DSGE Model Estimation
Challenges and (Some) Progress

Frank Schorfheide
University of Pennsylvania

MFM Summer Camp

June 12, 2017

Overview

PART 1. Some Challenges

- Nonlinear Model Solution
- Multiple Equilibria
- Extracting Latent States / Approximating the Likelihood Function

PART 2. Potential Compromises – An Example

This Talk is Based on

ACS B. Aruoba, P. Cuba-Borda, and F. Schorfheide (2017): "Macroeconomic Dynamics Near the ZLB: A Tale of Two Countries," *REStud*, forthcoming.

HS E. Herbst and F. Schorfheide (2016): "Tempered Particle Filtering," PIER Working Paper, 16-017.

SSY F. Schorfheide, D. Song, and A. Yaron (2017): "Identifying Long-Run Risks: A Bayesian Mixed-Frequency Approach," Manuscript, University of Pennsylvania.

These and other related papers are available on my webpage.

Solving a Model with Occasionally-Binding Constraint (ACS)

- Nonlinearities are important for macro-financial modeling.
- Occasionally binding constraints, e.g., ZLB
- ZLB/ELB for nominal interest rates

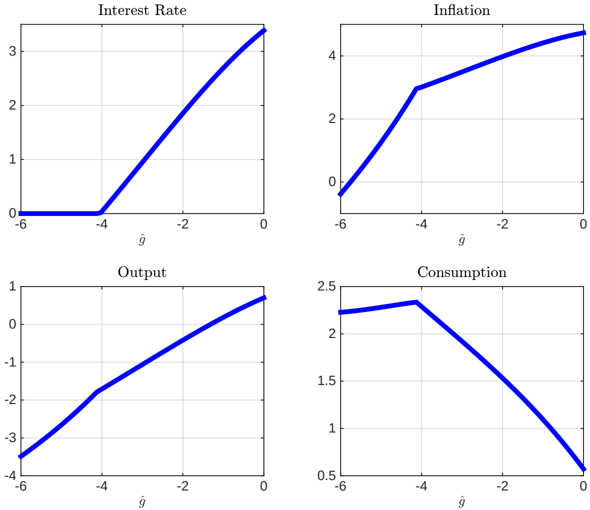
$$R_t = \max\left\{1,\ R_t^* e^{\epsilon_{R,t}}\right\}$$

where

$$R_t^* = \left[r\pi_* \left(\frac{\pi_t}{\pi_*}\right)^{\psi_1} \left(\frac{Y_t}{\gamma Y_{t-1}}\right)^{\psi_2} \right]^{1-\rho_R} R_{t-1}^{\rho_R}.$$

- Projection methods: ACS, Fernandez-Villaverde et al. (2015), Gust, Lopez-Salido and Smith (2012), Judd, Maliar, and Maliar (2010), Maliar and Maliar (2015)

Sample Decision Rules - Small-Scale NK Model for U.S. (ACS)



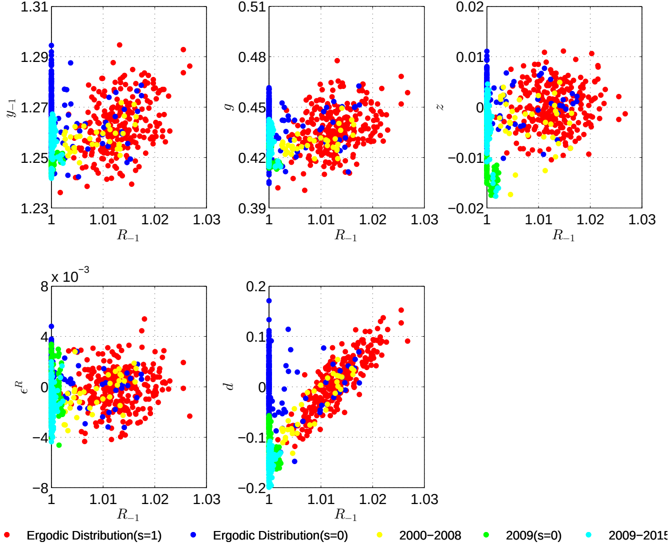
Sketch of Solution Method (ACS)

