Nonlinear Dynamic Factor Models with Financial and Macroeconomic Applications

Pablo A. Guerrón Quintana[1] Alexey Khazanov[2] Molin Zhong[3]

[1]Boston College
[2]The Hebrew University of Jerusalem
[3]Federal Reserve Board

June 2023

Motivation

- **Dynamic factor models** (DFM) are a fundamental tool in empirical economics for both structural and forecasting analysis (Stock and Watson, 2016).
- DFMs are mostly specified *linearly*.
  - *Linear* relationship between the factors today and their past values.
  - *Linear* relationship between the observables and the factors.
- However, the **Great Recession**, **Euro debt crisis**, and **COVID-19** crisis point to strong nonlinearities in the data.
  - Effective lower bound on interest rates.
  - Spikes in corporate and sovereign debt spreads.
  - Asymmetric tail risk behavior of GDP growth (Adrian, Boyarchenko, Giannone, 2019).

This paper

- We introduce a **nonlinear dynamic factor model.**
  - **Nonlinear** relationship between factors today and their past values.
  - **Nonlinear** relationship between observables and factors.

- Our nonlinear dynamic factor model is inspired by the pruned second-order state-space model of Kim et al. (2008) and Andreasen et al. (2019).

- The model can generate novel implications such as
  - Asymmetric and state-dependent IRFs (Andreasen et al., 2019)
  - Non-normal predictive distributions that feature time-varying volatility and asymmetric tail behavior.

- **Applications**
  - Nonlinear credit cycle
  - *Model the shadow rate indicator* à la Wu and Xia, 2016

Related Literature

- **Perturbation methods in DSGE models and their time series applications:** Kim et. al. (2008), Andreasen et. al. (2017), Aruoba et. al. (2017)
- **Dynamic factor models:** Chauvet (1998), Del Negro and Otrok (2008), Aruoba and Diebold (2010), Banbura and Modugno (2014), Cheng et. al. (2016), Carrasco and Rossi (2016), Stock and Watson (2016) and references therein... **Gorodnichenko Ng, 2017**
- **Macro and Financial Tail risks:** Adrian et. al. (2019), Carriero et. al. (2020 a,b), Caldara et. al. (2020), Cook and Doh (2020), Plagborg-Moller et. al. (2020), Caldara et. al. (2021)

NLDFM

Connection to theory: Pruned System

Following Andreasen, Fernandez-Villaverde, Rubio-Ramirez, 2019:

$$f_t = \mathcal{H}(f_{t-1}) + \sigma \nu_t$$

$$\downarrow$$

2^{nd} order approx.

$$\downarrow$$

$$f_t = h_x f_{t-1} + 0.5 h_{xx} (f_{t-1})^2 + \text{const} + \sigma \nu_t$$

Connection to theory: Pruned System

Following Andreasen, Fernandez-Villaverde, Rubio-Ramirez, 2019:

$$f_t = \mathcal{H}(f_{t-1}) + \sigma\nu_t$$

$$\downarrow$$

2^{nd} order approx.

$$\downarrow$$

$$f_t = h_x f_{t-1} + 0.5 h_{xx}(f_{t-1})^2 + \text{const} + \sigma\nu_t$$

- Substitute: $f_t = f_t^f + f_t^s$

$$f_t^f + f_t^s = h_x\left(f_{t-1}^f + f_{t-1}^s\right) + 0.5 h_{xx}\left(f_{t-1}^f + f_{t-1}^s\right)^2 + \text{const} + \sigma\nu_t \quad (1)$$

- Leave two equations: for 1st and 2nd order,

$$f_t^f = h_x f_{t-1}^f + \sigma\nu_t; \qquad f_t^s = h_x f_{t-1}^s + 0.5 h_{xx}\left(f_{t-1}^f\right)^2 + \text{const}$$

Connection to theory: Pruned System

Following Andreasen, Fernandez-Villaverde, Rubio-Ramirez, 2019:

$$f_t = \mathcal{H}(f_{t-1}) + \sigma\nu_t$$

$$\downarrow$$

2^{nd} order approx.

$$\downarrow$$

$$f_t = h_x f_{t-1} + 0.5 h_{xx}(f_{t-1})^2 + \text{const} + \sigma\nu_t$$

- Substitute: $f_t = f_t^f + f_t^s$

$$f_t^f + f_t^s = h_x\left(f_{t-1}^f + f_{t-1}^s\right) + 0.5 h_{xx}\left(f_{t-1}^f + f_{t-1}^s\right)^2 + \text{const} + \sigma\nu_t \quad (1)$$

