve heterogeneous-agent (HA) models in GE with aggregate shocks?

- When idiosyncratic risk $\gg$ aggregate risk, two leading options:
  1. Linearize wrt aggregate shocks, solve **linear state space system** [Reiter method]
  2. Assume perfect foresight wrt aggregate shocks, solve **nonlinear system** in space of aggregate sequences (**sequence space**) [MIT shock method]

  For small shocks, 1 $\Leftrightarrow$ 2 by certainty equivalence [Boppart, Krusell, Mitman]
- **Here**: directly solve **linear system** in the **sequence space**: same, but faster!
- **Our method:** three steps
  1. Write HA model as a collection of **blocks** along a **directed acyclic graph (DAG)**
  2. Compute the **Jacobian** of each block: key "**sufficient statistic**" for GE interactions
  3. Use Jacobians for: **IRFs, determinacy, full-info estimation, nonlinear transitions, ...**

Why is our method useful?

1. **Fast**: for state-of-the-art, two-asset HANK model,
   - First impulse response takes ∼**5s** (vs ∼**100s** with leading alternative methods)
   - Additional impulse responses take ∼**100ms** (vs **100s**) by **re-using Jacobians**
   - This makes **model estimation** possible
2. **Accurate**: no "model reduction" necessary, only error is from truncation
3. **Modular**: easy to build complex models by stitching blocks together
4. **Intuitive**: block Jacobians often have simple interpretation [eg MPCs]
5. **Accessible**: key steps automated in publicly available code [in Python]
   - Most ideas are also easily implemented in Matlab

Literature: our method combines several innovations

- Write equilibrium as linear system in aggregates
  [Reiter 2009, McKay and Reis 2016, Winberry 2018, Bayer, Luetticke, Pham-Dao and Tjaden 2019, Mongey and Williams 2017, Ahn, Kaplan, Moll, Winberry and Wolf 2018, ...]
  - $\rightarrow$ size of system now independent of underlying HA, no Schur decomposition that's costly for large state space
- Solve for impulse responses in sequence space
  [Auerbach and Kotlikoff 1987, Guerrieri and Lorenzoni 2017, McKay, Nakamura and Steinsson 2016, Kaplan, Moll and Violante 2018, Boppart, Krusell and Mitman 2018, ...]
  - $\rightarrow$ but now compute all in one go, no slowly-converging iteration
- Capture heterogeneity using GE sufficient statistics
  [Auclert and Rognlie 2018, Auclert, Rognlie and Straub 2018, Guren, McKay, Nakamura and Steinsson 2018, Koby and Wolf 2018, Wolf 2019]
  - $\rightarrow$ previously empirical or conceptual, now a computational tool

Roadmap

1. Models as collections of blocks arranged along a DAG
2. All you need are block Jacobians
3. Speeding up HA Jacobian computation

Models as collections of blocks arranged along a DAG

Introducing models as collections of blocks

- **Block:** Mapping from sequence of *inputs* to sequence of *outputs*

Introducing models as collections of blocks

- **Block:** Mapping from sequence of *inputs* to sequence of *outputs*

  Example 1: **heterogeneous household block** $\{r_t, w_t\} \to \{C_t\}$

  - Exogenous Markov chain for skills $\Pi(e'|e)$
  - Households

$$\max \mathbb{E}_0 \sum_t \beta^t u(c_{it})$$

$$c_{it} + k_{it} \leq (1 + r_t)k_{it-1} + w_t e_{it}$$

$$k_{it} \geq 0$$

Introducing models as collections of blocks

- **Block:** Mapping from sequence of *inputs* to sequence of *outputs*

  Example 1: **heterogeneous household block** $\{r_t, w_t\} \rightarrow \{C_t\}$
  - Exogenous Markov chain for skills $\Pi(e'|e)$
  - Households

$$\max \mathbb{E}_0 \sum_t \beta^t u(c_{it})$$

$$c_{it} + k_{it} \leq (1 + r_t) k_{it-1} + w_t e_{it}$$

$$k_{it} \geq 0$$

$\rightarrow$ Given initial distribution $D_0(e, k_-)$, path of aggregate consumption $C_t \equiv \int c_t(e, k_-) D_t(e, dk_-)$ only depends on $\{r_s, w_s\}_{s=0}^{\infty}$.
[Farhi-Werning 2017, Kaplan-Moll-Violante 2018, Auclert-Rognlie-Straub 2018]
(We'll assume $r_s = r$, $w_s = w$ for $s \geq T_0$.)

Introducing models as collections of blocks

- **Block:** Mapping from sequence of *inputs* to sequence of *outputs*

  Example 1: **heterogeneous household block** $\{r_t, w_t\} \to \{C_t\}$

  Example 2: **representative firm block** with $L = 1$ $\{K_t, Z_t\} \to \{Y_t, I_t, r_t, w_t\}$

$$
\begin{aligned}
Y_t &= Z_t K_{t-1}^{\alpha} \\
I_t &= K_t - (1-\delta) K_{t-1} \\
r_t &= \alpha Z_t K_{t-1}^{\alpha-1} - \delta \\
w_t &= (1-\alpha) Z_t K_{t-1}^{\alpha}
\end{aligned}
$$

$\to$ Given initial capital K_{-1}, path of $\{Y_t, I_t, r_t, w_t\}_{t=0}^{\infty}$ only depends on $\{K_s, Z_s\}_{s=0}^{\infty}$.

Introducing models as collections of blocks

- **Block:** Mapping from sequence of *inputs* to sequence of *outputs*

  Example 1: **heterogeneous household block** $\{r_t, w_t\} \to \{C_t\}$

  Example 2: **representative firm block** with $L = 1$ $\{K_t, Z_t\} \to \{Y_t, I_t, r_t, w_t\}$

  Example 3: **goods market clearing block** $\{Y_t, C_t, I_t\} \to \{H_t \equiv C_t + I_t - Y_t\}$

Introducing models as collections of blocks

- **Block:** Mapping from sequence of *inputs* to sequence of *outputs*

  Example 1: **heterogeneous household block** $\{r_t, w_t\} \to \{C_t\}$

  Example 2: **representative firm block** with $L = 1$ $\{K_t, Z_t\} \to \{Y_t, I_t, r_t, w_t\}$

  Example 3: **goods market clearing block** $\{Y_t, C_t, I_t\} \to \{H_t \equiv C_t + I_t - Y_t\}$

- **Model:** Set of blocks, arranged along a directed acyclic graph (DAG)
  - some inputs are exogenous **shocks**, e.g. $\{Z_t\}$
  - some inputs are endogenous **unknowns**, e.g. $\{K_t\}$
  - some outputs are **target** sequences that must equal zero in GE, e.g. $\{H_t\}$

    [must have as many targets as unknowns]

