Financial and Macroeconomic Data Through the Lens of a Nonlinear Dynamic Factor Model*

Pablo A. Guerrón Quintana† Alexey Khazanov‡ Molin Zhong§

May 15, 2023

Abstract

Through the lens of a nonlinear dynamic factor model, we study the role of exogenous shocks and internal propagation forces in driving the fluctuations of macroeconomic and financial data. The proposed model (1) allows for nonlinear dynamics in the state and measurement equations; (2) can generate asymmetric, state-dependent, and size-dependent responses of observables to shocks; and (3) can produce time-varying volatility and asymmetric tail risks in predictive distributions. We find evidence in favor of nonlinear dynamics in two important U.S. applications. The first uses interest rate data to extract a factor allowing for an effective lower bound and nonlinear dynamics. Our estimated factor coheres well with the historical narrative of monetary policy. We find that allowing for an effective lower bound constraint is crucial. The second recovers a credit cycle. The nonlinear component of the factor boosts credit growth in boom times while hindering its recovery post-crisis. Shocks in a credit crunch period are more amplified and persist for longer compared with shocks during a credit boom.

Keywords— Interest rates, effective lower bound, credit cycle, asymmetric dynamics, predictive distributions, tail risk

*We thank Todd Clark and Taeyoung Doh for insightful comments and suggestions. We also benefited from comments by Boragan Aruoba, Mark Bognanni, Thorsten Drautzburg, Jesus Fernandez-Villaverde, Andrew Foerster, Manuel Gonzales-Astudillo, Luca Guerrieri, Arthur Lewbel, Michele Modugno, Juan Rubio-Ramirez, Frank Schorfheide, Dongho Song, and seminar participants at the Board of Governors, European Central Bank, Bank of England, CEF 2021, IAAE 2021, NBER-NSF SBIES 2021, NBER-NSF Time Series 2021, System Econometrics Meetings 2021, for useful comments. All errors are ours. This paper does not necessarily reflect the views of the Federal Reserve System or its Board of Governors.

†Boston College, email: pguerron@gmail.com.

‡The Hebrew University of Jerusalem, email: alexey.khazanov@mail.huji.ac.il.

§Board of Governors, email: Molin.Zhong@frb.gov.

1 Introduction

What are the drivers of financial and macroeconomic variables – external shocks or endogenous propagation? This question is timeless in economics. Understanding the origins of economic fluctuations becomes ever more important during downturns because of their implications for the scope of economic policies. To address this question, dynamic factor models (DFMs) have played a crucial role (Stock and Watson, 2016). Since their introduction in the 1970s, DFMs have relied mostly on a vector autoregressive (VAR) representation that imposes a linear relation between the factor today and its past, as well as between the factor and the observables.[1] However, substantial evidence shows that economic variables possess significant nonlinearities (Baker et al., 2016, Fernández-Villaverde et al., 2015b, Justiniano and Primiceri, 2008). Furthermore, financial data are more prone to sudden changes, particularly so during times of crises (Gilchrist and Zakrajsek, 2012, Ludvigson et al., 2021). In this paper, we study macroeconomic and financial data through the lens of a DFM that incorporates nonlinearities in the measurement and state equations.

The nonlinearities in the DFM allow us to exploit many features emphasized in the recent macroeconomic and financial literature that previous work on factor models has largely left unexplored. For example, we can use our new nonlinear dynamic factor (NLDF) model to examine the importance of nonlinear dynamics in the state equation during moments of high volatility in the economy such as the Global Financial Crisis (GFC); to construct point and predictive density estimators in the presence of nonlinearities; and to study the truncated relation between factors and observables as in the shadow interest rate literature (Wu and Xia, 2016). Importantly, we rely on a coherent framework to simultaneously study all of these forces together.

Our nonlinear factor model is inspired by the pruned second order state-space (2nd SS) model discussed in Kim et al. (2008) and Andreasen et al. (2017) as the approximate solution of a nonlinear dynamic stochastic general equilibrium (DSGE) model. As discussed in Fernandez-Villaverde et al. (2016) and the references therein, an important feature of the pruned solution is that it can capture several types of nonlinearities with reasonable accuracy, a feature that we aim to exploit in this paper. We re-interpret the 2nd SS framework in the context of a DFM whose factor evolves according to the state-space model's state equation. We allow the measurement equation to be potentially nonlinear if economic theory suggests it. This adjustment accommodates situations where the observables are bounded and allows for the presence of non-additive measurement errors.

The nonlinear state dynamics generate novel implications both in terms of the impulse response functions (IRFs) and the predictive densities. Although the nonlinearities are specified at the factor

[1] The standard representation follows

$$Y_t = \Lambda F_t + e_t; \quad F_t = \Psi(L) F_{t-1} + v_t.$$

Here, the first and second expressions correspond to the measurement and state equations, respectively. The shocks e_t and v_t are assumed normally distributed and independent over time and cross-sectionally. L is the lag operator.

level, they are passed through to the observable variables via the measurement equation. The IRFs produced by the model have three interesting properties. First, they are asymmetric, meaning that positive shocks produce differently shaped IRFs compared with same-sized negative shocks. Second, they are state dependent, meaning that the shapes and magnitudes of the IRFs are different depending on initial conditions. Third, they are size dependent, meaning that one standard deviation shocks generate different shapes in the IRFs compared with two standard deviation shocks. The model has implications for higher-order moments as well, leading to rich distributional dynamics. Despite homoskedastic and normally distributed innovations, the unconditional distribution of the states is non-normal because of the asymmetries implied by the model. Moreover, the model generates predictive densities that have time-varying volatility and asymmetric tail risk movements. The latter fact is a key property documented by Adrian et al. (2019) in the macro data.