- Consider decision rule $\pi(\mathcal{S}_t)$, states $\mathcal{S}_t = (R_{t-1}, y^*_{t-1}, d_t, g_t, z_t, \epsilon_{R,t}, s_t)$
- "Stitch" two functions for each decision rule (endogenous "seam"):

$$\pi(\mathcal{S}_t;\Theta) = \begin{cases} f^1_\pi(\mathcal{S}_t;\Theta) & \text{if } R(\mathcal{S}_t) > 1 \\ f^2_\pi(\mathcal{S}_t;\Theta) & \text{if } R(\mathcal{S}_t) = 1 \end{cases}$$

- f^i_j are linear combinations of a complete set Chebyshev polynomials up to 4^{th} order, with weights Θ.
- Choose Θ to minimize sum squared residuals from the Euler Equations over a grid of points: Iterative process (ergodic-set method of Judd, Maliar and Maliar, 2010) the combines states from ergodic simulation and filtering.

Solution Grid - U.S. Data (ACS)

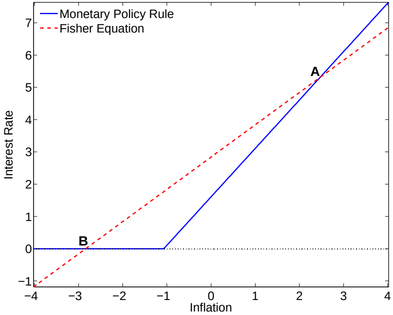


Multiple Equilibria (ACS)

- Local indeterminacy in linearized rational expectations systems, e.g.,

$$y_t = \frac{1}{\theta}\mathbb{E}_t[y_{t+1}] + \epsilon_t, \quad \epsilon_t \sim iid(0,1), \quad \theta \in (0,2].$$

- Multiple steady states:



- Other non-unique nonlinear dynamics, e.g., work of Benhabib, Schmitt-Grohe, and Uribe.

Multiple Equilibria (ACS)

How should one deal with multiplicity?

- Use theoretical arguments, e.g. learning, to eliminate some.
- Pick an interesting one and fit to data.
- Try to parameterize many/all of them and estimate.
- In macro we often introduce sunspot shocks:
  - Markov switching vs. AR(p);
  - exogenous vs. correlated with fundamentals

Example: Two-Equation Model (ACS)

- Euler equation / Fisher equation:

$$\hat{R}_t = \mathbb{E}_t \left[-\hat{M}_{t+1} + \hat{\pi}_{t+1} \right].$$

- Monetary policy rule:

$$\hat{R}_t = \max \left\{ -\log \left(r\pi_* \right), \psi \hat{\pi}_t \right\}.$$

- Exogenous real rate / discount factor:

$$\hat{M}_{t+1} = \rho \hat{M}_t + \sigma \varepsilon_{t+1}.$$

- Combine to single inflation equation:

$$\mathbb{E}_t \left[\hat{\pi}_{t+1} \right] = \max \Big\{ -\log \left(r\pi_* \right) + \rho \hat{M}_t, \psi \hat{\pi}_t + \rho \hat{M}_t \Big\}.$$

Various Equilibria in Simple Model (ACS)

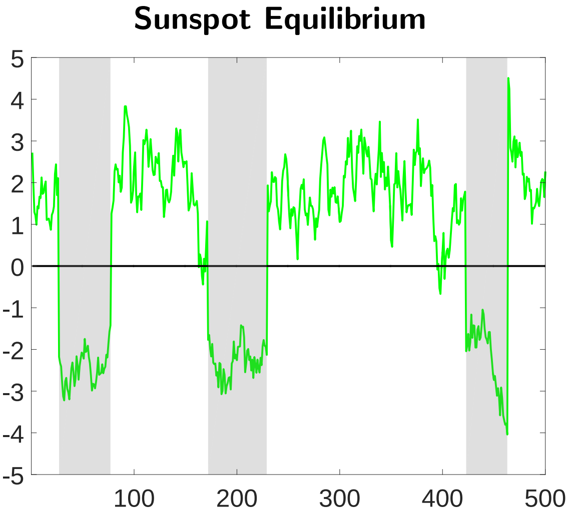
$$\mathbb{E}_t\left[\hat{\pi}_{t+1}\right] = \max\left\{ -\log\left(r\pi_*\right) + \rho\hat{M}_t, \psi\hat{\pi}_t + \rho\hat{M}_t \right\}.$$

- Targeted-Inflation Eq. and Deflation Eq. solutions: $\hat{\pi}_t = \theta_0 - \theta_1\hat{M}_t$.
- Sunspot/Regime-switching solution: $\hat{\pi}_t^{(s)} = \theta_0(s_t) - \theta_1(s_t)\hat{M}_t$, where $s_t \in \{0, 1\}$ is a Markov process.
- Numerical example: $\pi_* = 1.005$, $\psi = 1.5$, $r = 1.005$, $\sigma = 0.0007$, $\rho = 0.9$, $p_{11} = 0.99$ and $p_{00} = 0.95$.

