

DSGE Model Nonlinearities

Frank Schorfheide

University of Pennsylvania, CEPR, NBER, PIER

June 2017

JEDC Plenary Lecture
2017 SCE/CEF Conference

Papers and software available at [**https://web.sas.upenn.edu/schorf/**](https://web.sas.upenn.edu/schorf/)

Large body of recent work on DSGE model nonlinearities:

- stochastic volatility;
- effective lower bound on nominal interest rates;
- occasionally-binding financial constraints;
- general nonlinear dynamics in macro-financial models;
- (...)

Three Tasks

- ① Model Solution
- ② Model Estimation
- ③ Model Assessment

I will provide an overview of some of my recent collaborative research in these areas.

Task 1 – Model Solution

Nonlinear Model Solution

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

- Perturbation solutions capture some nonlinearities but not all
→ not well suited for occasionally-binding constraints.
- Example: ZLB/ELB for nominal interest rates

$$R_t = \max \{1, R_t^* e^{\epsilon_{R,t}}\}, \quad R_t^* = \left[r \pi_* \left(\frac{\pi_t}{\pi_*} \right)^{\psi_1} \left(\frac{Y_t}{Y_t^*} \right)^{\psi_2} \right]^{1-\rho_R} R_{t-1}^{\rho_R}.$$

- Two Challenges:
 - 1 capture "kinks" in decision rules;
 - 2 solution needs to be accurate in region of state-space that is relevant according to model AND according to data.
- Other issue in paper: multiplicity of equilibria, sunspots ...

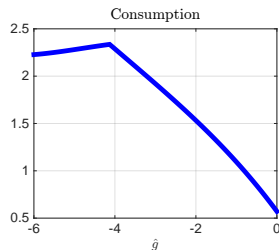
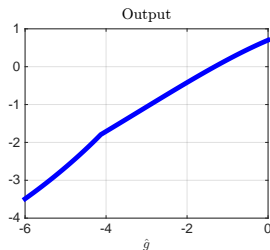
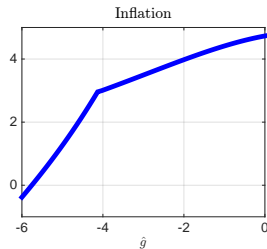
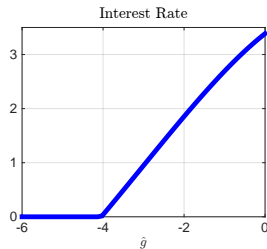
Challenge 1 – Kinks

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

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

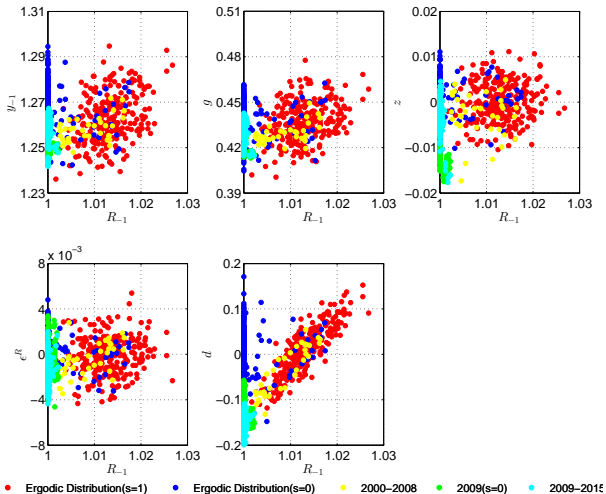
- f_j^i are linear combinations of a complete set Chebyshev polynomials up to 4th order, with weights Θ .

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



Challenge 2 – Accuracy Where it Matters

Choose Θ to minimize sum squared residuals from the (intertemporal) equilibrium conditions over particular grid of points in state space



Task 2 – Model Estimation

Model Estimation – 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.
- Large set of **computational tools** available.

- Bayesian inference is implemented by sampling draws θ^i from the posterior $p(\theta|Y)$.