- Result: stationary ($|h_x| < 1$) system with representation:

$$f_t = \text{const}' + \sum_{k=1}^{\infty} \phi_k\left(h_x, h_{xx}\right)\sigma^2\nu_{t-k}^2 + \sum_{k=1}^{\infty} h_x^k \sigma\nu_{t-k} + \sigma\nu_t$$

Nonlinear Dynamic Factor Model

- We take a pruned second-order approximation to a general nonlinear relationship:

$$f_t = \mathcal{H}(f_{t-1}) + \sigma\nu_t,$$

motivated by the work of Kim et. al. (2008), Andreasen et. al. (2019) and Aruoba et. al. (2017).

Measurement, e.g.	$Y_t = Gf_t + e_t$
Motion: pruned	$f_t = c + f_t^f + f_t^s$
	$f_t^f = h_x f_{t-1}^f + \sigma\nu_t$
	$f_t^s = h_x f_{t-1}^s + 0.5 h_{xx}\left(f_{t-1}^f \times f_{t-1}^f\right)$
Measurement errors	$e_t \sim N(0, \Omega_e)$
Motion innovation	$\nu_t \sim N(0, I)$

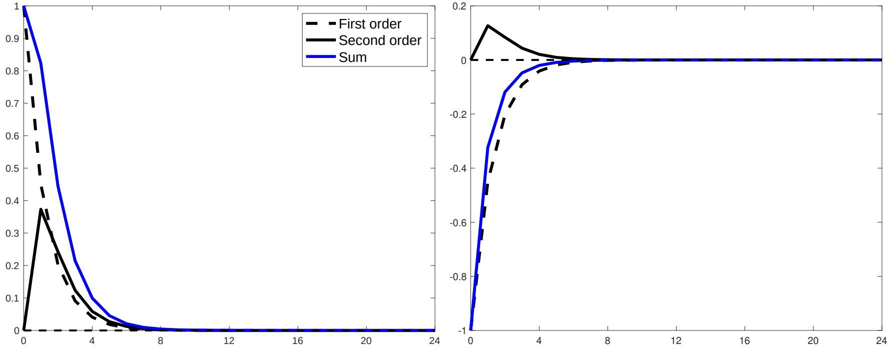
A useful state space representation

$$\begin{pmatrix} f_t^f \\ f_t^s \\ \left(f_t^f\right)^2 \end{pmatrix} = \begin{pmatrix} h_x & 0 & 0 \\ 0 & h_x & \frac{1}{2}h_{xx} \\ 0 & 0 & h_x^2 \end{pmatrix} \begin{pmatrix} f_{t-1}^f \\ f_{t-1}^s \\ \left(f_{t-1}^f\right)^2 \end{pmatrix} + \begin{pmatrix} \sigma & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \sigma^2 & 2\sigma h_x \end{pmatrix} \begin{pmatrix} \nu_t \\ \nu_t^2 \\ f_{t-1}^f \nu_t \end{pmatrix}$$

- Asymmetric responses to shocks from ν_t^2.
- State-dependent dynamics conditional on f_{t-1}^f.
  - Sign of f_{t-1}^f important determinant of IRF and conditional covariance dynamics.
  - Model features time-varying volatility.
- Correlation between mean and volatility.
  - Unconditional correlation between f_t^s and $\left(f_t^f\right)^2$ is non-zero.

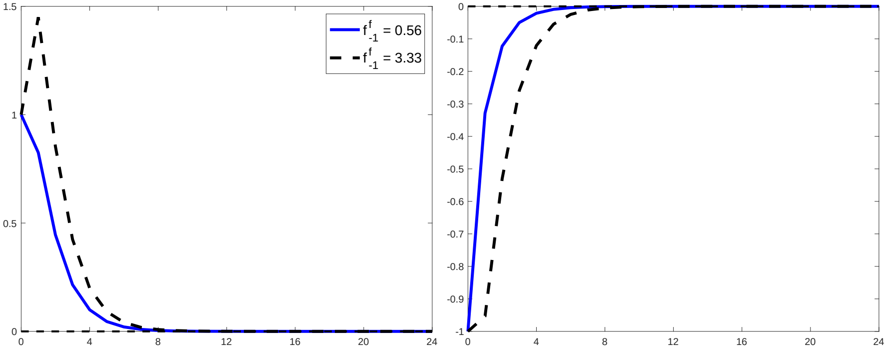
Impulse response functions

- Model generates asymmetric IRFs controlled by the sign of h_{xx}.