Table: Decision Rule Coefficients

Targeted-Inflation Equilibrium	$\theta_0^* = 0$	$\theta_1^* = 1.5$
Deflation Equilibrium	$\theta_0^D = -0.01$	$\theta_1^D = -1$
Sunspot Equilibrium	$\theta_0(1) = -0.0002$ $\theta_0(0) = -0.0105$	$\theta_1(1) = 1.4611$ $\theta_1(0) = -1.1295$

A Sunspot Equilibrium in Simple Model (ACS)



Filtering

- Occasionally binding constraint causes strong nonlinearity.
- Nonlinear state-space representation of DSGE model requires nonlinear filter:

$$\begin{aligned} y_t &= \Psi(s_t, t; \theta) + u_t, \quad u_t \sim F_u(\cdot;\theta) \\ s_t &= \Phi(s_{t-1}, \epsilon_t; \theta), \quad \epsilon_t \sim F_\epsilon(\cdot;\theta). \end{aligned}$$

- Filtering:
  - Start with $p(s_{t-1}|Y_{1:t-1})$
  - Forecast s_t: $p(s_t|Y_{1:t-1}) = \int p(s_t|s_{t-1})p(s_{t-1}|Y_{1:t-1})ds_{t-1}$
  - Forecast y_t: $p(y_t|Y_{1:t-1}) = \int p(y_t|s_t)p(s_t|Y_{1:t-1})ds_t$
  - Updating $p(s_t|y_t, Y_{1:t-1}) \propto p(y_t|s_t)p(s_t|Y_{1:t-1})$
- Particle Filtering: represent $p(s_{t-1}|Y_{1:t-1})$ by $\{s^j_{t-1}, W^j_{t-1}\}$ such that

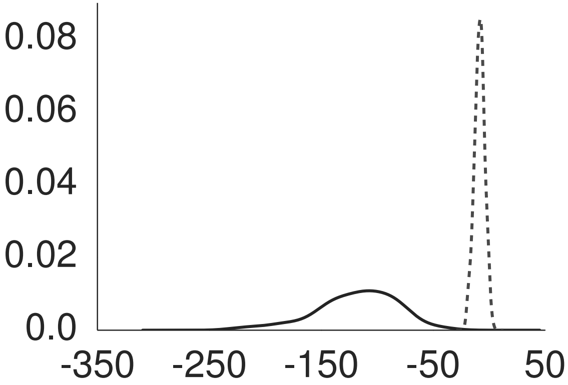
$$\frac{1}{M}\sum_{j=1}^{M} h(s^j_{t-1})W^j_{t-1} \approx \int h(s_{t-1})p(s_{t-1}|Y_{1:t-1})ds_{t-1}.$$

Particle Filtering (HS)

- Algorithm involves
  - Mutation/Forecasting: turn s^j_{t-1} into $\tilde{s}^j_t$. Bootstrap PF: sample $\tilde{s}^j_t \sim p(s_t|s^j_{t-1})$.
  - Correction/Updating: change particle weights to (Bootstrap PF): $\tilde{W}^j_t \propto p(y_t|\tilde{s}^j_t)W^j_{t-1}$.
  - Selection (Optional): Resample to turn $\{\tilde{s}^j_t, \tilde{W}^j_t\}$ into $\{s^j_t, W^j_t = 1\}$.
- Problem: naive forward simulation of Bootstrap PF leads to uneven particle weights.
- (Mostly) Infeasible Solution (Conditionally Optimal): mutate based on $p(s_t|y_t, s^j_{t-1})$.
- Feasible Solution: introduce a sequence of tempering steps.