- Posterior samplers require evaluation of likelihood function:

$\theta \longrightarrow$ model solution \longrightarrow state-space representation $\longrightarrow p(Y|\theta)$.

- State-space representation $\longrightarrow p(Y, S|\theta)$:

$$y_t = \Psi(s_t, t; \theta) + u_t, \quad u_t \sim F_u(\cdot; \theta)$$

$$s_t = \Phi(s_{t-1}, \epsilon_t; \theta), \quad \epsilon_t \sim F_\epsilon(\cdot; \theta).$$

- In order to obtain $p(Y|\theta) = \prod_{t=1}^T p(y_t|Y_{1:t-1}, \theta)$
we need to integrate out latent states S from $p(Y, S|\theta) \longrightarrow$ use filter:

- Initialization: $p(s_{t-1}|Y_{1:t-1}, \theta)$
- Forecasting: $p(s_t|Y_{1:t-1}, \theta)$, $p(y_t|Y_{t-1})$
- Updating: $p(s_t|y_t, Y_{1:t-1}) = p(s_t|Y_{1:t})$.

- **Particle Filtering**: represent $p(s_{t-1}|Y_{1:t-1})$ by $\{s_{t-1}^j, W_{t-1}^j\}_{j=1}^M$ such that

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

- Example: Bootstrap particle filter

- **Mutation/Forecasting**: turn s_{t-1}^j into \tilde{s}_t^j : sample $\tilde{s}_t^j \sim p(s_t|s_{t-1}^j)$.
- **Correction/Updating**: change particle weights to: $\tilde{W}_t^j \propto p(y_t|\tilde{s}_t^j) W_{t-1}^j$.
- **Selection** (Optional): Resample to turn $\{\tilde{s}_t^j, \tilde{W}_t^j\}_{j=1}^M$ into $\{s_t^j, W_t^j = 1\}_{j=1}^M$.
- **Problem**: naive forward simulation of Bootstrap PF leads to uneven particle weights
→ **inaccurate likelihood approximation!**

Tempered Particle Filter

Reference: E. Herbst and F. Schorfheide (2017): “Tempered Particle Filtering,” *NBER Working Paper*, 23448.

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

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

- Bridge posteriors given s_{t-1} :

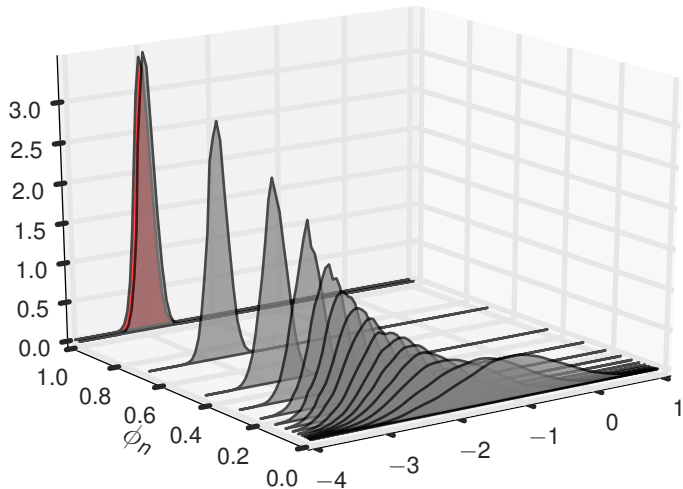
$$p_n(s_t | y_t, s_{t-1}, \theta) \propto p_n(y_t | s_t, \theta) p(s_t | s_{t-1}, \theta).$$