Introducing models as collections of blocks

- **Block:** Mapping from sequence of *inputs* to sequence of *outputs*

  Example 1: **heterogeneous household block** $\{r_t, w_t\} \to \{C_t\}$

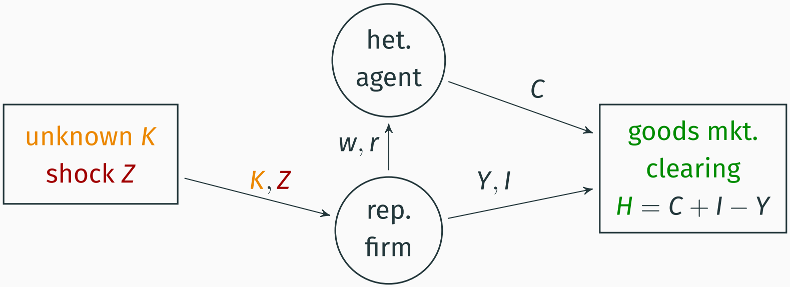
  Example 2: **representative firm block** with $L = 1$ $\{K_t, Z_t\} \to \{Y_t, I_t, r_t, w_t\}$

  Example 3: **goods market clearing block** $\{Y_t, C_t, I_t\} \to \{H_t \equiv C_t + I_t - Y_t\}$
- **Model:** Set of blocks, arranged along a directed acyclic graph (DAG)
  - some inputs are exogenous **shocks**, e.g. $\{Z_t\}$
  - some inputs are endogenous **unknowns**, e.g. $\{K_t\}$
  - some outputs are **target** sequences that must equal zero in GE, e.g. $\{H_t\}$

    [must have as many targets as unknowns]
- Many models can be written in this way.
  - Key restriction: agents interact via **limited set of aggregate variables**

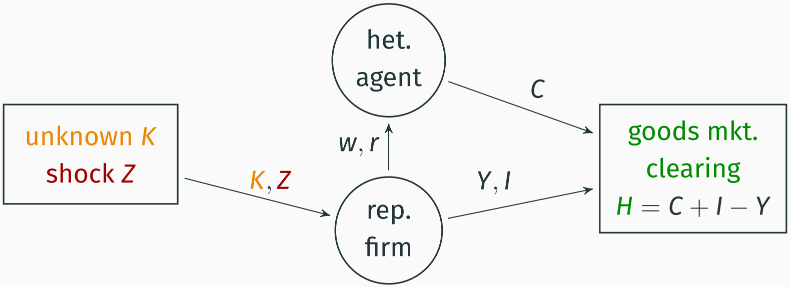
Krusell-Smith model DAG

Capital market clearing



Krusell-Smith model DAG

▸ Capital market clearing

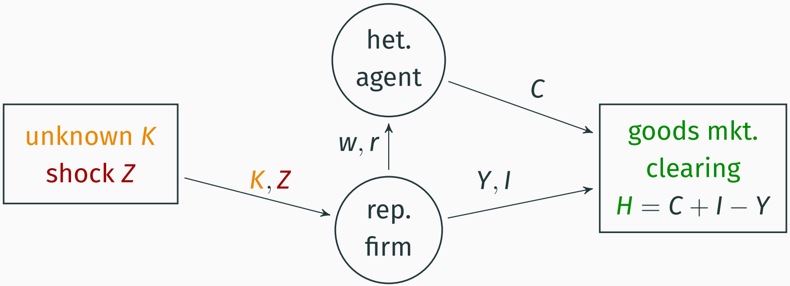


- DAG can be collapsed into mapping

$$H_t(\{K_s\}, \{Z_s\}) = C_t + I_t - Y_t$$

Krusell-Smith model DAG

Capital market clearing

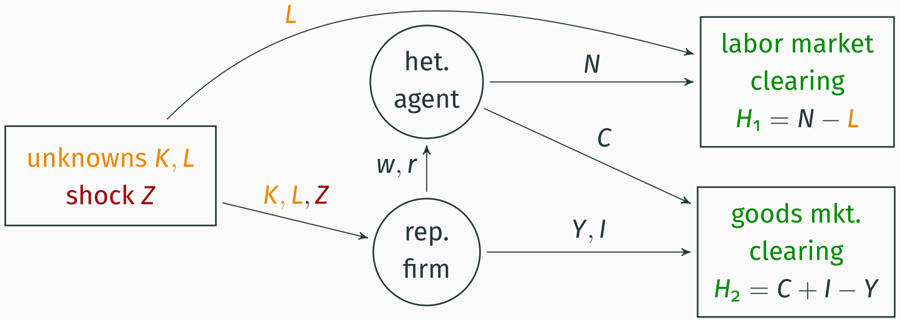


- DAG can be collapsed into mapping

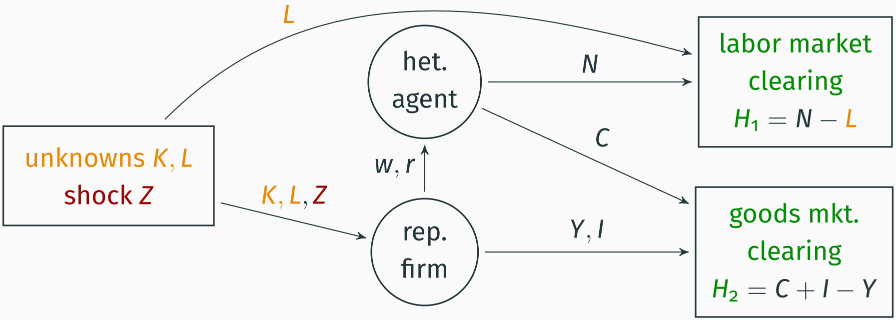
$$H_t\left(\{K_s\}, \{Z_s\}\right) = C_t + I_t - Y_t$$

- GE path of $\{K_s\}$ achieves $H_t\left(\{K_s\}, \{Z_s\}\right) = 0$

Dealing with endogenous labor: add an unknown and a target



Dealing with endogenous labor: add an unknown and a target

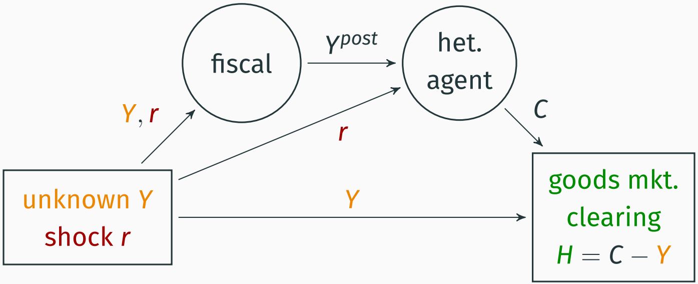


- DAG can be collapsed into mapping

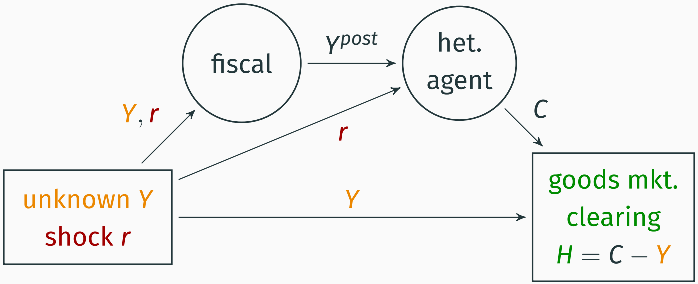
$$\mathbf{H}_t(\{K_s, L_s\}, \{Z_s\}) = \{C_t + I_t - Y_t, N_s - L_s\}$$