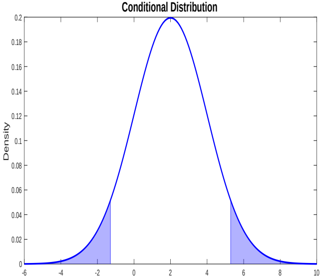


Impulse response functions

- Model generates state-dependent IRFs
- ...and therefore time-varying volatility.



Distributional Moments of Interest



- **Standard deviation:** Width, Dispersion
- **Shortfall:** Left tail behavior

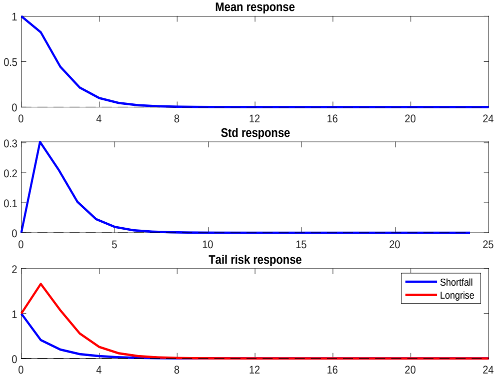
$$SF_\alpha(h) = E_t\left[x_{t+h} | x_{t+h} < q_\alpha(x_{t+h})\right]$$

- **Longrise:** Right tail behavior

$$LR_\alpha(h) = E_t\left[x_{t+h} | x_{t+h} > q_{1-\alpha}(x_{t+h})\right]$$

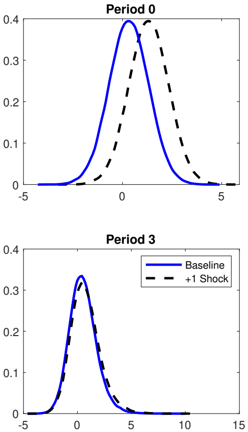
SF, LR

Higher-order moment responses



- Standard deviation increases with a one period lag.
- Simultaneous increase in the mean and standard deviation generates a larger movement in the 5% longrise relative to the 5% shortfall.

Predictive distribution responses



- Initial shock only shifts the mean of the forecast and the distribution is normal.
- Second-order factor responds with a one period lag, which leads to non-normalities.

Estimation Details: 2 cases

Credit growth	Interest rates
Linear measurement	Effective lower bound
Gibbs sampler with particle smoother	Metropolis Hastings with bootstrap particle filter
Fewer particles due to ancestor sampling	Many particles
Uses all the data	Retains the information structure
Approximation of the model Details	Model itself Details

Application-1: Nonlinear Credit Cycle

Data

- Normalized real credit growth in the US
- Nonfinancial business, household, financial and government sectors (Z.1 Financial Accounts data by the Fed)

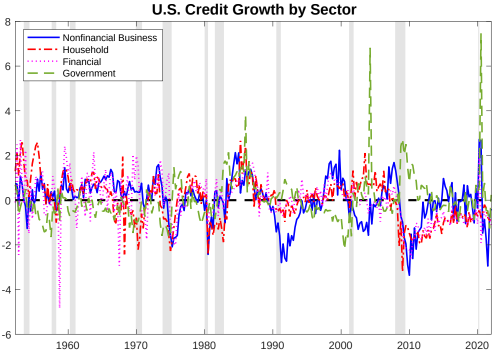


Figure: Normalized real credit growth by sector in the United States: 1952:Q1-2021:Q4 with National Bureau of Economic Research recession shading.

Figure: Estimated Credit Cycle and the Contribution of the Second-Order Factor

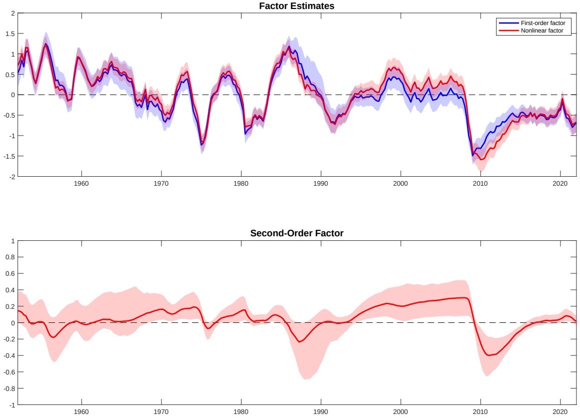


Figure: Shaded areas denote 68% credible sets.