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

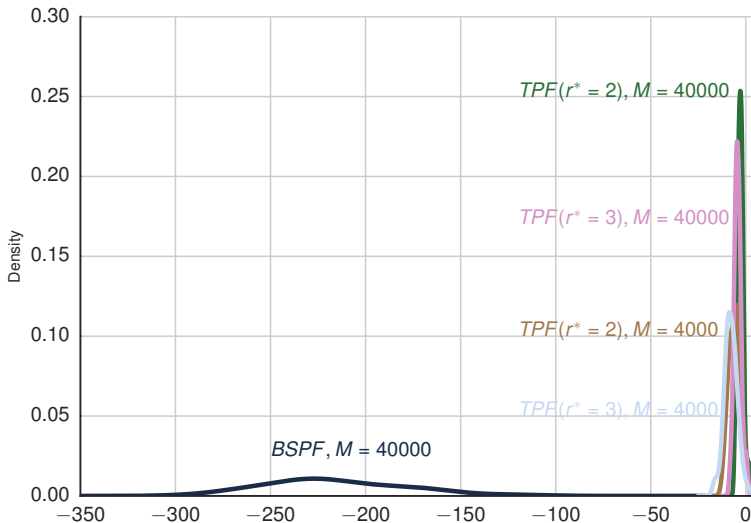
$$p_n(s_t | Y_{1:t}) = \int p_n(s_t | y_t, s_{t-1}, \theta) p(s_{t-1} | Y_{1:t-1}) ds_{t-1}.$$

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

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



Distribution of Log-lh Approx Error – Great Recession Sample



Putting it All Together

- Once a reasonably accurate likelihood approximation has been obtained, it can be embedded in a posterior sampler.
- **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.
- Schorfheide, Song, Yaron (2017): **slight short-cut in model solution** → **conditionally-linear state-space representation** → **efficient particle filter approximation of likelihood** → full Bayesian estimation.

Part III – Model Assessment

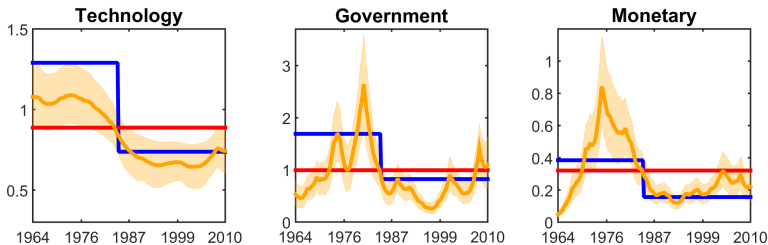
Can the nonlinearities in DSGE models correctly reproduce the nonlinearities in the data?

- **Model Assessment** (see Fernandez-Villaverde, Rubio-Ramirez, and Schorfheide (2016): "Solution and Estimation of DSGE Models," *Handbook of Macroeconomics*, Vol 2., Elsevier):
 - **Relative fit**: comparison with other models.
 - **Absolute fit**: can the DSGE reproduce salient features of the data? Violation of over-identifying restrictions?
- **Linear VARs** have been useful benchmark / reference model for the **evaluation of linearized DSGE models**:
 - testing over-identifying restrictions; model odds
 - comparison of model-implied autocovariances, spectra, impulse responses
 - (...)
- **No obvious benchmark for evaluation of nonlinear models**: generalized autoregressive models? bilinear models? ARCH-M? LARCH? regime-switching models? time-varying coefficient models? threshold autoregressions? smooth transition autoregressions?

Assessing Nonlinearities Can Be Delicate... An Example

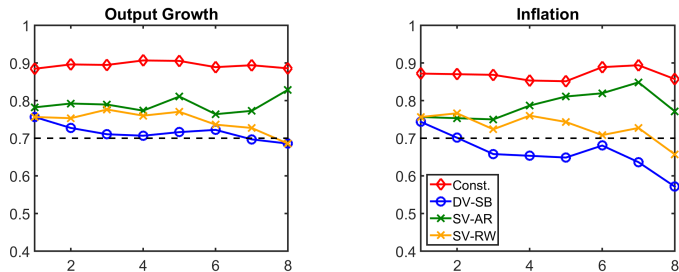
- Many papers argue that it is **important to incorporate stochastic volatility in DSGE models**.
- **Important nonlinearity or device to capture rare events** like Great Moderation, Great Recession?
- **Diebold, Schorfheide, Shin (2016)**: Evaluate forecast performance of DSGE model with **stochastic volatility** versus **structural break in volatility**