- GE path of $\{K_s, L_s\}$ achieves $\mathbf{H}_t(\{K_s, L_s\}, \{Z_s\}) = 0$

Simple one-asset HANK model with sticky wages: another DAG with one unknown



Simple one-asset HANK model with sticky wages: another DAG with one unknown

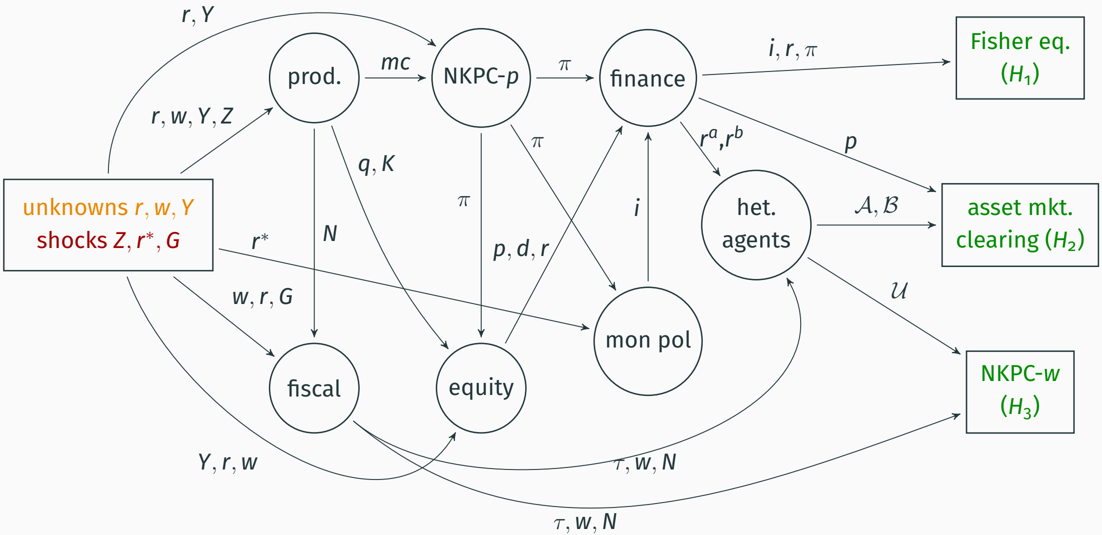


- DAG can be collapsed into mapping

$$H_t(\{Y_s\}, \{r_s\}) = C_t - Y_t$$

- GE path of $\{Y_s\}$ achieves $H_t(\{Y_s\}, \{r_s\}) = 0$

Two-asset HANK model in paper: richer DAG with three unknowns



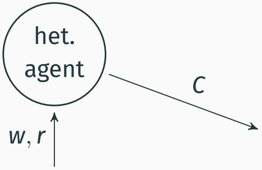
All you need are block Jacobians

Block Jacobians

- Suppose we have set the DAG and initial conditions [typically the steady state]
- Define a block **Jacobian** as the derivatives of its outputs wrt its inputs

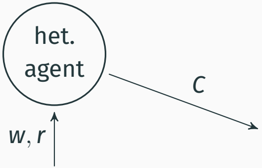
Block Jacobians

- Suppose we have set the DAG and initial conditions [typically the steady state]
- Define a block **Jacobian** as the derivatives of its outputs wrt its inputs
  - e.g. household block



Block Jacobians

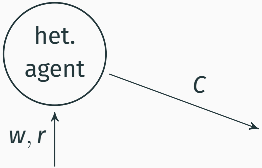
- Suppose we have set the DAG and initial conditions [typically the steady state]
- Define a block **Jacobian** as the derivatives of its outputs wrt its inputs
  - e.g. household block



$\rightarrow$ two Jacobians: $\mathcal{J}^{C,w}_{t,s} \equiv \frac{\partial C_t}{\partial w_s}$ [iMPCs, Auclert-Rognlie-Straub] and $\mathcal{J}^{C,r}_{t,s} \equiv \frac{\partial C_t}{\partial r_s}$

Block Jacobians

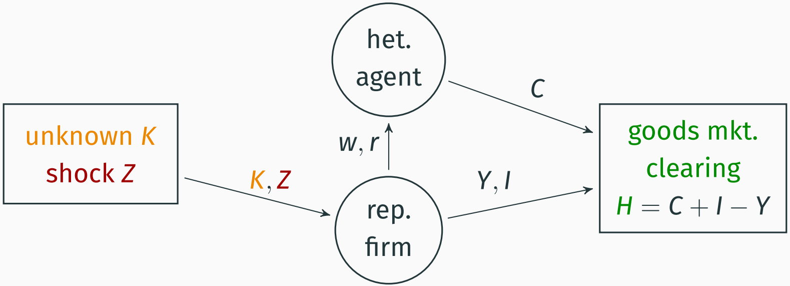
- Suppose we have set the DAG and initial conditions [typically the steady state]
- Define a block **Jacobian** as the derivatives of its outputs wrt its inputs
  - e.g. household block



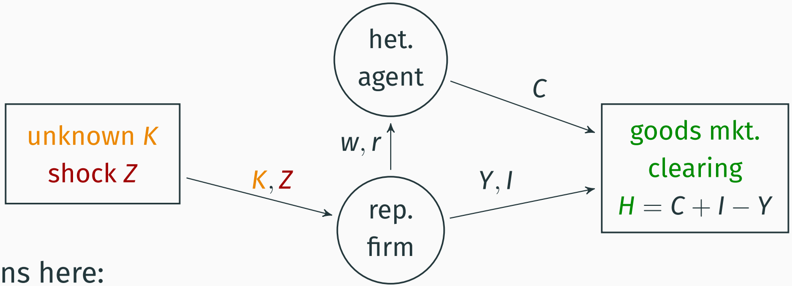
$\rightarrow$ two Jacobians: $\mathcal{J}^{C,w}_{t,s} \equiv \frac{\partial C_t}{\partial w_s}$ [iMPCs, Auclert-Rognlie-Straub] and $\mathcal{J}^{C,r}_{t,s} \equiv \frac{\partial C_t}{\partial r_s}$