With the NLDF model in hand, we analyze the role of external forces and internal propagation in two macroeconomic and financial cases. In our first case, we estimate the shadow interest rate model along the lines of Wu and Xia (2016). Our exercise extracts a common factor from a series of U.S. forward rates while respecting the effective lower bound (ELB) in the short-maturity rates. There are two key differences from Wu and Xia (2016). First, we model the yields in first differences, following the recommendations of Onatski and Wang (2021) and Crump and Gospodinov (2022). Second, we investigate the possibility of nonlinear factor dynamics in the interest rates. This exercise is motivated by the literature debating whether there were structural changes in the behavior of longer-term yields brought about by the ELB constraining monetary policy (Swanson and Williams, 2014). Our nonlinear factor model provides one avenue to test this question empirically because it allows for state-dependence.

We find that allowing for the ELB constraint is crucial, both in terms of estimating a yield curve factor that coheres with the historical narrative of monetary policy and in model fit. This result is in line with results found in Wu and Xia (2016). The presence of the ELB affects the estimation results in two time periods. The first is in the late 2003 and early 2004 period, when the fed funds rate declined to 1 percent. The second is the long spell in the zero lower bound that began in the GFC. After allowing for the ELB constraint, however, we do not find much evidence of nonlinear factor dynamics, which suggests that the dynamics of the entire yield curve did not appreciably change upon entering the ELB period. The key nonlinearity to account for is the lower bound constraint for the shorter-maturity yields.

Our second case estimates a nonlinear credit cycle as the common component of U.S. credit growth dynamics across four sectors: nonfinancial business, household, financial, and public sectors. The investigation is motivated by the extensive theoretical and empirical literature documenting the importance of credit growth and leverage in understanding the credit cycle and especially nonlinear amplification of shocks (Bernanke et al., 1999, Brunnermeier and Sannikov, 2014, Schularick and Taylor, 2012). We investigate the presence of common nonlinear dynamics across sectors in a time series context. Our nonlinear factor captures the slow rise and rapid declines common

across credit sectors. The results highlight the importance of a slow-moving second-order factor that boosts credit growth in boom times – notably beginning in the mid 1990s until the onset of the GFC. Its collapse in the GFC plays a key role in the sluggish post-crisis recovery of credit. The effects of the same-sized shock to the credit cycle factor are different depending on the state of the economy. Shocks in a credit crunch period are more amplified and persist for longer compared with shocks during a credit boom. Negative shocks lead to increases in the standard deviation of the credit factor predictive distribution. The combination of the decline in mean and increase in volatility generates larger movements in downside tail risk compared with upside risk.

Related Literature. The literature has considered some forms of nonlinearities in factor models like Markov switching, time-varying parameters, and stochastic volatility. However, to the best of our knowledge, there is no work on models that allows for second-order dynamics in the state equation or general nonlinearities in the measurement equation. The classic example these days of a nonlinearity in the data is the zero lower bound imposed on short-term interest rates. In general, we find lower and upper bounds when we deal with percentages like labor market tightness, transition probabilities, and job-finding and separation rates, so investigating a factor model that can deal with these situations is important.

Our work is particularly close to two recent papers in the literature. First, Aruoba et al. (2017a) introduce the quadratic autoregressive process (QAR), which allows quadratic terms in lagged regressors as well as GARCH (general autoregressive conditional heteroskedasticity) features. Like our approach, they rely on the pruned representation to generate a stable model, but their study concentrates on univariate models and posits that the observables follow the QAR. In contrast, our factor model framework can be modified to admit different classes of nonlinearities like the one introduced by the ELB.

Second, Gorodnichenko and Ng (2017) use the insights from the second order solution of DSGE models to obtain restrictions on the dynamics of observables and its squared values. From this analysis, the authors extract a factor that mimics the dynamics of a level state variable and another one that displays stochastic volatility features. There are important differences between our papers. While the approach of Gorodnichenko and Ng (2017) is based on the approximated solution proposed by Benigno et al. (2013), our representation arises from the perturbed solution of a nonlinear DSGE model (Andreasen et al., 2017). They extract factors based on singular value decomposition. In contrast, we estimate the nonlinear system using likelihood-based methods, which allows us, among other things, to build predictive densities, filter the most likely state of the economy, or report IRFs conditional on the state of the economy. Finally, the methodology of Gorodnichenko and Ng (2017) requires that observables be measured without error and rules out the possibility of kinks in the data.

More broadly, we contribute to the large literature on factor models. Stock and Watson (2016) give an in-depth review of linear factor models. Among the recent advances, Banbura and Mod-

ugno (2014) allow for missing data with arbitrary patterns in estimating linear factor models by using an expectations maximization algorithm. Chauvet (1998) uses a linear factor model with regime switches to estimate business cycles. Aruoba and Diebold (2010) leave nonlinear factors as a to-do task, although with a focus on Markov switching regimes rather than the type we propose. Shintani (2005) estimates a nonparametric diffusion model for forecasting Japanese data. Chen et al. (2021) analyze a semiparametric panel data model in which latent factors are modeled in a nonparametric fashion. Cheng et al. (2016) propose a linear DFM that allows for breaks in loadings and the number of factors, which is an alternative view of the world. They find that the Great Recession led to a change in the factor loadings and the emergence of a new factor. Carrasco and Rossi (2016) consider forecasting with misspecified factor models.

Finally, our study is also related to work that departs from Gaussian shocks (Gourieroux et al., 2019). Aruoba et al. (2021) estimate a structural VAR model that allows coefficients to switch depending on whether the economy is at the ELB. The results on the distributional implications of the NLDF also connect us to a growing literature on tail risks and distributional asymmetries (Adrian et al., 2019).

The rest of the paper is organized as follows. The next section discusses the NLDF model using a simple example with two observables. We motivate the factor model by connecting it to the pruned solution of a nonlinear DSGE model. In Section 3, we highlight the novel implications of the nonlinear DFM for moments, IRFs, and predictive densities. Section 4 discusses our two empirical applications. Some concluding remarks are in the final section.