Posterior Mean Structural Shock Volatilities / Final Data Vintage

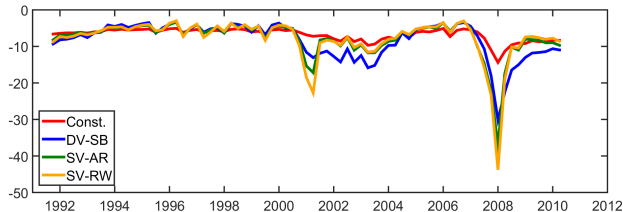


Assessing Nonlinearities Can Be Delicate... An Example: DSS (2016)

Coverage Rates of 70% Interval Forecasts, $h = 1, \dots, 8$



Log Predictive Scores, $h = 4$



Assessing a DSGE Model With Smooth Nonlinearities

Reference: B. Aruoba, L. Bocola, and F. Schorfheide (2017): “Assessing DSGE Model Nonlinearities,” *Manuscript*.

- Small-scale DSGE model with **nominal price and wage rigidities**.
- **Price and wage adjustment costs are potentially asymmetric** to capture downward rigidity, see Kim and Ruge-Murcia (JME, 2009):

$$\Phi(x) = \varphi \left(\frac{\exp(-\psi(x - x_*)) + \psi(x - x_*) - 1}{\psi^2} \right).$$

- **Model consists of**
 - households
 - intermediate goods producers
 - final goods producers
 - central bank / fiscal authority

Some evidence for downward nominal rigidities:

Parameter	1960:Q1-2007:Q4		1984:Q1-2007:Q4	
	Mean	90% Interval	Mean	90% Interval
Price Rigidity				
PC Slope $\kappa(\varphi_p)$	0.02	[0.01, 0.04]	0.21	[0.12, 0.35]
Price Asymmetry ψ_p	150	[130, 175]	165	[130, 192]
Wage Rigidity				
Wage Adj Costs φ_w	18.7	[8.47, 38.1]	11.7	[5.34, 20.2]
Wage Asymmetry ψ_w	67.4	[33.2, 99.5]	59.4	[21.7, 90.9]

$$\Phi(x) = \varphi \left(\frac{\exp(-\psi(x - x_*)) + \psi(x - x_*) - 1}{\psi^2} \right).$$

- State ($s_{i,t}$) and control ($c_{i,t}$) variables evolve according to

$$\begin{aligned}c_{i,t} &= \psi_{1i}(\theta) + \psi_{2ij}(\theta)s_{j,t} + \psi_{3ijk}(\theta)s_{j,t}s_{k,t} \\ s_{i,t+1}^{\text{end}} &= \zeta_{1i}^{\text{end}}(\theta) + \zeta_{2ij}^{\text{end}}(\theta)s_{j,t} + \zeta_{3ijk}^{\text{end}}(\theta)s_{j,t}s_{k,t} \\ s_{i,t+1}^{\text{exo}} &= \zeta_{2i}^{\text{exo}}(\theta)s_{i,t}^{\text{exo}} + \zeta_{3i}^{\text{exo}}(\theta)\epsilon_{i,t+1}.\end{aligned}$$

- Perturbation solutions are easy to compute (DYNARE), improve accuracy near steady state (though not necessarily globally).
- BUT: are $\psi_{3ijk}(\theta)s_{j,t}s_{k,t}$ and $\zeta_{3ijk}^{\text{end}}(\theta)s_{j,t}s_{k,t}$ consistent with data?

Our Approach – Posterior Predictive Checks

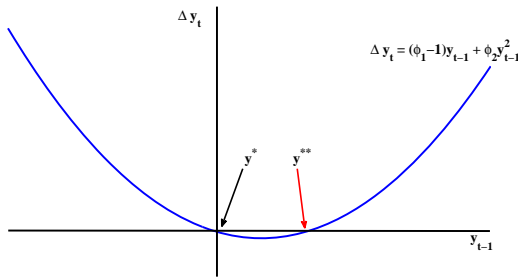
- Develop a nonlinear time series model that mimics structure of DSGE solution.
- Compare estimates of this model based on actual data and DSGE model-generated data.
- Alternative approaches:
 - Barnichon and Matthes (2016): create nonlinear benchmark or DSGE evaluation using Gaussian mixture approximation of moving ave. representation.
 - Ruge-Murcia (2016): indirect inference based on a VAR with higher-order terms and some DSGE model-implied zero restrictions.
 - Time-variation as / versus nonlinearity: literature on TVP VARs building on Cogley and Sargent (2002, 2005) and Primiceri (2005); evidence from time-varying model weights as in Del Negro, Hasegawa, and Schorfheide (2016).