Exhibit 1: Smets-Wouters Model (HS)

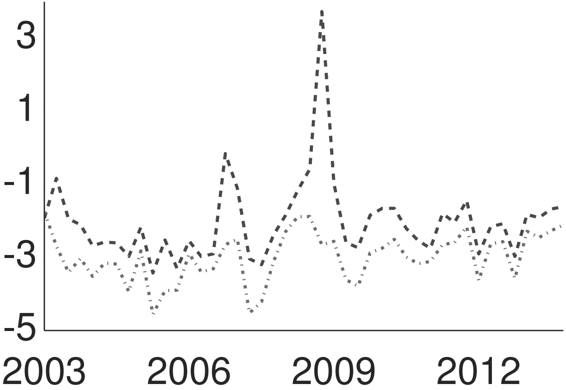
BS ($M = 400,000$) versus CO ($M = 4,000$)



Density estimates of $\hat{\Delta}_1 = \ln \hat{p}(Y|\theta) - \ln p(Y|\theta)$ based on $N_{run} = 100$.

Exhibit 2: Small-Scale NK DSGE Model (HS)

Log Standard Dev of Log-Likelihood Increments



Bootstrap PF ($M = 40,000$) is dashed; Cond-opt. PF ($M = 400$) is dotted.

Use Tempered Particle Filter (HS)

- Construct a sequence "bridge distributions" with inflated measurement errors. Define

$$p_n(y_t|s_t,\theta) \quad \propto \quad \phi_n^{d/2}|\Sigma_u(\theta)|^{-1/2}\exp\bigg\{-\frac{1}{2}(y_t-\Psi(s_t,t;\theta))' \times\phi_n\Sigma_u^{-1}(\theta)(y_t-\Psi(s_t,t;\theta))\bigg\}, \quad \phi_1<\phi_2<\ldots<\phi_{N_\phi}=1.$$

- Bridge posteriors given s_{t-1}:

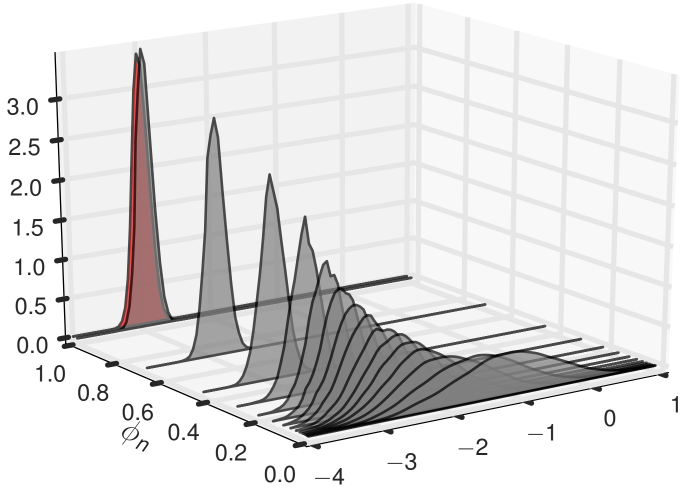
$$p_n(s_t|y_t,s_{t-1},\theta)\propto p_n(y_t|s_t,\theta)p(s_t|s_{t-1},\theta).$$

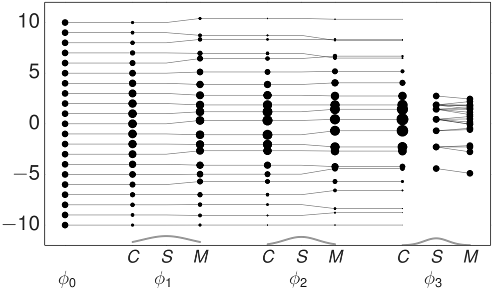
- Bridge posteriors given $Y_{1:t-1}$:

$$p_n(s_t|Y_{1:t})=\int p_n(s_t|y_t,s_{t-1},\theta)p(s_{t-1}|Y_{1:t-1})ds_{t-1}.$$

- Traverse these bridge distributions with "static" Sequential Monte Carlo method (Chopin, 2002). References in stats lit: Godsill and Clapp (2001), Johansen (2016)

Bridge Posteriors: $p_n(s_t|Y_{1:t})$, $n = 1, \ldots, N_\phi$ (HS)