2 The Nonlinear Dynamic Factor Model

In this section, we introduce the NLDF model and then discuss the motivation for using our factor dynamics. Next, we discuss possible specifications of the measurement equation. Finally, we close the section by presenting our estimation algorithms.

2.1 Model Specification

We consider the following NLDF model:

$$\text{Measurement: } y_t = \mathcal{G}(f_t) + \eta\epsilon_t \tag{1}$$

$$\text{Factor dynamics: } f_t = \mathcal{H}(f_{t-1}) + \sigma\nu_t. \tag{2}$$

Here, ϵ_t is an $N\times 1$ vector of $iid\ N(0, I_N)$ innovations, ν_t is a $K\times 1$ vector of $iid\ N(0, I_K)$ innovations, y_t is an $N \times 1$ vector of observed variables, and f_t is the $K \times 1$ underlying factor. $\mathcal{G}(\cdot)$ and $\mathcal{H}(\cdot)$ are general, possibly nonlinear functions. In addition, we assume that the $\mathcal{H}(\cdot)$ function is at least twice differentiable. η is an $N \times N$ diagonal matrix of standard deviations, and σ is a $K \times K$ matrix that is the square root of a variance-covariance matrix. The additive measurement error assumption is for

ease of exposition. Our framework can easily handle multiplicative errors, like in Hwang (1986), or nonadditive errors.

For empirical applications involving data such as unemployment, gross domestic product (GDP) growth, or inflation rates, a typical desirable feature of the factor process is stationarity. A generic $\mathcal{H}(\cdot)$ function, estimated using a limited data range, may imply explosive dynamics of the factor and, thus, of the observables. To avoid this problem, we use the pruned motion equation (Kim et al., 2008, Andreasen et al., 2017) that has easily verifiable stationarity conditions.

To make the analysis more concrete, consider an NLDF model in which the measurement equation is linear in a single underlying factor and we take a pruned second-order approximation to the function $\mathcal{H}(\cdot)$. We adopt the single factor specification for the rest of the paper. Let f_t denote the underlying factor and f_t^f and f_t^s its first- and second-order terms such that $f_t = c + f_t^f + f_t^s$. Then the pruned system is

$$
\begin{aligned}
&y_t = G f_t + \eta \epsilon_t \\
&\begin{cases}
f_t = c + f_t^f + f_t^s \text{ – first- and second-order factors} \\
f_t^f = h_x f_{t-1}^f + \sigma \nu_t \\
f_t^s = h_x f_{t-1}^s + \frac{1}{2} h_{xx} \left(f_{t-1}^f \right)^2 .
\end{cases}
\end{aligned}
\tag{3}
$$

The first-order term follows the same process as a linear DFM with the persistence parameter governed by h_x. The exogenous shocks ν_t perturb the first order term on impact. The second-order term depends on the square of the lagged first-order term, with h_{xx} modulating the importance of this relationship. The second-order term also has persistence determined by h_x. The exact structure of this process is discussed in detail as the second-order solution to a dynamic equilibrium model that prevents explosive paths (Andreasen et al., 2017).

This structure has two main attractive properties. First, the model allows for rich nonlinearities due to the presence of the second-order term. Specifically, the model can generate asymmetric, state-dependent, and size-dependent IRFs. Moreover, the model can generate time-varying volatility through the state dependence. We will illustrate these properties in detail in Section 3. Second, the model has easily verifiable stationarity conditions. As long as $|h_x| < 1$, the model is stationary.

The presence of the quadratic term h_{xx} introduces an additional distinction relative to the linear factor model. Even if the shocks have zero mean, the factor has a mean different from zero. This property can be verified by applying the expectation operator on the second-order term in Equation 3. In the dynamic equilibrium models literature, h_{xx} makes the model's deterministic steady state different from its stochastic steady state. We include a constant c in the factor's law of motion to adjust the overall factor to have zero mean in our applications, although this parameter may alternatively be estimated.

2.2 Motivating the Nonlinear Dynamic Factor Model

Our time series model is closely connected to the nonlinear solution of a DSGE model, which we view as an important strength of our framework. Nonlinear DSGE models have rapidly grown in popularity, spurred on by an ample body of empirical research that documents that the U.S. economy has nonlinear features such as stochastic volatility (Justiniano and Primiceri, 2008, Fernández-Villaverde and Rubio-Ramírez, 2007, Bloom, 2009, Fernández-Villaverde et al., 2015a), time-varying monetary policy (Fernández-Villaverde et al., 2015b), and the zero lower bound on short-term interest rates (Fernandez-Villaverde et al., 2015, Gust et al., 2017, Wu and Xia, 2016).

Our NLDF model is the direct time series analogue of the pruned second-order perturbation DSGE solution. Suppose the DSGE modle has only one state variable (denoted f_t). Then, its dynamic equation is approximated by

$$\begin{aligned} \text{State equation: } f_t &= f_t^f + f_t^s \\ f_t^f &= h_1 f_{t-1}^f + \sigma \nu_t \\ f_t^s &= h_1 f_{t-1}^s + \frac{1}{2} h_2 \left(f_{t-1}^f \right)^2 . \end{aligned} \tag{4}$$

Here, h_1 and h_2 are coefficients and ν is a normally distributed innovation. Comparing Equations 3 and 4, we see that their structures are the same. The key difference is that in the DSGE model solution, h_1 and h_2 are known given the deep parameters of the model. In our time series model, the corresponding parameters are estimated from the data.

If the researcher believes that the fundamental driver of the data is the one in Equation 4, it seems natural to advocate for the extraction of factors based on an approach that departs from linearity. Incorrectly assuming a linear factor model, and thereby ignoring the f_t^s term, results in an estimated factor driven by counterfactually volatile shocks – that is, the researcher would conclude that fluctuations are in a large part due to exogenous events as opposed to endogenous propagation.

2.3 Specification of the Measurement Equation

Given the nonlinearities modeled in the latent factor, a natural benchmark case is for the measurement equation to be linear, as in Equation 3. The linear measurement equation allows us to prove that the latent factor is not identified and an additional normalization is needed. We close the section by discussing some nonlinear measurement equation extensions.