- Next: block Jacobians are **sufficient to compute GE impulse responses**

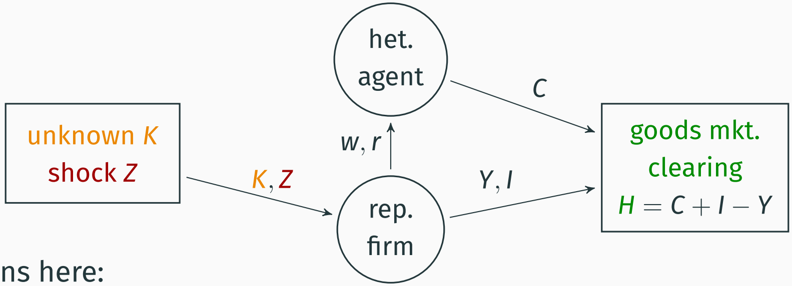
Krusell-Smith model Jacobians



Krusell-Smith model Jacobians

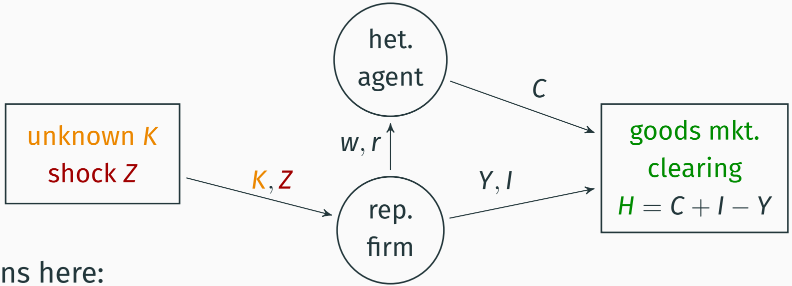


- Jacobians here:
  - het. agent: $\left\{\frac{\partial C_t}{\partial w_s}\right\}, \left\{\frac{\partial C_t}{\partial r_s}\right\} \rightsquigarrow$ denote $\mathcal{J}^{C,w}, \mathcal{J}^{C,r}$



- Jacobians here:
  - het. agent: $\left\{\frac{\partial C_t}{\partial w_s}\right\}, \left\{\frac{\partial C_t}{\partial r_s}\right\} \rightsquigarrow$ denote $\mathcal{J}^{C,w}, \mathcal{J}^{C,r}$
  - rep. firm: $\left\{\frac{\partial w_t}{\partial K_s}\right\}, \left\{\frac{\partial w_t}{\partial Z_s}\right\}, \left\{\frac{\partial r_t}{\partial K_s}\right\}, \left\{\frac{\partial r_t}{\partial Z_s}\right\}, \ldots \rightsquigarrow$ denote $\mathcal{J}^{w,K}, \mathcal{J}^{w,Z}, \mathcal{J}^{r,K}, \mathcal{J}^{r,Z}, \ldots$

Krusell-Smith model Jacobians



- Jacobians here:
  - het. agent: $\left\{\frac{\partial C_t}{\partial w_s}\right\}, \left\{\frac{\partial C_t}{\partial r_s}\right\} \rightsquigarrow$ denote $\mathcal{J}^{C,w}, \mathcal{J}^{C,r}$
  - rep. firm: $\left\{\frac{\partial w_t}{\partial K_s}\right\}, \left\{\frac{\partial w_t}{\partial Z_s}\right\}, \left\{\frac{\partial r_t}{\partial K_s}\right\}, \left\{\frac{\partial r_t}{\partial Z_s}\right\}, \ldots \rightsquigarrow$ denote $\mathcal{J}^{w,K}, \mathcal{J}^{w,Z}, \mathcal{J}^{r,K}, \mathcal{J}^{r,Z}, \ldots$
- We can then **chain the Jacobians along the DAG** to get the Jacobians of **H**:

$$\frac{\partial \mathbf{H}}{\partial \mathbf{K}} = \mathcal{J}^{C,r}\mathcal{J}^{r,K} + \mathcal{J}^{C,w}\mathcal{J}^{w,K} + \mathcal{J}^{I,K} - \mathcal{J}^{Y,K} \qquad \frac{\partial \mathbf{H}}{\partial \mathbf{Z}} = \mathcal{J}^{C,r}\mathcal{J}^{r,Z} + \mathcal{J}^{C,w}\mathcal{J}^{w,Z} + \mathcal{J}^{I,Z} - \mathcal{J}^{Y,Z}$$

From block Jacobians to impulse responses

Once Jacobians are chained to give $\frac{\partial \mathbf{H}}{\partial \mathbf{K}}$ and $\frac{\partial \mathbf{H}}{\partial \mathbf{Z}}$, we are done:

Suppose shock is $d\mathbf{Z} = \{dZ_t\}$ [with $dZ_t = 0$, $t \geq T_0$], what are the impulse responses?

From block Jacobians to impulse responses

Once Jacobians are chained to give $\frac{\partial \mathbf{H}}{\partial \mathbf{K}}$ and $\frac{\partial \mathbf{H}}{\partial \mathbf{Z}}$, we are done:

Suppose shock is $d\mathbf{Z} = \{dZ_t\}$ [with $dZ_t = 0$, $t \geq T_0$], what are the impulse responses?

1. $\mathbf{H}(\mathbf{K}, \mathbf{Z}) = 0$ after the shock

$$\frac{\partial \mathbf{H}}{\partial \mathbf{K}} d\mathbf{K} + \frac{\partial \mathbf{H}}{\partial \mathbf{Z}} d\mathbf{Z} = 0$$

From block Jacobians to impulse responses

Once Jacobians are chained to give $\frac{\partial \mathbf{H}}{\partial \mathbf{K}}$ and $\frac{\partial \mathbf{H}}{\partial \mathbf{Z}}$, we are done:

Suppose shock is $d\mathbf{Z} = \{dZ_t\}$ [with $dZ_t = 0$, $t \geq T_0$], what are the impulse responses?

1. $\mathbf{H}(\mathbf{K}, \mathbf{Z}) = 0$ after the shock. Solve for unknown $d\mathbf{K} \Rightarrow$

$$d\mathbf{K} = -\left(\frac{\partial \mathbf{H}}{\partial \mathbf{K}}\right)^{-1} \frac{\partial \mathbf{H}}{\partial \mathbf{Z}} d\mathbf{Z}$$

From block Jacobians to impulse responses

Once Jacobians are chained to give $\frac{\partial \mathbf{H}}{\partial \mathbf{K}}$ and $\frac{\partial \mathbf{H}}{\partial \mathbf{Z}}$, we are done:

Suppose shock is $d\mathbf{Z} = \{dZ_t\}$ [with $dZ_t = 0$, $t \geq T_0$], what are the impulse responses?

1. $\mathbf{H}(\mathbf{K}, \mathbf{Z}) = 0$ after the shock. Solve for unknown $d\mathbf{K} \Rightarrow$

$$d\mathbf{K} = -\left(\frac{\partial \mathbf{H}}{\partial \mathbf{K}}\right)^{-1} \frac{\partial \mathbf{H}}{\partial \mathbf{Z}} d\mathbf{Z}$$

2. Use Jacobians to back out any IRF of interest, e.g. IRF of output

$$d\mathbf{Y} = \mathcal{J}^{Y,K} d\mathbf{K} + \mathcal{J}^{Y,Z} d\mathbf{Z}$$

From block Jacobians to impulse responses

Once Jacobians are chained to give $\frac{\partial \mathbf{H}}{\partial \mathbf{K}}$ and $\frac{\partial \mathbf{H}}{\partial \mathbf{Z}}$, we are done:

Suppose shock is $d\mathbf{Z} = \{dZ_t\}$ [with $dZ_t = 0$, $t \geq T_0$], what are the impulse responses?