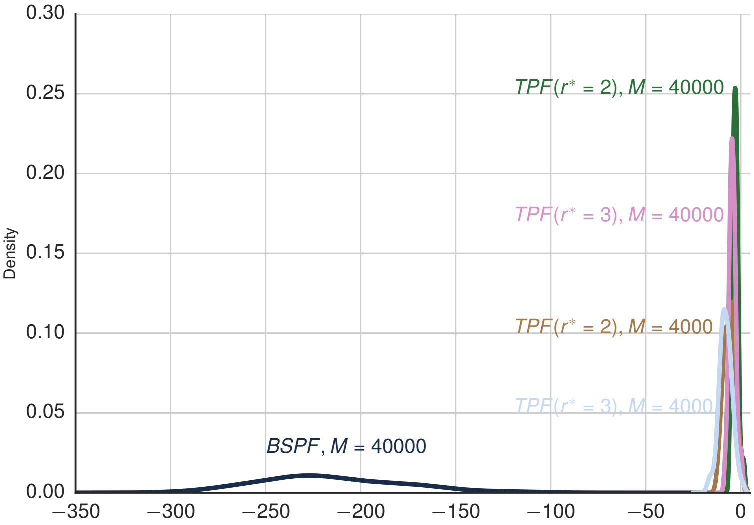
- C is Correction; S is Selection; and M is Mutation.
- $p_n(s_t|Y_{1:t})$ is represented by a swarm of particles $\{s_t^{j,n}, W_t^{j,n}\}_{j=1}^M$:

$$\frac{1}{M}\sum_{j=1}^{M} W_t^{j,n} h(s_t^{j,n}) \xrightarrow{a.s.} \int h(s_t) p_n(s_t|Y_{1:t}) ds_t.$$

Small-Scale Model, Pre Great Recession: Performance Statistics (HS)

	BSPF	TPF			
Number of Particles M	40k	4k	4k	40k	40k
Target Ineff. Ratio r^*		2	3	2	3
High Posterior Density: $\theta = \theta^m$					
Bias	-1.4	-0.9	-1.5	-0.3	-.05
StdD	1.9	1.4	1.7	0.4	0.6
$T^{-1}\sum_{t=1}^{T} N_{\phi,t}$	1.0	4.3	3.2	4.3	3.2
Average Run Time (s)	0.8	0.4	0.3	4.0	3.3
Low Posterior Density: $\theta = \theta^l$					
Bias	-6.5	-2.1	-3.1	-0.3	-0.6
StdD	5.3	2.1	2.6	0.8	1.0
$T^{-1}\sum_{t=1}^{T} N_{\phi,t}$	1.0	4.4	3.3	4.4	3.3
Average Run Time (s)	1.6	0.4	0.3	3.7	2.9

Great Recession: Distribution of Log-lh Approx Error (HS)



So far...

- We've solved the model (maybe multiple solutions).
- Conditional on parameters, we've extracted latent states and approximated likelihood function.
- We still need to
  - parameterize the model;
  - assess its fit;
  - conduct substantive economic analysis;
  - compare to other models / aggregate conclusions across models & data sets etc.

A Plug for Bayesian Inference...

$$p(\theta|Y) = \frac{p(Y|\theta)p(\theta)}{\int p(Y|\theta)p(\theta)d\theta}$$

- Treat uncertainty with respect to shocks, latent states, parameters, and model specifications uncertainty symmetrically.
- Condition inference on what you know (the data Y) instead of what you don't know (the parameter θ).
- Make optimal decision conditional on observed data.

Compromises

- The Full Monty is a real pain: see Gust, C., E. Herbst, D. Lopez-Salido, and M. E. Smith (2017): "The Empirical Implications of the Interest-Rate Lower Bound," *American Economic Review*, forthcoming.
- Potential shortcuts:
  - less accurate model solution;
  - cruder state extraction / likelihood approximation;
  - non-likelihood-based parameterization of model.

Estimation of a Long-Run-Risks Model (SSY)

- Exogenous cash flow processes:

$$
\begin{aligned}
g_{c,t+1} &= \mu_c + x_t + \sigma_{c,t}\eta_{c,t+1} \\
g_{d,t+1} &= \mu_d + \phi x_t + \pi\sigma_{c,t}\eta_{c,t+1} + \sigma_{d,t}\eta_{d,t+1}
\end{aligned}
$$

- Common predictable component:

$$
x_{t+1} = \rho x_t + \sqrt{1-\rho^2}\sigma_{x,t}\eta_{x,t+1}
$$

- Three volatility processes: $i \in \{c, x, d\}$

$$
h_{i,t+1} = \rho_{h_i} h_{i,t} + \sigma_{h_i} w_{h_i,t+1}, \quad \sigma_{i,t} = \varphi_i \sigma \exp(h_{i,t})
$$