A “Naive” Quadratic Autoregressive Model

- Generalized autoregressive (GAR) models (e.g. Mitnik, 1990): add quadratic terms to a standard autoregressive:

$$y_t = \phi_0 + \phi_1(y_{t-1} - \phi_0) + \phi_2(y_{t-1} - \phi_0)^2 + \sigma u_t$$

- Unattractive features: (i) multiple steady states; (ii) explosive dynamics.



- Problem is well-known in DSGE model solution literature → pruning, e.g., Kim, Kim, Schaumburg, and Sims (2008), Lombardo (2010), Andreasen, Fernandez-Villaverde, and Rubio-Ramirez (2016) (...).

QAR(1,1): Specification

- We set $f_{uu} = 0$ to maintain a conditional Gaussian distribution and consider the system as a nonlinear state-space model:

$$y_t = \phi_0 + \phi_1(y_{t-1} - \phi_0) + \phi_2 s_{t-1}^2 + (1 + \gamma s_{t-1})\sigma u_t$$

$$s_t = \phi_1 s_{t-1} + \sigma u_t \quad u_t \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$$

- Important properties:
 - Conditional linear structure facilitates calculation of moments; see Andreasen et al.
 - Stationary if $|\phi_1| < 1$.
 - Nonlinear impulse responses and conditional heteroskedasticity.

Relationship Between Nonlinear Dynamics and TVP Models

- Write nonlinear model as

$$y_t = f(y_{t-1}) + \sigma u_t = \underbrace{\phi(y_{t-1})}_{\phi_t} y_{t-1} + \sigma u_t.$$

- Could treat the estimation of $\phi(y_{t-1})$ nonparametrically, e.g., with prior

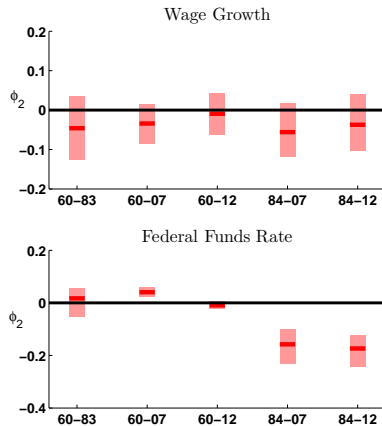
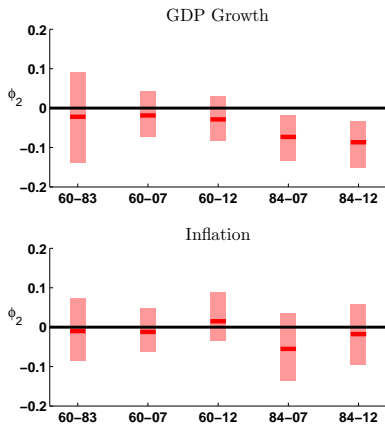
$$\phi(0) \sim N(\rho, \lambda), \quad \phi(y) - \phi(0) \sim N(0, \delta|y|).$$

- This is *ex ante* different from assuming that

$$y_t = \phi_t y_{t-1} + \sigma u_t, \quad \phi_t = \phi_{t-1} + \sigma_\eta \eta_t.$$