1. $\mathbf{H}(\mathbf{K}, \mathbf{Z}) = 0$ after the shock. Solve for unknown $d\mathbf{K} \Rightarrow$

$$d\mathbf{K} = -\left(\frac{\partial \mathbf{H}}{\partial \mathbf{K}}\right)^{-1} \frac{\partial \mathbf{H}}{\partial \mathbf{Z}} d\mathbf{Z}$$

2. Use Jacobians to back out any IRF of interest, e.g. IRF of output

$$d\mathbf{Y} = \mathcal{J}^{Y,K} d\mathbf{K} + \mathcal{J}^{Y,Z} d\mathbf{Z}$$

$\Rightarrow$ **Block Jacobians are sufficient to obtain all GE impulse responses**

Can also compute moments of the distribution $D_t(e, k_-)$ this way

From block Jacobians to impulse responses

Once Jacobians are chained to give $\frac{\partial \mathbf{H}}{\partial \mathbf{K}}$ and $\frac{\partial \mathbf{H}}{\partial \mathbf{Z}}$, we are done:

Suppose shock is $d\mathbf{Z} = \{dZ_t\}$ [with $dZ_t = 0$, $t \geq T_0$], what are the impulse responses?

1. $\mathbf{H}(\mathbf{K}, \mathbf{Z}) = 0$ after the shock. Solve for unknown $d\mathbf{K} \Rightarrow$

$$d\mathbf{K} = -\left(\frac{\partial \mathbf{H}}{\partial \mathbf{K}}\right)^{-1} \frac{\partial \mathbf{H}}{\partial \mathbf{Z}} d\mathbf{Z}$$

2. Use Jacobians to back out any IRF of interest, e.g. IRF of output

$$d\mathbf{Y} = \mathcal{J}^{Y,K} d\mathbf{K} + \mathcal{J}^{Y,Z} d\mathbf{Z}$$

$\Rightarrow$ **Block Jacobians are sufficient to obtain all GE impulse responses**

Can also compute moments of the distribution $D_t(e, k_-)$ this way

[in paper: generalize using **automatic differentiation** along the DAG]

Aggregate shocks and $MA(\infty)$ representation

- Certainty equivalence $\Rightarrow d\mathbf{K}$ is also the $MA(\infty)$ representation in model with aggregate shocks:

Aggregate shocks and $MA(\infty)$ representation

- Certainty equivalence $\Rightarrow$ $d\mathbf{K}$ is also the $MA(\infty)$ representation in model with aggregate shocks:
- Suppose $\{d\tilde{Z}_t\}$ is $MA(\infty)$ in iid structural innovation vectors $\{\epsilon_t\}$:

$$d\tilde{Z}_t = \sum_{s=0}^{\infty} dZ_s \epsilon_{t-s}$$

then

$$d\tilde{K}_t = \sum_{s=0}^{\infty} dK_s \epsilon_{t-s}$$

Aggregate shocks and $MA(\infty)$ representation

- Certainty equivalence $\Rightarrow$ $d\mathbf{K}$ is also the $MA(\infty)$ representation in model with aggregate shocks:
- Suppose $\{d\tilde{Z}_t\}$ is $MA(\infty)$ in iid structural innovation vectors $\{\epsilon_t\}$:

$$d\tilde{Z}_t = \sum_{s=0}^{\infty} dZ_s \epsilon_{t-s}$$

then

$$d\tilde{K}_t = \sum_{s=0}^{\infty} dK_s \epsilon_{t-s}$$

→ Applications:

1. Simulation method (immediate)
2. Analytical second moments for any X, Y: $\text{Cov}(d\tilde{X}_t, d\tilde{Y}_{t'}) = \sigma_\epsilon^2 \sum_{s=0}^{T-(t'-t)} dX_s dY_{s+t'-t}$
3. Estimation (**next**)

Estimation

- Let $\mathbf{V}(\theta)$ be the covariance matrix for a set of k outputs, where $\theta \equiv$ parameters
- Assuming Gaussian innovations, log-likelihood of observed data $\mathbf{Y}$ given θ:

$$\mathcal{L}(\mathbf{Y};\theta) = -\frac{1}{2}\log\det\mathbf{V}(\theta) - \frac{1}{2}\mathbf{Y}'\mathbf{V}(\theta)^{-1}\mathbf{Y}$$

Estimation

- Let $\mathbf{V}(\theta)$ be the covariance matrix for a set of k outputs, where $\theta \equiv$ parameters
- Assuming Gaussian innovations, log-likelihood of observed data $\mathbf{Y}$ given θ:

$$\mathcal{L}(\mathbf{Y};\theta) = -\frac{1}{2}\log\det\mathbf{V}(\theta) - \frac{1}{2}\mathbf{Y}'\mathbf{V}(\theta)^{-1}\mathbf{Y}$$

- No need for Kalman filter! Old estimation strategy in time series.
  - several recent revivals in DSGE [e.g. Mankiw and Reis 2007]
  - [in practice: use Cholesky or Levinson on $\mathbf{V}$, or Whittle approx when T is large]
  - first application to het agents, perfectly suited for sequence-space methods

Estimation

- Let $\mathbf{V}(\theta)$ be the covariance matrix for a set of k outputs, where $\theta \equiv$ parameters
- Assuming Gaussian innovations, log-likelihood of observed data $\mathbf{Y}$ given θ:

$$\mathcal{L}(\mathbf{Y};\theta) = -\frac{1}{2}\log\det\mathbf{V}(\theta) - \frac{1}{2}\mathbf{Y}'\mathbf{V}(\theta)^{-1}\mathbf{Y}$$

- No need for Kalman filter! Old estimation strategy in time series.
  - several recent revivals in DSGE [e.g. Mankiw and Reis 2007]
  - [in practice: use Cholesky or Levinson on $\mathbf{V}$, or Whittle approx when T is large]
  - first application to het agents, perfectly suited for sequence-space methods
- Estimating shock processes $d\mathbf{Z}$ almost free: use same Jacobians for any $d\mathbf{Z}$!

Estimation

- Let $\mathbf{V}(\theta)$ be the covariance matrix for a set of k outputs, where $\theta \equiv$ parameters
- Assuming Gaussian innovations, log-likelihood of observed data $\mathbf{Y}$ given θ:

$$\mathcal{L}(\mathbf{Y};\theta) = -\frac{1}{2}\log\det \mathbf{V}(\theta) - \frac{1}{2}\mathbf{Y}'\mathbf{V}(\theta)^{-1}\mathbf{Y}$$

- No need for Kalman filter! Old estimation strategy in time series.
  - several recent revivals in DSGE [e.g. Mankiw and Reis 2007]
  - [in practice: use Cholesky or Levinson on **V**, or Whittle approx when T is large]
  - first application to het agents, perfectly suited for sequence-space methods
- Estimating shock processes $d\mathbf{Z}$ almost free: use same Jacobians for any $d\mathbf{Z}$!
- Other estimation still **very fast** as long as we don't need to recalculate HA s.s.
  [eg, cap. adjustment costs, degree of price stickiness, ...]
  $\rightarrow$ can use the same HA Jacobians $\mathcal{J}^{C,w}$, $\mathcal{J}^{C,r}$, etc.