Figure: State-Dependent Impulse Response Functions in Three Periods

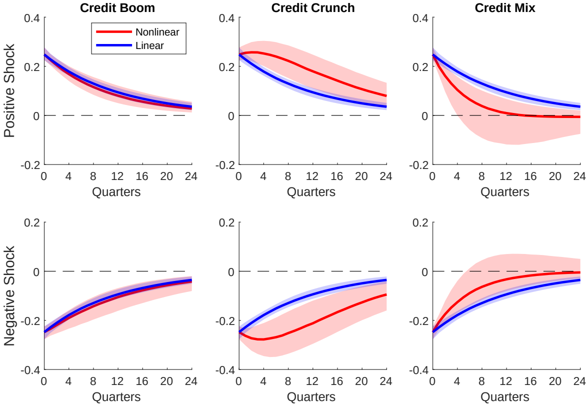


Figure: *

Columns: mid-2000s, credit bust period 2010, mixed case before the early 1990s recession.

Shaded areas: 68% credible sets.

Figure: Impulse Response Functions of the Mean, Standard Deviation, and Tail Risk During the Credit Crunch

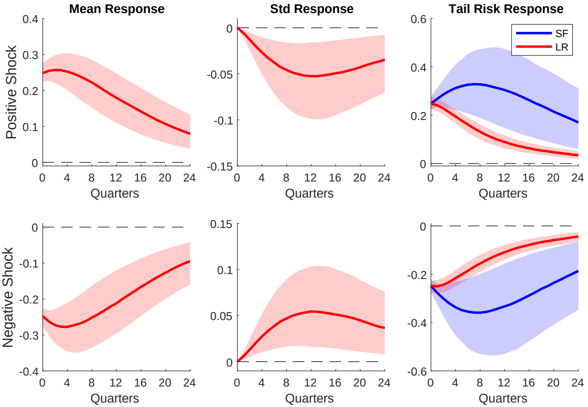


Figure: *

Credit crunch period of 2010. Shaded areas: 68% credible sets.

Jump to conclusions

Application 2: Shadow Interest rate

- Data: one-month forward rates constructed using Wu and Xia (2016) approach to Gurkaynak et al. (2007) data: 1990-2019.

- Measurement:

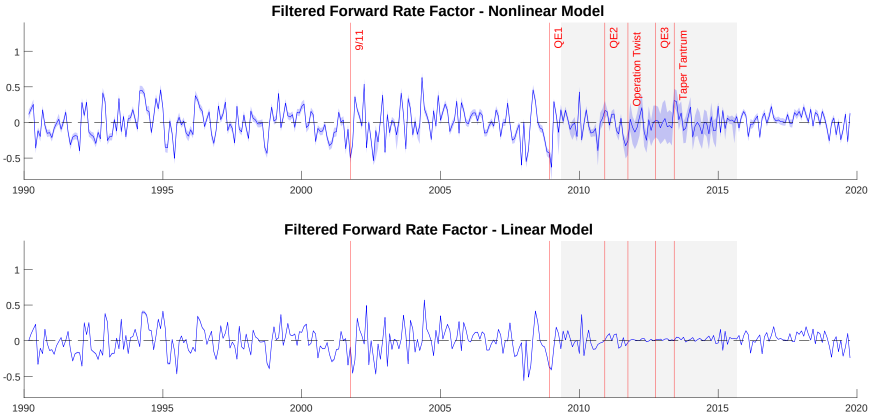
$$\Delta forward_t^h = m_h + \begin{cases} G_h(c + f_t^f + f_t^s) + \eta^h \varepsilon_t^h & \text{if } \widehat{S}_t^h >= 0.3 \\ -m_h + \eta^h \varepsilon_t^h & \text{otherwise} \end{cases} \tag{2}$$

where $\widehat{S}_t^h = \sum_{\tau=2}^{t} \left(m_h + G_h\left(c + f_\tau^f + f_\tau^s\right)\right) + forward_1^h$
$\Delta forward_t^h = forward_t^h - forward_{t-1}^h$, h– maturity