Identification with Linear Measurement Equation We begin by discussing an identification issue with the linear measurement equation case. To show the issue, it is convenient to use a two-observable version of our model with a single factor:

$$
\begin{aligned}
\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} &= \begin{pmatrix} G_1 \\ G_2 \end{pmatrix} f_t + \eta \epsilon_t, \\
f_t &= c + f_t^f + f_t^s, \\
f_t^f &= h_x f_{t-1}^f + \sigma \nu_t, \\
f_t^s &= h_x f_{t-1}^s + \frac{1}{2} h_{xx} \left(f_{t-1}^f \right)^2 .
\end{aligned} \tag{5}
$$

In this system, the unknown parameters are the loading components G_1 and G_2, the factor's linear and quadratic terms h_x and h_{xx}, the standard deviation of the shock σ, and the standard deviations of the measurement errors η.

Proposition 2.1. The sign and scale of the factor in the nonlinear factor model in Equation 5 are not identified.

Proof. Consider the following constant $n \neq 0$ and scale the system in Equation 5 as follows:

$$
\begin{aligned}
\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} &= \begin{pmatrix} G_1 \\ G_2 \end{pmatrix} \left(\frac{1}{n} \right) n f_t + \eta \epsilon_t \\
n f_t &= nc + n f_t^f + n f_t^s \\
n f_t^f &= h_x n f_{t-1}^f + n \sigma v_t \\
n f_t^s &= h_x n f_{t-1}^s + n \frac{1}{2} h_{xx} \left(\frac{1}{n} \right)^2 \left(n f_{t-1}^f \right)^2 .
\end{aligned}
$$

Next, define $\widetilde{f}_t = n f_t$, $\widetilde{G}_1 = \left(\frac{1}{n}\right) G_1$, $\widetilde{G}_2 = \left(\frac{1}{n}\right) G_2$, $\widetilde{\sigma} = n\sigma$, $\widetilde{h}_{xx} = \left(\frac{1}{n}\right) h_{xx}$, $\widetilde{c} = nc$ and rewrite the system as follows:

$$
\begin{aligned}
\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} &= \begin{pmatrix} \widetilde{G}_1 \\ \widetilde{G}_2 \end{pmatrix} \widetilde{f}_t + \epsilon_t \\
\widetilde{f}_t &= \widetilde{c} + \widetilde{f}_t^f + \widetilde{f}_t^s \\
\widetilde{f}_t^f &= h_x \widetilde{f}_{t-1}^f + \widetilde{\sigma} \nu_t \\
\widetilde{f}_t^s &= h_x \widetilde{f}_{t-1}^s + \frac{1}{2} \widetilde{h}_{xx} \left(\widetilde{f}_{t-1}^f \right)^2
\end{aligned} \tag{6}
$$

This formulation shows that the "tilde" model produces exactly the same observables as generated by the baseline model. Therefore, not all parameters in the model are identified. Moreover, as n can be negative, we do not have sign or magnitude identification. □

In our applications, we fix $G_1 = 1$, which corresponds to the named factor approach in the DFM literature (Stock and Watson (2016)).

Nonlinear Measurement Equation. Although the linear measurement equation is a leading case, in some instances, economic theory may suggest specifying nonlinearities in the measurement equation. Our model can accommodate these more complex dynamics. For instance, one could allow for a fully nonlinear measurement equation:

$$
\begin{aligned}
\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} &= \mathcal{G}\left(\begin{bmatrix} f_t, \eta_1 \epsilon_{1,t} \\ f_t, \eta_2 \epsilon_{2,t} \end{bmatrix}\right), \\
f_t &= c + f_t^f + f_t^s, \\
f_t^f &= h_x f_{t-1}^f + \sigma \nu_t, \\
f_t^s &= h_x f_{t-1}^s + \frac{1}{2} h_{xx} \left(f_{t-1}^f\right)^2.
\end{aligned}
$$

Here, $\mathcal{G}$ is the nonlinear function mapping from measurement errors and factors to observables.

In our shadow interest rates application, we specify a condition that restricts the level of interest rates from going below a lower bound via a nonlinear measurement equation, consistent with the zero lower bound restriction on nominal rates.

2.4 Estimation Algorithms

We use Bayesian methods to estimate the model. As our model is nonlinear, we rely on particle filtering methods to approximate the likelihood (Särkkä, 2013). We use two algorithms in our empirical illustrations: a Metropolis Hastings combined with the bootstrap particle filter and a Gibbs sampling combined with the particle smoother. We believe each has their strengths. The particle filtering algorithm readily delivers the filtered factor, which may be useful in situations where maintaining the information structure of the filtered variable is important, such as when the filtered variable is included in a VAR (Fernández-Villaverde et al., 2015a). A caveat of the bootstrap particle filter is that it demands hundreds of thousands of particles to characterize accurately the likelihood function. The particle Gibbs sampling algorithm delivers the smoothed estimate, which is the most accurate estimate of the factor given all of the data. Also, through its exploitation of ancestor sampling (Lindsten et al., 2014), the algorithm has good mixing properties even with relatively few particles – in the order of hundreds. The disadvantage is that the sampler is only approximate for our model, although Lindsten et al. (2014) show that its performance is still good.[2] Further details about both algorithms and their computational implementation can be found in Appendix Section A.

Monte Carlo. In Appendix Section B, we conduct a Monte Carlo exercise to study the estimation performance. First, we show that if the data-generating process is the nonlinear factor model itself,

[2]See the associated discussion in Section 7.2 of that paper.

our estimation strategy can recover the true parameter values. Second, we find that the likelihood implied by a linear factor model is below the likelihood from the nonlinear model if the data were generated from our NLDF model. By ignoring the nonlinear dynamics, the linear model tends to estimate an excessively persistent linear factor.