- Agents maximize life-time utility, which is defined recursively:

$$
V_t = \left[(1-\delta)\lambda_t C_t^{\frac{1-\gamma}{\theta}} + \delta \left(\mathbb{E}_t[V_{t+1}^{1-\gamma}] \right)^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}}
$$

- Budget constraint:

$$
W_{t+1} = (W_t - C_t) R_{c,t+1}
$$

where W_t is the wealth of the agent, $R_{c,t}$ is the return on all invested wealth.

Solution Short Cuts (SSY)

- Log normal approximation of Euler equation:

$$\mathbb{E}_t[m_{t+1} + r_{i,t+1}] + \frac{1}{2} V_t[m_{t+1} + r_{i,t+1}] \approx 0.$$

- Campbell-Shiller (1988) approximation:

$$\begin{aligned} r_{c,t+1} &\approx \kappa_0 + \kappa_1 pc_{t+1} - pc_t + g_{c,t+1} \\ r_{m,t+1} &\approx \kappa_{0,m} + \kappa_{1,m} pd_{t+1} - pd_t + g_{d,t+1}. \end{aligned}$$

- Approximate volatility process:

$$\sigma^2_{i,t} - (\varphi_i \sigma)^2 \approx (\varphi_i \sigma)^2 (1 - \nu_i) + \nu_i \sigma^2_{i,t} + \sigma_{w_i} w_{i,t+1}, \quad i = \{c, x, d\}$$

- Approximate solution can be computed very fast... looks like:

$$\begin{aligned} pc_t &= A_0 + A_1 x_t + A_{2,c} \sigma^2_{c,t} + A_{2,x} \sigma^2_{x,t} \\ r_{f,t} &= B_0 + B_1 x_t + B_{1,\lambda} x_{\lambda,t} + B_{2,c} \sigma^2_{c,t} + B_{2,x} \sigma^2_{x,t} \\ &\ldots \end{aligned}$$

Empirical (State-Space) Model (SSY)

- Relating model to data:

$$\text{observables} = \text{model variables} + \text{measurement errors}$$

- Observables:
  - Consumption growth (annual or monthly)
  - Dividend growth (monthly)
  - Real returns of the CRSP value-weighted portfolio of all stocks traded on NYSE, AMEX, NASDAQ.
  - Ex ante real risk-free rate

- Measurement errors

Conditionally Linear State-Space Representation (SSY)

- Collecting the bits and pieces:
  - Observables: y_t
  - Model variables: s_t
  - Volatilities: s_t^V

- LRR model delivers state-transition equation:

$$s_{t+1} = \Phi s_t + v_{t+1}(h_t), \quad h_{t+1} = \Psi h_t + \Sigma_h w_{t+1}, \quad w_{t+1} \sim iidN(0, I),$$

- Measurement equation can be written as

$$y_{t+1} = A_{t+1}\big(D + Z s_{t+1} + Z^v s^v_{t+1}(h_{t+1}, h_t) + \Sigma^u u_{t+1}\big), \quad u_{t+1} \sim iidN(0, I).$$

where A_{t+1} is a selection matrix that accounts for the deterministic changes in the data availability.

Bayesian Estimation (SSY)

- Replace likelihood $p(Y|\theta)$ in Bayes Theorem by particle-filter approximation $\hat{p}(Y|\theta)$.
- Conditional on the volatility states our state-space model is linear, which allows for efficient implementation of the filter.
  (see Chen and Liu, 2000 and Shephard, 2013)
- Draws from posterior $p(\theta|Y)$ are generated using a Metropolis-Hastings algorithm.
- Despite use of approximate likelihood $\hat{p}(Y|\theta)$ the Markov chain converges to "exact" posterior $p(\theta|Y)$ (see Andrieu, Doucet, and Holenstein, 2010).
- BOTTOM LINE: linearizations speed up solution and likelihood approximation while retaining accuracy of model solution for reasonable parameter values.

To Recap...

PART 1. Some Challenges

- Nonlinear Model Solution
- Multiple Equilibria
- Extracting Latent States / Approximating the Likelihood Function

PART 2. Potential Compromises – An Example