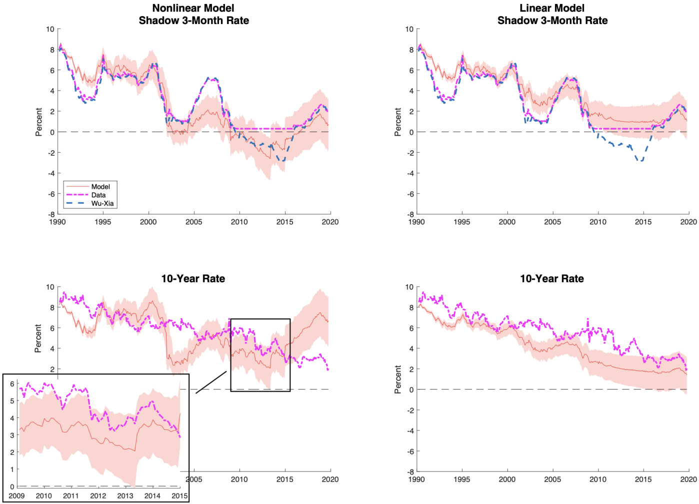
- Motion:

$$\begin{aligned} f_t^f &= h_x f_{t-1}^f + \sigma \nu_t \\ f_t^s &= h_x f_{t-1}^s + \frac{1}{2} h_{xx} \left(f_{t-1}^f\right)^2 \end{aligned} \tag{3}$$

Figure: Filtered Estimates of the Forward Rate Factor from the Nonlinear and Linear Models



- Very similar outside the ZLB
- Shadow rate changes are non-zero at the ELB.
- Next: Sum to get the shadow rate itself



- Outside ELB: tracks mostly 3m rate; at ELB: longer maturities.
- LR test: $h_{xx} = 0$ – no change in relation at the ELB

Conclusion

- Introduce a nonlinear dynamic factor model.
- The nonlinear dynamic factor model can generate asymmetric and state-dependent impulse response functions.
- The model can also generate non-normal predictive distributions that feature time-varying volatility and asymmetric tail behavior.
- Two applications:
  - Credit growth
  - ELB

Appendix

Gibbs Sampler with Particle Smoother

Step 1. Draw G, η given f_t^f, f_t^s, and y_t. This step follows a standard linear regression model.

Step 2. Draw h_x, h_{xx} given σ, G, η, f_t^f, f_t^s, and y_t. RW Metropolis to draw h_x and h_{xx}. Given h_x and h_{xx}, the proposal:

$$\begin{pmatrix} h_x^{prop} \\ h_{xx}^{prop} \end{pmatrix} = \begin{pmatrix} h_x \\ h_{xx} \end{pmatrix} + S^h \zeta,\ \zeta \sim N(0, I).$$

Given proposed h_x^{prop} and h_{xx}^{prop}, calculate its likelihood. Update c and f_t^s

$$f_t^{s,prop} = h_x^{prop} f_{t-1}^{s,prop} + \frac{1}{2} h_{xx}^{prop} \left(f_{t-1}^f\right)^2 .$$

We initialize $f_0^{s,prop} = f_0^s$. Likelihood of the proposal, two steps:

$$y_t - G\left(c^{prop} + f_t^f + f_t^{s,prop}\right) = \eta \epsilon_t; \quad f_t^f - h_x^{prop} f_{t-1}^f = \sigma \nu_t. \tag{4}$$

Accept with probability:

$$prob = \max \left\{ \frac{\prod_{t=1}^T p\left(y_t | c^{prop}, G, \eta, f_t^f, f_t^{s,prop}\right) p_{trans}\left(f_t^f | h_x^{prop}, \sigma, f_{t-1}^f\right) g\left(h_x^{prop}, h_{xx}^{prop}\right)}{\cdots}, 1 \right\} \tag{5}$$

Draw σ^2 given G, η, f_t^f, f_t^s, and y_t using a RW Metropolis step. Given σ^2, :

$$\sigma^{2,prop} = \sigma^2 + S^\sigma \iota, \ \iota \sim N(0, I).$$

Likelihood of the proposal:

$$\begin{aligned} y_t - G\left(c^{prop} + f_t^f + f_t^s\right) &= \eta \epsilon_t \\ f_t^f - h_x f_{t-1}^f &= \sigma^{prop} \nu_t. \end{aligned} \tag{6}$$

Accept probability:

$$prob = \max\left\{ \frac{\prod_{t=1}^T p\left(y_t | c^{prop}, G, \eta, f_t^f, f_t^s\right) p_{trans}\left(f_t^f | h_x, \sigma^{2,prop}, f_{t-1}^f\right) g\left(\sigma^{2,prop}\right)}{\prod_{t=1}^T p\left(y_t | c^{curr}, G, \eta, f_t^f, f_t^s\right) p_{trans}\left(f_t^f | h_x, \sigma^{2,curr}, f_{t-1}^f\right) g\left(\sigma^{2,curr}\right)}, 1 \right\} \tag{7}$$

Step 3. Draw f_t^f, f_t^s given σ, G, η, h_x, h_{xx}, and y_t using the particle Gibbs sampler with ancestor sampling.