3 Properties of the Nonlinear Dynamic Factor Model

We now move on to some key properties generated by the NLDF model. Our focus is on the latent factor with the understanding that these properties propagate through to the observables via the measurement equation.[3] We focus on three novel features that our NLDF model brings to the table: asymmetric responses to positive versus negative shocks, state-dependent responses, and size-dependent responses. These properties of IRFs were previously discussed in the context of solution methods to DSGE models in Andreasen et al. (2017), but we find it instructive to review them here. A linear DFM cannot deliver these types of IRFs.

3.1 Analytical Properties of the Model

We begin by discussing the analytical properties of the model. We can write the factor dynamics in a useful state-space form, first presented in Andreasen et al. (2017), to make analytical progress on the model's implications for the moments of the factors. To facilitate the exposition, let us continue to assume a one-dimensional factor. Then one can write the factor dynamics as follows:

$$f_t = c + f_t^f + f_t^s \tag{7}$$

$$\begin{pmatrix} f_t^f \\ f_t^s \\ \left(f_t^f\right)^2 \end{pmatrix} = \begin{pmatrix} h_x & 0 & 0 \\ 0 & h_x & \frac{1}{2}h_{xx} \\ 0 & 0 & h_x^2 \end{pmatrix} \begin{pmatrix} f_{t-1}^f \\ f_{t-1}^s \\ \left(f_{t-1}^f\right)^2 \end{pmatrix} + \begin{pmatrix} \sigma & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \sigma^2 & 2\sigma h_x \end{pmatrix} \begin{pmatrix} \epsilon_t \\ \epsilon_t^2 \\ f_{t-1}^f \epsilon_t \end{pmatrix}. \tag{8}$$

For ease of notation, we write the state space as follows:

$$z_t = Az_{t-1} + B\zeta_t, \tag{9}$$

where $z_t = \begin{pmatrix} f_t^f \\ f_t^s \\ \left(f_t^f\right)^2 \end{pmatrix}$ and $\zeta_t = \begin{pmatrix} \epsilon_t \\ \epsilon_t^2 \\ f_{t-1}^f \epsilon_t \end{pmatrix}$. The matrices A and B contain the corresponding parameters in the state equation. It is possible to show that the innovations ζ_t are intertemporally

[3] The relationship between the properties of the factor and the observables is most straightforward in the benchmark case of a linear measurement equation. With a nonlinear measurement equation, there will, in general, be a nonlinear transformation of the latent factor to the observable variables.

uncorrelated – in other words, $E\left[\zeta_{t+i}\zeta_{t+j}\right] = 0$ for $i \neq j$. In this section, when we discuss conditional moments at time t, the conditioning set is the past history of factors $\{f_t^f, f_t^s, f_{t-1}^f, f_{t-1}^s, ...\}$.

First Moment Dynamics. We begin by discussing first moment dynamics. We emphasize four parts: the persistence of the factor, asymmetry in the impulse response to a positive versus negative shock, state dependence in the response, and size dependence in the response.

The overall factor has persistence from both the first- and second-order factors. The first-order factor, f_t^f, has a persistence equal to h_x. The second-order factor, f_t^s, has a persistence greater than h_x. The second-order factor has more persistence because while f_t^s follows an autoregressive process with parameter h_x, its "innovation" depends on $\left(f_{t-1}^f\right)^2$, which itself is persistent with parameter h_x^2. As long as $h_x > 0$, which is usually the case for macroeconomic and financial data, the second-order factor is more persistent than the first-order one.

The ϵ_t^2 component of the innovation generates an asymmetric response to a positive versus negative shock because while a positive shock increases ϵ_t – and therefore f_t^f – and a negative shock decreases it, both a positive and negative shock increase ϵ_t^2. The important parameter governing the direction of the asymmetry is h_{xx}. This parameter governs how $\left(f_{t-1}^f\right)^2$ relates to f_t^s and therefore how the effects of ϵ_t^2 pass through to the overall factor. If the sign of h_{xx} is positive, then a positive shock increases f_t^f and the response of ϵ_t^2 increases f_t^s. A negative shock decreases f_t^f, but the response of ϵ_t^2 still increases f_t^s. The effect is reversed if the sign of h_{xx} is negative.

The impulse response is state dependent, which comes from the $f_{t-1}^f \epsilon_t$ term in the innovation. The sign of f_{t-1}^f determines the effect of a shock to ϵ_t on $\left(f_t^f\right)^2$. Additionally, the magnitude of f_{t-1}^f determines the amount of time-varying volatility, which we discuss more in the next subsection.

The ϵ_t^2 term also creates size dependencies in the response to a shock, meaning that a two standard deviation shock does not generate double the responses of a one standard deviation shock. This effect follows straightforwardly from the quadratic transformation. This feature of our model is important because it can generate strong amplification to shocks during downturn episodes.

The general formula for conditional mean dynamics at horizon h is shown in Equation 10. On top of the previous discussion, there are two additional points to mention from this equation. First, the h_x term determines the persistence property of the entire system, as is expected. If $|h_x| < 1$, the conditional mean responses converge. Second, the ϵ_t^2 innovation produces a nonzero long-run mean for $\left(f_t^f\right)^2$.

$$E_t z_{t+h} = A^h z_t + \sum_{i=0}^{h-1} A^i \begin{pmatrix} 0 \\ 0 \\ \sigma^2 \end{pmatrix} \tag{10}$$

Volatility Dynamics. The model generates time-varying volatility via the state dependence inherent in the second-order factor. The magnitude of f_{t-1}^f modulates the effect of ϵ_t on $\left(f_t^f\right)^2$. A

larger value of f_{t-1}^f means that the same-sized shock generates a larger response in $\left(f_t^f\right)^2$. Intuitively, the quadratic component of the model responds by more the further f_{t-1}^f is away from 0.

$$V_t\left(z_{t+h}\right) = \sum_{i=0}^{h-1} A^i B V_t\left(\zeta_{t+h-i}\right) B' \left(A'\right)^i \tag{11}$$

$$V_t\left(\zeta_{t+h}\right) = \begin{pmatrix} 1 & 0 & h_x^{h-1} f_t^f \\ 0 & 2 & 0 \\ h_x^{h-1} f_t^f & 0 & E_t\left(\left(f_{t+h-1}^f\right)^2\right) \end{pmatrix} \tag{12}$$

The formula for the h-step-ahead conditional variance of the system is shown in Equations 11 and 12. The latter equation shows that the $f_{t-1}^f \epsilon_t$ term generates time-varying volatility in the system. The conditional variance from this term depends on $\left(f_{t-1}^f\right)^2$ and therefore has a persistence of h_x^2. As Equation 11 shows, the conditional variance of z_t at various horizons is then a discounted sum of the conditional variance of ζ_t from $t+1$ up through $t+h$.