Remarks

▸ Determinacy ▸ Nonlinear transitions

1. In practice, our method involves the inversion of $nT \times nT$ matrix $\frac{\partial \mathbf{H}}{\partial \mathbf{K}}$, where n = # unknowns and T = truncation horizon [typically $T \simeq$ 300-500]
   - very fast as long as DAG doesn't have too many unknowns
   - key benefit of DAGs: reduce n without any loss in accuracy [typically $n \leq 3$]
   - in practice, choice of T depends on persistence of exogenous variables
2. This matrix is invertible if the model is locally **determinate**
   - simple test based on the winding number criterion of Onatski (2006) [see paper]
3. Jacobians are also useful to get the **nonlinear perfect-foresight** solution
   - Solve $\mathbf{H}(\mathbf{K}, \mathbf{Z}) = 0$ using Newton's method with s.s. Jacobian $\frac{\partial \mathbf{H}}{\partial \mathbf{K}}$ [see paper]

1. In practice, our method involves the inversion of $nT \times nT$ matrix $\frac{\partial \mathbf{H}}{\partial \mathbf{K}}$, where n = # unknowns and T = truncation horizon [typically $T \simeq$ 300-500]
   - very fast as long as DAG doesn't have too many unknowns
   - key benefit of DAGs: reduce n without any loss in accuracy [typically $n \leq 3$]
   - in practice, choice of T depends on persistence of exogenous variables
2. This matrix is invertible if the model is locally **determinate**
   - simple test based on the winding number criterion of Onatski (2006) [see paper]
3. Jacobians are also useful to get the **nonlinear perfect-foresight** solution
   - Solve $\mathbf{H}(\mathbf{K}, \mathbf{Z}) = 0$ using Newton's method with s.s. Jacobian $\frac{\partial \mathbf{H}}{\partial \mathbf{K}}$ [see paper]

Next: how to rapidly compute the Jacobians of heterogeneous-agent blocks

Speeding up HA Jacobian computation

Computing heterogeneous-agent Jacobians

So far: DAG + Jacobians $\Rightarrow$ IRFs, determinacy, estimation, nonlinear transitions

But how do we get the block **Jacobians?**

Computing heterogeneous-agent Jacobians

So far: DAG + Jacobians $\Rightarrow$ IRFs, determinacy, estimation, nonlinear transitions

But how do we get the block **Jacobians?**

- *simple blocks*: (e.g. representative firms) simple, sparse matrix

Computing heterogeneous-agent Jacobians

So far: DAG + Jacobians $\Rightarrow$ IRFs, determinacy, estimation, nonlinear transitions

But how do we get the block **Jacobians?**

- *simple blocks*: (e.g. representative firms) simple, sparse matrix
- *HA blocks?* $\to$ **next**

Jacobian of consumption with respect to wage

- Want to know $\mathcal{J}_{t,s} \equiv \frac{\partial C_t}{\partial w_s}$ for $s, t \in \{0, \ldots, T-1\}$ [intertemporal MPCs]
    - Assume initial condition is s.s., with $r_t = r$, $w_t = w$, $D_0(e, k_-) = D(e, k_-)$

Jacobian of consumption with respect to wage

- Want to know $\mathcal{J}_{t,s} \equiv \frac{\partial C_t}{\partial w_s}$ for $s, t \in \{0, \ldots, T-1\}$ [intertemporal MPCs]
  - Assume initial condition is s.s., with $r_t = r$, $w_t = w$, $D_0(e, k_-) = D(e, k_-)$
- **Direct algorithm**: perturb $w_s \equiv w + \epsilon$

Jacobian of consumption with respect to wage

- Want to know $\mathcal{J}_{t,s} \equiv \frac{\partial C_t}{\partial w_s}$ for $s, t \in \{0, \ldots, T-1\}$ [intertemporal MPCs]
    - Assume initial condition is s.s., with $r_t = r$, $w_t = w$, $D_0(e, k_-) = D(e, k_-)$

- **Direct algorithm**: perturb $w_s \equiv w + \epsilon$
    1. iterate backwards to get perturbed policies: $\mathbf{c}_t^s(e, k_-), \mathbf{k}_t^s(e, k_-)$
    2. iterate forward to get perturbed distributions $D_t^s(e, k_-)$
    3. put together to get perturbed aggregate consumption: $C_t^s = \int \mathbf{c}_t^s(e, k_-) D_t^s(e, dk_-)$
    4. compute $\mathcal{J}$ from $\mathcal{J}_{t,s} \equiv (C_t^s - C)/\epsilon$

Jacobian of consumption with respect to wage

- Want to know $\mathcal{J}_{t,s} \equiv \frac{\partial C_t}{\partial w_s}$ for $s, t \in \{0, \ldots, T-1\}$ [intertemporal MPCs]
  - Assume initial condition is s.s., with $r_t = r$, $w_t = w$, $D_0(e, k_-) = D(e, k_-)$
- **Direct algorithm**: perturb $w_s \equiv w + \epsilon$
  1. iterate backwards to get perturbed policies: $\mathbf{c}_t^s(e, k_-), \mathbf{k}_t^s(e, k_-)$
  2. iterate forward to get perturbed distributions $D_t^s(e, k_-)$
  3. put together to get perturbed aggregate consumption: $C_t^s = \int \mathbf{c}_t^s(e, k_-) D_t^s(e, dk_-)$
  4. compute $\mathcal{J}$ from $\mathcal{J}_{t,s} \equiv (C_t^s - C)/\epsilon$
- This is **slow**, since 1–4 needs to be done T times, once for each s

Jacobian of consumption with respect to wage

- Want to know $\mathcal{J}_{t,s} \equiv \frac{\partial C_t}{\partial w_s}$ for $s, t \in \{0, \ldots, T-1\}$ [intertemporal MPCs]
  - Assume initial condition is s.s., with $r_t = r$, $w_t = w$, $D_0(e, k_-) = D(e, k_-)$

- **Direct algorithm**: perturb $w_s \equiv w + \epsilon$
  1. iterate backwards to get perturbed policies: $\mathbf{c}_t^s(e, k_-), \mathbf{k}_t^s(e, k_-)$
  2. iterate forward to get perturbed distributions $D_t^s(e, k_-)$
  3. put together to get perturbed aggregate consumption: $C_t^s = \int \mathbf{c}_t^s(e, k_-) D_t^s(e, dk_-)$
  4. compute $\mathcal{J}$ from $\mathcal{J}_{t,s} \equiv (C_t^s - C)/\epsilon$

- This is **slow**, since 1–4 needs to be done T times, once for each s

- Paper proposes **fake news algorithm** that is T times faster:
  - requires **single** backward iteration & **single** forward iteration
  - key idea: exploit **time symmetries** around the steady-state