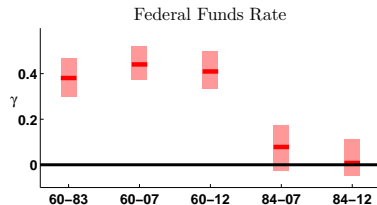
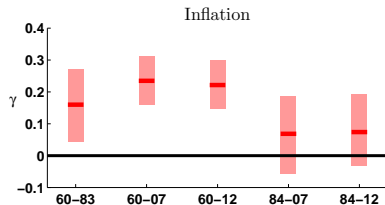
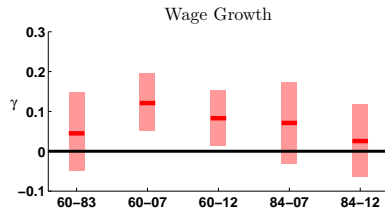
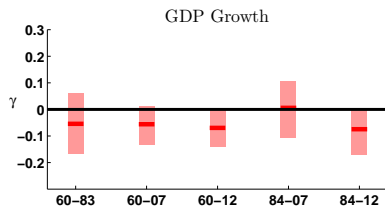
- State (y_t) dependence versus time dependence of ϕ_t .

Estimation of QAR(1,1) Model on U.S. Data – Φ_2



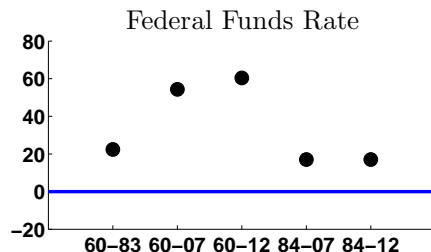
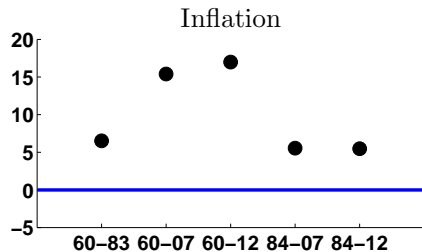
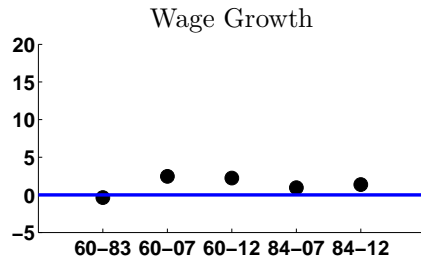
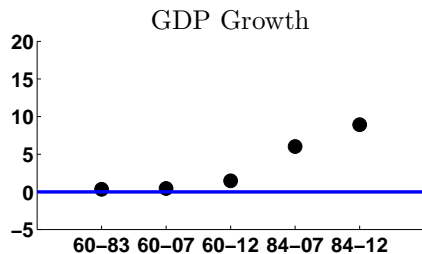
$$y_t = \phi_0 + \phi_1(y_{t-1} - \phi_0) + \phi_2 s_{t-1}^2 + (1 + \gamma s_{t-1})\sigma u_t, \quad s_t = \phi_1 s_{t-1} + \sigma u_t$$

Estimation of QAR(1,1) Model on U.S. Data – γ

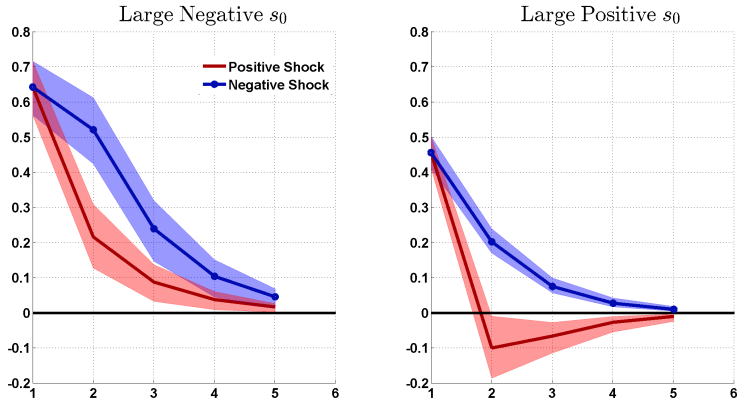


$$y_t = \phi_0 + \phi_1(y_{t-1} - \phi_0) + \phi_2 s_{t-1}^2 + (1 + \gamma s_{t-1})\sigma u_t, \quad s_t = \phi_1 s_{t-1} + \sigma u_t$$