Relationship Between First and Second Moments. Our model generates a non-zero correlation between conditional first and second moments. This relationship can be seen by noticing that $\left(f_{t-1}^f\right)^2$ simultaneously determines the conditional mean of f_t^s and the conditional volatility of shocks to $\left(f_t^f\right)^2$. Simultaneous movements in mean and volatility are an important mechanism identified in the growth-at-risk literature to generate asymmetric tail risk behavior (Adrian et al., 2019).

More formally, we can examine the conditional covariances between f_t^f and $\left(f_t^f\right)^2$ and f_t^s.[4] We begin by discussing short-run conditional correlations and then move on to unconditional correlations. These expressions can be derived from the general formulas in Equations 11 and 12. At one step ahead, the conditional covariance between the first order factor and its square is

$$Cov_t\left(f_{t+1}^f, \left(f_{t+1}^f\right)^2\right) = 2\sigma^2 h_x f_t^f. \tag{13}$$

As suggested by the earlier discussion on conditional mean dynamics, the sign of f_t^f is important. Specifically, if f_t^f is positive, the conditional correlation between the two terms is positive as well and vice versa.[5] To understand why, note that the shock ϵ_{t+1} determines f_{t+1}^f, while ϵ_{t+1}^2 and $f_t^f \epsilon_{t+1}$ determine $\left(f_{t+1}^f\right)^2$ conditional on knowing f_t^f. The shocks ϵ_{t+1} and ϵ_{t+1}^2 are uncorrelated, so any nonzero covariance must come from $f_t^f \epsilon_{t+1}$. When f_t^f is positive, then further increases

[4] While we discuss the correlations between f_t^f and $\left(f_t^f\right)^2$ and f_t^s, keep in mind that $\left(f_{t-1}^f\right)^2$ affects the overall factor f_t.

[5] This result is assuming $h_x > 0$, which is the empirically realistic range.

in the first-order factor increase $\left(f_{t+1}^{f}\right)^{2}$ and the correlation between the first-order factor and its square is positive, whereas when f_t^f is negative, further decreases in the first-order factor increase $\left(f_{t+1}^{f}\right)^{2}$ and the correlation is negative.

At two steps ahead, the conditional covariance is:

$$Cov_t\left(f_{t+2}^{f},\left(f_{t+2}^{f}\right)^{2}\right)=2\sigma^{2}h_{x}^{2}\left(1+h_{x}^{2}\right)f_{t}^{f}. \tag{14}$$

Qualitatively, the same mechanisms are at play as in the one-step-ahead case. The correlation is positive if $f_t^f > 0$ and is negative otherwise because both the first-order factor and its square have persistent dynamics. For example, if f_t^f is positive, f_{t+1}^f is expected to remain positive and therefore continue producing a positive co-movement between the first-order factor and its square. The overall covariance is a sum of two terms because it takes into account shocks at $t+1$ and $t+2$.

The conditional covariance between the second-order factor and the first-order factor squared is 0 at one step ahead, as the second-order factor at time $t+1$ is predetermined given time t information. The conditional covariance becomes nonzero at two steps ahead and its value equals

$$Cov_t\left(f_{t+2}^{s},\left(f_{t+2}^{f}\right)^{2}\right)=\frac{1}{2}h_{xx}\underbrace{\left(E_t\left(\left(f_{t+1}^{f}\right)^{2}\left(f_{t+2}^{f}\right)^{2}\right)-E_t\left(\left(f_{t+1}^{f}\right)^{2}\right)E_t\left(\left(f_{t+2}^{f}\right)^{2}\right)\right)}_{\text{Conditional one-step ahead autocovariance of }\left(f_{t+1}^{f}\right)^{2}}. \tag{15}$$

The two terms are tightly related because the second-order factor directly loads onto past values of the first-order factor squared. As can be seen by examining Equation 15, the conditional covariance is determined by two components: h_{xx} and the conditional one-step-ahead autocovariance of $\left(f_{t+2}^{f}\right)^{2}$. If $h_{xx} > 0$, then increases in the first-order factor squared generate increases in the second-order factor, and the conditional covariance is positive. The effects are reversed if $h_{xx} < 0$. The autocovariance term appears because the time $t+2$ value of the second-order factor loads onto the time $t+1$ value of the squared first-order factor, so intertemporal dynamics play a role.

We compare the short-run co-movement behavior with unconditional co-movements. The unconditional variance-covariance matrix of the system is given by Equations 16 through 18:

$$V\left(z\right)=AV\left(z\right)A'+BV\left(\zeta\right)B' \tag{16}$$

$$V(\zeta)=\begin{pmatrix}1 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & E\left(f^{f}\right)^{2}\end{pmatrix} \tag{17}$$

$$E(z) = (I - A)^{-1} \begin{pmatrix} 0 \\ 0 \\ \sigma^2 \end{pmatrix}. \tag{18}$$

The general form of $V(z)$ is as follows:

$$V(z) = \begin{pmatrix} X & 0 & 0 \\ 0 & X & X \\ 0 & X & X \end{pmatrix}, \tag{19}$$

where the X denotes nonzero values. Unconditionally, there is no correlation between the first-order factor and its square, or between the first- and second-order factors. Although the first-order factor determines the time-varying volatility in the system, it unconditionally has zero correlation with volatility because the time-varying volatility depends only on the magnitude of f^f_{t-1}, not its sign. As the unconditional distribution of f^f_t is symmetric around zero, this correlation is also zero unconditionally.