(0) The fake news matrix

- We can think of $\mathcal{J} \equiv \left(\frac{\partial C_t}{\partial w_s}\right)$ as a **news matrix**
  - column $s =$ response to news that shock hits in period s
- Define a new auxiliary matrix:

$$\mathcal{F}_{t,s} \equiv \begin{cases} \frac{\partial C_t}{\partial w_s} & s = 0 \text{ or } t = 0 \\ \frac{\partial C_t}{\partial w_s} - \frac{\partial C_{t-1}}{\partial w_{s-1}} & s, t > 0 \end{cases}$$

- Can think of this as **fake news matrix:**
  - at $t = 0$: news shock that period s shock hits $\rightarrow \frac{\partial C_0}{\partial w_s}$
  - at $t = 1$: news shock that there won't be a shock at $s \rightarrow \frac{\partial C_1}{\partial w_s} - \frac{\partial C_0}{\partial w_{s-1}}$
  - useful: starting in $t = 1$, **agents' policy functions are unchanged** by fake news shock
- Can recover $\mathcal{J}$ from $\mathcal{F}$: news shock = sequence of fake news shocks

(o) The fake news matrix

$$\mathcal{J} = \begin{pmatrix} \mathcal{J}_{00} & \mathcal{J}_{01} & \mathcal{J}_{02} & \cdots \\ \mathcal{J}_{10} & \mathcal{J}_{11} & \mathcal{J}_{12} & \cdots \\ \mathcal{J}_{20} & \mathcal{J}_{12} & \mathcal{J}_{22} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \qquad \mathcal{F} = \begin{pmatrix} \mathcal{J}_{00} & \mathcal{J}_{01} & \mathcal{J}_{02} & \cdots \\ \mathcal{J}_{10} & \mathcal{J}_{11} - \mathcal{J}_{00} & \mathcal{J}_{12} - \mathcal{J}_{01} & \cdots \\ \mathcal{J}_{20} & \mathcal{J}_{12} - \mathcal{J}_{10} & \mathcal{J}_{22} - \mathcal{J}_{11} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

- Can recover $\mathcal{J}$ from $\mathcal{F}$ by adding elements from top left diagonal

(1) Single backward iteration

- **Claim:** Single backward iteration is enough to recover $\mathbf{c}_t^s(e, k_-), \mathbf{k}_t^s(e, k_-)$

(1) Single backward iteration

- **Claim:** Single backward iteration is enough to recover $\mathbf{c}_t^s(e, k_-), \mathbf{k}_t^s(e, k_-)$
- Why? only the **time** $s - t$ **until the perturbation matters**

$$\mathbf{c}_t^s(e, k_-) = \begin{cases} \mathbf{c}(e, k_-) & s < t \\ \mathbf{c}_{T-1-(s-t)}^{T-1}(e, k_-) & s \geq t \end{cases}$$

- Thus, only need a single backward iteration with $s = T - 1$ to get all the $\mathbf{c}_t^s$

(1) Single backward iteration

- **Claim:** Single backward iteration is enough to recover $\mathbf{c}_t^s(e, k_-), \mathbf{k}_t^s(e, k_-)$
- Why? only the **time** $s - t$ **until the perturbation matters**

$$\mathbf{c}_t^s(e, k_-) = \begin{cases} \mathbf{c}(e, k_-) & s < t \\ \mathbf{c}_{T-1-(s-t)}^{T-1}(e, k_-) & s \geq t \end{cases}$$

- Thus, only need a single backward iteration with $s = T - 1$ to get all the $\mathbf{c}_t^s$
- From these we get:
  - $C_0^s = \int \mathbf{c}_0^s(e, k_-) D(e, dk_-)$, so first row of Jacobian $\mathcal{J}_{0s} = \frac{\partial C_0}{\partial w_s} = \mathcal{F}_{0s}$
  - $D_1^s(e, dk_-)$, distributions at date 1 implied by new policy $\mathbf{c}_0^s$ at date 0

(2) Single forward iteration

- Let's iterate those distributions forward using **s.s. policies**

$$D_1^s(e, dk_-) \mapsto D_2^s(e, dk_-) \mapsto D_3^s(e, dk_-) \mapsto \dots$$

  - this is just a **linear map**: $\mathbf{D}_t^s = (\Lambda')^{t-1}\,\mathbf{D}_1^s$ where Λ is s.s. transition matrix

(2) Single forward iteration

- Let's iterate those distributions forward using **s.s. policies**

$$D_1^s(e, dk_-) \mapsto D_2^s(e, dk_-) \mapsto D_3^s(e, dk_-) \mapsto \dots$$

  - this is just a **linear map**: $\mathbf{D}_t^s = (\Lambda')^{t-1}\,\mathbf{D}_1^s$ where Λ is s.s. transition matrix

- Now construct aggregate consumption using **s.s. policies c**

$$C_t^s \equiv \int \mathbf{c}(e, k_-)\, D_t^s(e, dk_-) \quad \Rightarrow \quad C_t^s = \mathbf{c}'\,(\Lambda')^{t-1}\,\mathbf{D}_1^s$$

  - this only requires computing $\mathbf{c}'$, $\mathbf{c}'\Lambda'$, $\mathbf{c}'\,(\Lambda')^2, \dots \rightarrow$ like a **single** forward iteration!

(2) Single forward iteration

- Let's iterate those distributions forward using **s.s. policies**

$$D_1^s(e, dk_-) \mapsto D_2^s(e, dk_-) \mapsto D_3^s(e, dk_-) \mapsto \dots$$

  - this is just a **linear map**: $\mathbf{D}_t^s = (\Lambda')^{t-1}\mathbf{D}_1^s$ where Λ is s.s. transition matrix

- Now construct aggregate consumption using **s.s. policies c**

$$C_t^s \equiv \int \mathbf{c}(e, k_-)\, D_t^s(e, dk_-) \quad \Rightarrow \quad C_t^s = \mathbf{c}' \left(\Lambda'\right)^{t-1} \mathbf{D}_1^s$$

  - this only requires computing $\mathbf{c}'$, $\mathbf{c}'\Lambda'$, $\mathbf{c}'\left(\Lambda'\right)^2, \dots \rightarrow$ like a **single** forward iteration!

- This is exactly the **fake news matrix**

$$\mathcal{F}_{t,s} = (C_t^s - C)/\epsilon$$

How long does this take?