- **Initialize particle smoother:** For particles $j = 1, ..., N-1$. simulate for 500 periods and use the final period of the simulation to determine: $f_0^{f,(j)}, f_0^{s,(j)}, f_1^{s,(j)}$. Note that $f_1^{s,(j)}$ is a function of $f_0^{f,(j)}, f_0^{s,(j)}$, so it is known.
- **Draw first period:** For particles $j = 1, ..., N-1$. We determine $f_1^{f,(j)}, f_2^{s,(j)}$ by simulation.
- **Fix final particle:** Fix $f_0^{f,(N)}, f_0^{s,(N)}, f_1^{f,(N)}, f_1^{s,(N)}$, and $f_2^{s,(N)}$ equal to $f_0^{f,*}, f_0^{s,*}, f_1^{f,*}, f_1^{s,*}$, and $f_2^{s,*}$, where $*$ is the accepted previous draw.
- **Set weights:** Compute $w_1^{(j)} = \frac{p(y_1|f_1^{f,(j)}, f_1^{s,(j)})}{\sum_{jj=1}^N p(y_1|f_1^{f,(jj)}, f_1^{s,(jj)})}$ for $j = 1, ..., N$. For $t = 2, ..., T$:
- **Sample indices to set ancestors for each particle:** For particles $j = 1, ..., N-1$. Draw $a_t^{(j)}$ from the distribution w_{t-1}. Simulate the following:

$$f_t^{f,(j)} = h_x f_{t-1}^{f,(a_t^{(j)})} + \sigma \nu_t \quad (8)$$

- **Fix the final particle:** Fix $f_t^{f,(N)}$ equal to $f_t^{f,*}$.
- **Compute auxiliary weights for the fixed particle:** For $j = 1, ..., N$. We compute the auxiliary weights for the fixed particle as follows:

$$w_t^{aux,(j)} = w_{t-1}^{(j)} p(y_t | f_t^{f,(N)}, f_t^{s,(j)}) g(f_t^{f,(N)} | f_{t-1}^{f,(j)}) p(y_{t+1} | f_{t+1}^{f,(N)}, f_{t+1}^{s,(N')}) g(f \tag{9}$$

When calculating $f_{t+1}^{s,(N')}$, we have to take into account that $f_{t+1}^{s,(N')}$ depends on $f_t^{s,(j)}$. Therefore, $f_{t+1}^{s,(N')}$ does not equal $f_{t+1}^{s,(N)}$. The formula is

$$f_{t+1}^{s,(N')} = h_x f_t^{s,(j)} + \frac{1}{2} h_{xx} \left(f_t^{f,(N)} \right)^2. \tag{10}$$

- **Sample the associated ancestor index for particle N:** We sample $a_t^{(N)}$ from the distribution w_t^{aux}. Note that we have to update $f_{t+1}^{s,(N)}$ to make it consistent with the selected ancestor:

$$f_{t+1}^{s,(N)} = h_x f_t^{s,(a_t^{(N)})} + \frac{1}{2} h_{xx} \left(f_t^{f,(N)} \right)^2 . \tag{11}$$

- **Set weights:** Compute $w_t^{(j)} = \frac{p(y_t | f_t^{f,(j)}, f_t^{s,(j)})}{\sum_{jj=1}^{N} p(y_t | f_t^{f,(jj)}, f_t^{s,(jj)})}$ for $j = 1, ..., N$.
- **Sample selected states:** Sample $*$ according to w_T. Set $f_t^{f,*}, f_t^{s,*}$ equal to the sampled state.

Metropolis Hasting with Bootstrap Particle Filter: Details

Propose a new set of parameters $\Theta^{prop} = \left\{ G^{prop}, \eta^{prop}, h_x^{prop}, h_{xx}^{prop}, \sigma^{2,prop} \right\}$.
3 blocks:

- Block 1 (factor equation) $\Theta_1^{prop} = \left\{ h_x^{prop}, h_{xx}^{prop}, \sigma^{2,prop} \right\}$;
- Block 2 (measurement equation loadings) $\Theta_2^{prop} = \left\{ G^{prop} \right\}$;
- Block 3 (measurement equation variances) $\Theta_3^{prop} = \left\{ \eta^{prop} \right\}$.