There is dependence, however, between the second-order factor and the volatility in the system, which induces a relationship between level and volatility, even unconditionally. This dependence arises because the first-order factor squared enters as the driving force of the second-order factor. Unsurprisingly, the sign of h_{xx} is important in governing this relationship, with a positive h_{xx} generating a positive dependence between level and volatility and a negative h_{xx} generating a negative dependence.

3.2 Simulations

We use a parameterized version of the nonlinear model and simulation methods to further illustrate its properties. To this end, let us consider the baseline model with two observables:

$$\begin{aligned} \begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} &= \begin{pmatrix} G_1 \\ G_2 \end{pmatrix} f_t + \eta \epsilon_t \\ f_t &= c + f^f_t + f^s_t \\ f^f_t &= h_x f^f_{t-1} + \sigma v_t \\ f^s_t &= h_x f^s_{t-1} + \frac{1}{2} h_{xx} \left(f^f_{t-1} \right)^2, \end{aligned} \tag{20}$$

The baseline parameterization is $c = -\frac{1}{2}\frac{h_{xx}\sigma^2}{(1-h_x)(1-h_x^2)}$, $G_1 = 1$, $G_2 = 2$, $h_x = 0.45$, $h_{xx} = 0.5$, and $\sigma = 1$. The setting of c guarantees that the overall factor f has zero mean. We start our discussion with the IRFs implied by the model. Then, we move on to the distributional implications of the

shocks. A key parameter in the analysis is h_{xx}. Its sign determines the direction of the asymmetry. We set $h_{xx} > 0$, but if h_{xx} were to be negative, then the asymmetries would be flipped.

Impulse Response Functions. We focus on the three novel properties that our factor model can deliver: asymmetric, state-dependent, and size-dependent shocks.

As our model is nonlinear and state dependent, there are two main issues when computing IRFs. First, as discussed in Koop et al. (1996) and Goncalves et al. (2021), in the presence of nonlinearities, the different notions of IRFs do not necessarily coincide. We use the definition of IRFs suggested by Goncalves et al. (2021):

$$IRF_{\delta,t-1}\left[f_{t+h}\right] = E\left[f_{t+h}\left(\delta\right) - f_{t+h}|\Omega_{t-1}\right], \tag{21}$$

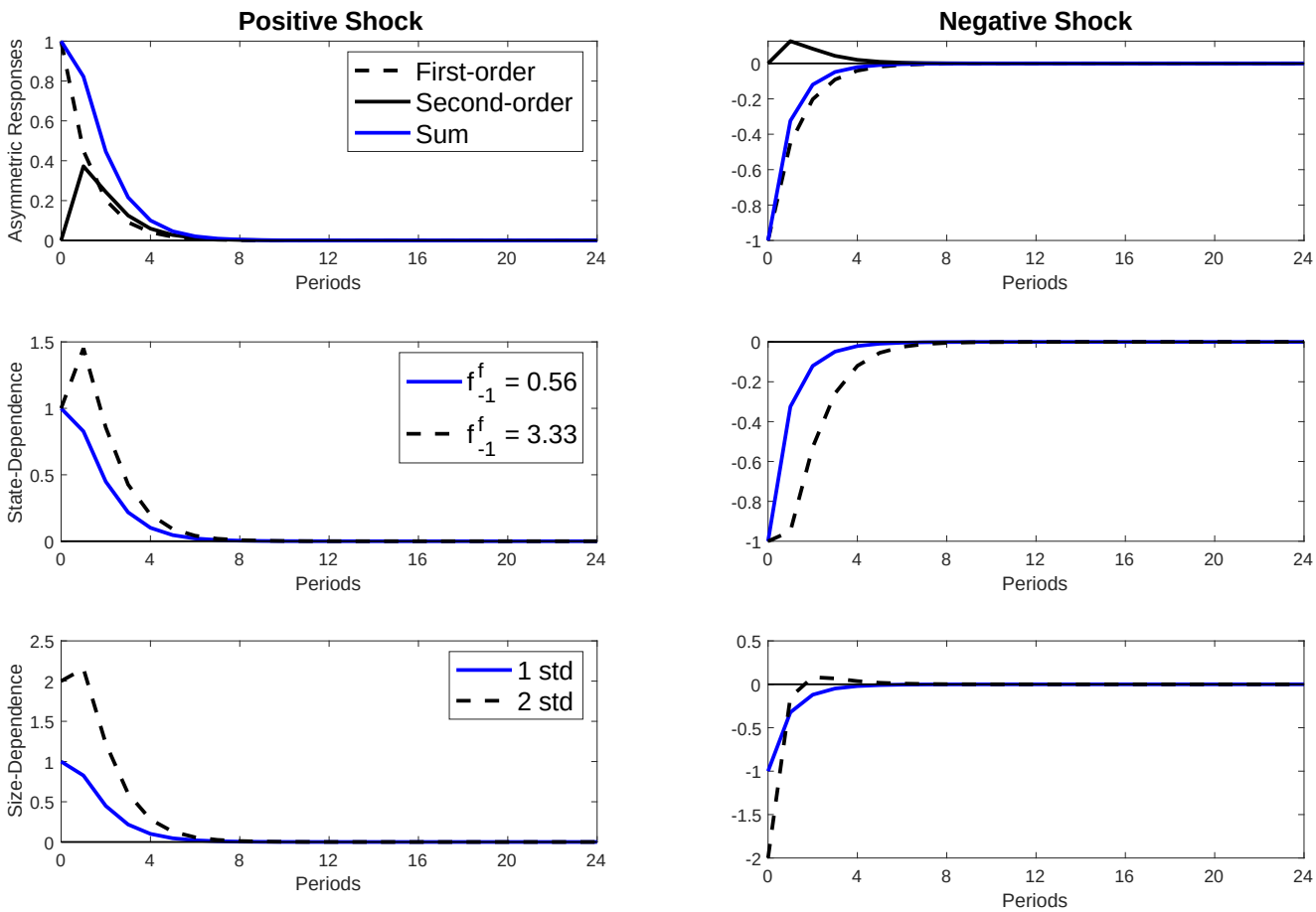
where $\Omega_{t-1} = \{f^f_{t-1}, f^f_{t-2}, ..., f^s_{t-1}, f^s_{t-2}, ...\}$, δ is the size of the innovation, f_{t+h} is the baseline value conditional on a path of shocks $\{\nu_t, \nu_{t+1}, ...\}$, and $f_{t+h}\left(\delta\right)$ is the counterfactual value conditional on the same path of shocks except with the addition of δ at time t $\{\nu_t + \delta, \nu_{t+1}, ...\}$.