Log Marginal Data Density Differentials: QAR(1,1) versus AR(1)



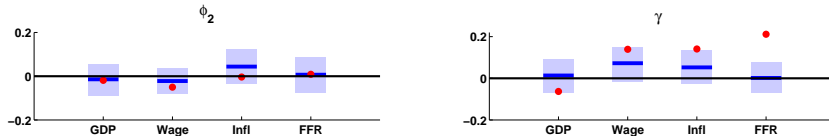
Some Properties: Impulse Responses of GDP Growth (in Absolute Terms), 1984-2012 Sample



$$y_t = 0.53 + 0.36(y_{t-1} - \phi_0) - 0.09s_{t-1}^2 + (1 - 0.07s_{t-1})0.28u_t$$

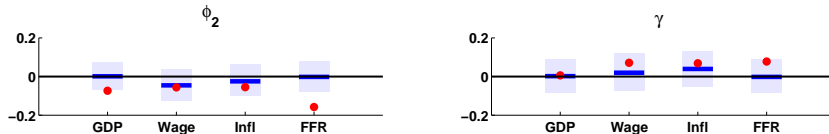
$$s_t = 0.36s_{t-1} + 0.28u_t$$

1960-2007 Sample



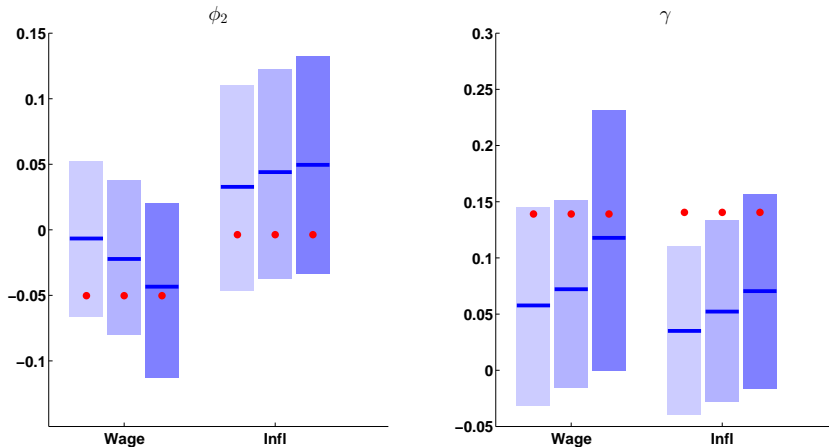
- QAR estimates from actual and model-generated data are similar.
- Only interest rates exhibit noticeable differences.
- Except for wage and inflation $\hat{\gamma}$, nonlinearities are generally weak.

1984-2007 Sample



- Model does not generate nonlinearity ($\hat{\phi}_2$) in GDP dynamics.

Effect of Adjustment Costs on Nonlinearities: 1960-2007 Sample



No asymmetric costs is $\psi_p = \psi_w = 0$ (light blue); high asymmetric costs is $\psi_p = \psi_w = 300$ (dark blue). Large red dots correspond to posterior median estimates based on U.S. data.

Summary of Empirical Results

- Some nonlinearities in U.S. data:
 - Post 1983: output growth displays sharp declines and slow recoveries.
 - 1960-2007: inflation and nominal wage growth display conditional heteroskedasticity.
 - Post 1983: downward adjustments in FFR seem to be easier than upward adjustments.
- DSGE model captures some but not all nonlinearities:
 - Conditional heteroskedasticity in inflation and nominal wage growth through asymmetric adjustment costs that penalize downward movements.
 - But no nonlinearities in model-implied output growth and FFR.

- Literature on methods and applications for DSGE models is well and alive!
- Significant progress in area of model solution and estimation techniques.
- More work needed on the model assessment:
 - Do nonlinearities in one area of model correctly propagate to other areas?
 - Does model perform well in crisis times?
 - Are nonlinearities strong enough so that they are measurable in “short” samples?