For each block 50 draws, holding other parameters at their previously accepted values.

$$\Theta_i^{prop} = \Theta_i^{curr} + 0.95 S_{i,1} \zeta_1 + 0.05 S_{i,2} \zeta_2, \ \zeta_i \sim N(0, I) \quad i = 1, 2, 3, 4$$

We tune the variance-covariance matrix of the proposals $S_{i,1}$ and $S_{i,2}$ in an adaptive fashion over the first $30,000$ draws of the algorithm. $S_{i,1}$ is calculated using the variance-covariance matrix from all of the previous draws multiplied by a scaling parameter that decreases if the previous 250 draws within the block had an acceptance rate less than 10%. $S_{i,2}$ is a diagonal matrix that is meant to introduce some independent noise within the proposal. It is multiplied by a separate scaling parameter that decreases if the

Bootstrap Particle filter

- **Initialize the particle filter:** For particles $j = 1, ..., N$. To take a draw from the unconditional distribution, we simulate the model for 500 periods and use the final period of the simulation to determine: $f_0^{f,(j)}, f_0^{s,(j)}, f_1^{s,(j)}$.
  For $t = 1, ..., T$:
- **Prediction step:**

  Given particles and weights at $t-1$: $\left\{ f_{t-1}^{f,(j)}, f_t^{s,(j)}, w_{t-1}^{(j)} \right\}$.

  1. For particles $j = 1, ..., N$. Draw a new particle $\left\{ f_t^{f,(j)}, f_{t+1}^{s,(j)} \right\}$ from

$$f_t^{f,(j)} = h_x f_{t-1}^{f,(j)} + \sigma \nu_t$$
$$f_{t+1}^{s,(j)} = h_x f_t^{s,(j)} + \frac{1}{2} h_{xx} \left(f_t^{f,(j)} \right)^2.$$

  2. Calculate weights:

$$\omega_t^{(j)} = p(y_t | f_t^{f,(j)}, f_t^{s,(j)}), \quad j = 1, \ldots, N.$$

- **Update step:**
  1. Define normalized weights: $\widetilde{w}_t^{(j)} = \frac{\omega_t^{(j)} w_{t-1}^{(j)}}{\frac{1}{N}\sum \omega_t^{(j)} w_{t-1}^{(j)}}$.
  2. Resample from multinomial distribution $\left\{\omega_t^{(j)}, \widetilde{w}_t^{(j)}\right\}$ and set $w_t^{(j)} = 1$.
- **Compute conditional likelihood:**

$$p(y_t | Y_{1:t-1}) \approx \frac{1}{N} \sum_{i=1}^{N} \omega_t^{(j)} w_{t-1}^{(j)}. \qquad (12)$$

The overall likelihood is then $p(y|\Theta_i^{prop}, \Theta_{-i}^{curr}) = \prod_{t=1}^{T} p(y_t | Y_{1:t-1})$.

Appendix: Definitions of Shortfall and Longrise

Shortfall:

$$SF_{\alpha}(h) = E_t\left[x_{t+h} | x_{t+h} < q_{\alpha}(x_{t+h})\right]$$

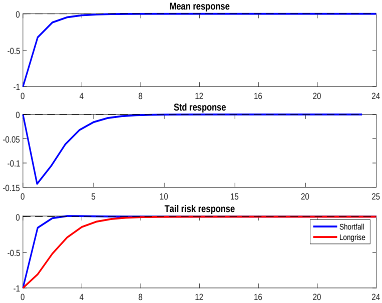
Longrise:

$$LR_{\alpha}(h) = E_t\left[x_{t+h} | x_{t+h} > q_{1-\alpha}(x_{t+h})\right]$$

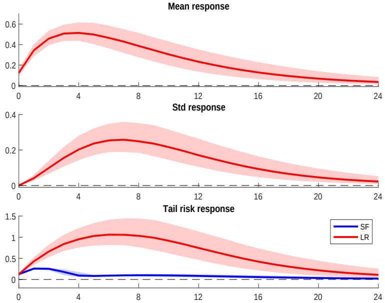
- h is horizon
- q_{α} is the α quantile of the distribution