Algorithm	Krusell-Smith	HD Krusell-Smith	one-asset HANK	two-asset HANK
Direct	**26 s**	**1939 s**	**176 s**	**2107 s**
step 1 (backward)	16 s	1338 s	150 s	1291 s
step 2 (forward)	10 s	601 s	27 s	815 s
Fake news	**0.104 s**	**8.429 s**	**0.646 s**	**5.697 s**
step 1 (backward)	0.067 s	5.433 s	0.525 s	5.206 s
step 2 (forward)	0.010 s	1.546 s	0.021 s	0.122 s
step 3	0.023 s	1.445 s	0.092 s	0.346 s
step 4	0.004 s	0.004 s	0.008 s	0.023 s
Gridpoints n_g	3,500	250,000	3,500	10,500

Conclusion

What we do in this paper:

Computing times for:	Krusell-Smith	HD Krusell-Smith	one-asset HANK	two-asset HANK
Heterogeneous-agent Jacobians	0.10 s	8.4 s	0.65 s	5.7 s
One impulse response	0.0012 s	0.0012 s	0.017 s	0.120 s
All impulse responses	0.0068 s	0.0068 s	0.097 s	0.400 s
Bayesian estimation (shocks)				
single likelihood evaluation	0.00088 s	0.00088 s	0.0021 s	0.058 s
entire estimation	0.12 s	0.12 s	0.50 s	21 s
Bayesian estimation (shocks + model)				
single likelihood evaluation	—	—	0.011 s	0.18 s
entire estimation	—	—	16 s	570 s
Determinacy test	252 μs	252 μs	631 μs	631 μs
Nonlinear impulse responses	0.18 s	13.76 s	0.96 s	27 s

Conclusion

- New method to **simulate, estimate & analyze** HA models

    1. model as **collection of blocks**
    2. **block Jacobians** as **sufficient statistics** for GE
    3. **fast & accurate:** IRFs, determinacy, full-info estimation, nonlinear transitions

Conclusion

- New method to **simulate, estimate & analyze** HA models

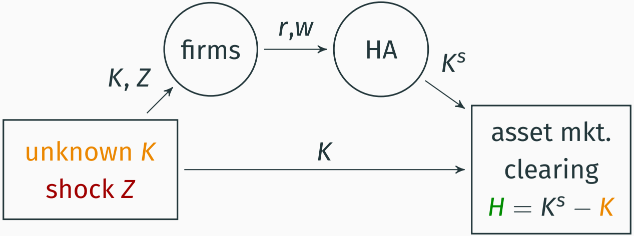
  1. model as **collection of blocks**
  2. **block Jacobians** as **sufficient statistics** for GE
  3. **fast & accurate:** IRFs, determinacy, full-info estimation, nonlinear transitions

https://github.com/shade-econ/sequence-jacobian

Comments welcome!

Krusell-Smith model DAG alternative

Back

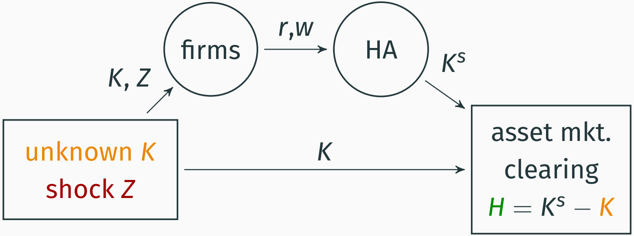


- By Walras's law, alternative target is capital market clearing:

$$H_t\left(\{K_s\}, \{Z_s\}\right) = K_t^s - K_t$$

Krusell-Smith model DAG alternative

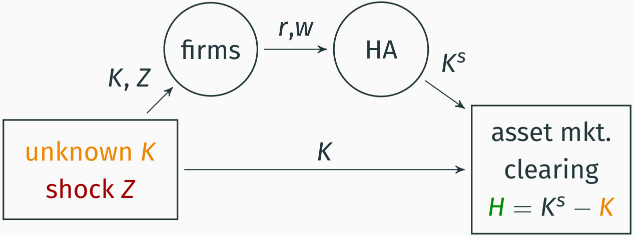
Back



- By Walras's law, alternative target is capital market clearing:

$$H_t\left(\{K_s\}, \{Z_s\}\right) = K_t^s - K_t$$

Krusell-Smith model DAG alternative



- By Walras's law, alternative target is capital market clearing:

$$H_t\left(\{K_s\},\{Z_s\}\right) = K_t^s - K_t$$

- GE path of $\{K_s\}$ achieves $H_t\left(\{K_s\},\{Z_s\}\right) = 0 \Rightarrow$ same solution as above.

Determinacy

◂ Back

- In state space, have e.g. Blanchard-Kahn: count stable roots
  - What analogue in sequence space?
  - Could test singularity of $\mathbf{H}_U$: works, but slow and imprecise
- Asymptotic time invariance for the Jacobians of SHADE models:

$$[\mathbf{H}_U]_{t,s} \to A_{t-s} \quad \text{as } t, s \to \infty$$

- **Winding number criterion**: precise and fast
- **Local determinacy** for generic model if winding number of

$$\det A(\lambda) \equiv \det \sum A_j e^{ij\lambda}; \qquad \lambda \in [0, 2\pi]$$

  around the origin is zero
  - Generalizes criterion for exactly time invariant models [Onatski 2006]
  - Given As, sample many λ and test in less than 1 ms using FFT

Nonlinear perfect foresight transitions

- Given Jacobian $\frac{\partial \mathbf{H}}{\partial \mathbf{K}}$, can compute full nonlinear solution to

$$H(\mathbf{K}, \mathbf{Z}) = 0$$

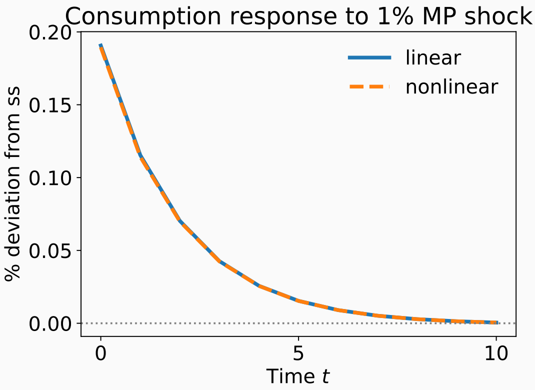
- Idea: use (quasi)-**Newton method**
- Start from $\mathbf{K}^{(0)} = \mathbf{K}_{ss}$ and iterate using

$$\mathbf{K}^{(n)} = \mathbf{K}^{(n-1)} - \left(\frac{\partial \mathbf{H}}{\partial \mathbf{K}}\right)^{-1} H\left(\mathbf{K}^{(n-1)}, \mathbf{Z}\right)$$

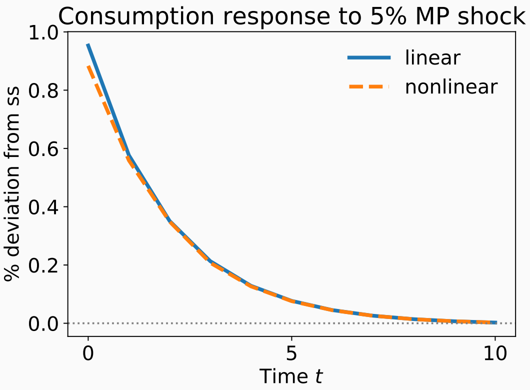
  where $\frac{\partial \mathbf{H}}{\partial \mathbf{K}}$ is the steady state Jacobian computed with our method

Nonlinear perfect foresight transitions: example

◂ Back



(5 iterations)



(8 iterations)