Return

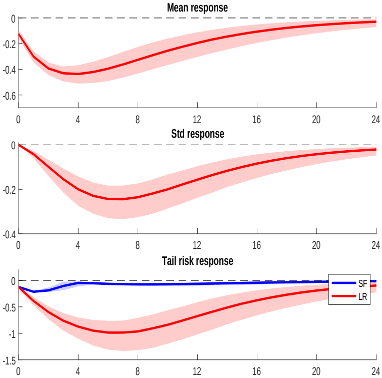
Appendix: Higher-Order Moment Responses, Negative Shock



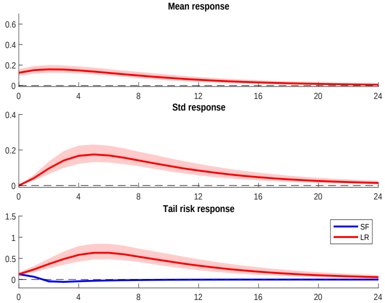
Appendix: Higher-Order Moment Responses, High risk EDC



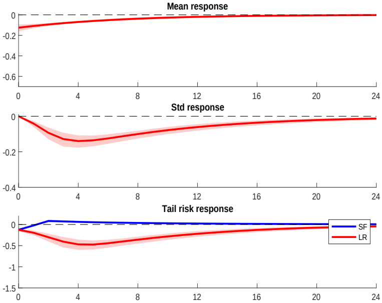
Appendix: Higher-Order Moment Responses, High risk EDC



Appendix: Higher-Order Moment Responses, Low risk



Appendix: Higher-Order Moment Responses, Low risk



Appendix: Parameter Estimates

Table 3: MCMC estimates

	Linear	Non-linear
h_x	0.972	0.847
	(0.952 , 0.992)	(0.818 , 0.876)
h_{xx}	0.000	2.609
	(0.000 , 0.000)	(2.268 , 3.167)
σ	0.213	0.126
	(0.182 , 0.243)	(0.093 , 0.161)
constant	0.000	-0.850
	(0.000 , 0.000)	(-0.995 , -0.755)
g_1	1.000	1.000
	(1.000 , 1.000)	(1.000 , 1.000)
g_2	1.070	1.087
	(1.016 , 1.174)	(0.990 , 1.204)
g_3	1.008	0.972
	(0.945 , 1.119)	(0.900 , 1.141)
g_4	1.028	0.946
	(0.918 , 1.116)	(0.864 , 1.157)
g_5	0.982	0.983
	(0.889 , 1.081)	(0.860 , 1.121)
η_1	0.416	0.450
	(0.374 , 0.487)	(0.354 , 0.520)
η_2	0.193	0.122
	(0.136 , 0.239)	(0.109 , 0.239)
η_3	0.382	0.422
	(0.345 , 0.432)	(0.341 , 0.464)
η_4	0.406	0.426
	(0.365 , 0.469)	(0.363 , 0.508)
η_5	0.476	0.410
	(0.422 , 0.533)	(0.400 , 0.542)
Log likelihood	-478.375	-438.474

Maximum likelihood estimates and 95% confidence intervals.

Appendix: Particle Filter

Bootstrap filter version (Sarkka, 2013)

- Prediction

  Given particles and weights at $t-1$: $\left\{x_{t-1}^{i}, W_{t-1}^{i}\right\}$

  1. Draw a new particle $x_t^{(i)}$ for each point in the sample set $\{x_{t-1}^{(i)} : i = 1, \dots, N\}$ from

  $$x_t^{(i)} \sim p(x_t|x_{t-1}^{(i)}), \quad i = 1, \dots, N$$

  2. Calculate weights:

  $$\omega_t^{(i)} = p(y_t|x_t^{(i)}), \quad i = 1, \dots, N$$

- Update

  1. Define normalized weights: $\widetilde{W}_t^{(i)} = \frac{\omega_t^{(}i) W_{t-1}^{(i)}}{\frac{1}{N}\sum \omega_t^{(i)} W_{t-1}^{(i)}}$.
  2. Resample from multinomial distribution $\left\{\omega_t^{(i)}, \widetilde{W}_t^{(i)}\right\}$ and set $W_t^{(i)} = 1$.

Approximate state distribution and likelihood are:

$$p(x_t|Y_t) \approx \sum^{N} \omega^{(i)} \delta(x_t - x^{(i)}) \quad p(y_t|Y_{t-1}) \approx \frac{1}{N} \sum^{N} \omega^{i} W^{i} \quad (13)$$