This definition of IRFs has several properties. First, it integrates out the effects of future shock uncertainty, similarly to the generalized IRF. Second, as the time t shock is added onto a baseline path, the exercise is best thought of as showing the effects of perturbing the time t innovation ν_t by δ. This thought experiment is a slightly different when compared with the generalized IRF one, in which the time t shock is fixed at δ in the counterfactual case (Koop et al., 1996). We prefer the definition suggested by Goncalves et al. (2021) because it maintains randomness in the period of the shock, which has important implications when thinking about higher-order moments and the distribution. Finally, our IRFs are state dependent. We are explicit in our conditioning set to make this dependence clear to the reader.

We begin by discussing the asymmetric responses to shocks. The first row in Figure 1 shows the IRFs of the factor following a one standard deviation positive innovation (left panel) and a negative one (right panel), initializing the first-order factor f^f_{-1} at 0.56 and the second-order factor f^s_{-1} at its unconditional mean value. This calibration is illustrative, and we choose a nonzero lag of the first-order factor to showcase the state-dependence in the IRFs. The figure plots the dynamics of the factor (blue line) and its first- (dashed black line) and second-order (solid black line) components. These IRFs illustrate the asymmetry that the model can generate. Specifically, with this parameterization and initial condition, a positive shock persists for longer than a negative shock. This divergence can be seen by comparing the impulse response with its first-order component. The first-order component is linear and therefore symmetric. It produces impulse responses that are representative of those that come from a standard DFM. The blue line is formed by adding the first- and second-order components together. Given the initial conditions and the fact that $h_{xx} > 0$, the second-order term is always positive in this example. This finding means that the blue line is always above the first-order response, no matter whether the shock is positive or negative.

The second row of the figure illustrates the next important property that our model can pro-

Figure 1: Nonlinear Dynamic Factor Model Can Generate Asymmetric, State-Dependent, and Size-Dependent Responses to Shocks



NOTE: **Row 1**: Asymmetric responses to the same-sized shock in a calibrated model. This figure shows the effects of a positive shock on the left panel in blue. The dashed black line is the first-order response, while the solid black line is the second-order response. **Row 2**: State-dependent responses in a calibrated model. The responses to a positive shock are shown in the left panel. The blue line is the same as in Row 1, and the dashed black line is the response to the same shock but at a different initial condition. **Row 3**: Size-dependent responses in a calibrated model. The responses to a positive shock are shown in the left panel. The blue line is the same as in Row 1, and the dashed black line is the response to a shock twice the size as the one that generates the blue line. In all rows, the right panel shows the responses to a negative shock.

duce: state dependence. The solid blue line is the same response as in the top row. The dashed black line now shows the responses to the same-sized shock but starting at an initial condition of $f^f_{-1} = 3.33$ and f^s_{-1} at its unconditional mean. Although the initial impulse is the same, as can be seen by the identical response at time 0, the effects of the state dependence kick in with a one period lag. Starting from the different initial condition with an elevated first-order factor, both positive and negative shocks generate larger magnitude of responses in the factor. This difference can be

understood by examining Equation 8. The lag of the first-order factor determines the volatility of ϵ_t in the third equation governing $\left(f_t^f\right)^2$, with a f_{t-1}^f that is larger in magnitude leading to a higher variance of ϵ_t.

The responses shown in the second row of the figure lead us to another related fact, which we illustrate in more detail in the section on distributional responses to shocks but is worth mentioning here. The different amplification of shocks is indicative of time-varying volatility. This example illustrates that when the first-order component starts out at a larger value in magnitude, the overall nonlinear factor also becomes more volatile. We contrast this state dependence with factor models of exogenous stochastic volatility, such as Del Negro and Otrok (2008), where movements in the volatility of the factor are due to separate shocks.

Finally, the third row of Figure 1 shows the size dependence of the IRFs. What we mean by size dependence is that the shape of the IRF changes depending on the size of the shock. The blue line is the same one that we have carried over from the previous rows. The dashed black line is the response at the same initial conditions but to a shock that is two standard deviations in size instead of one standard deviation. On impact, the response is double that of the one standard deviation shock. In the next period, however, the shapes of the IRF change for positive and negative shocks. This fact is especially clear in this example after a negative shock, in which the IRF turns positive two periods after impact following a two standard deviation shock, while it stays negative following a one standard deviation shock. In essence, a two standard deviation change to the first-order component changes the second-order component by more than double that of the first-order component. This greater-than-proportionate response of the second-order component generates the different shape of the overall IRF.

Distributional Implications

Despite being driven by normal, homoskedastic shocks, the NLDF model produces rich non-normalities in the distribution of the factor, which then feeds into the distribution of the observables. We begin by discussing the unconditional distribution of the factor. Next, we move on to the time-varying volatility and tail risk that our model can produce.

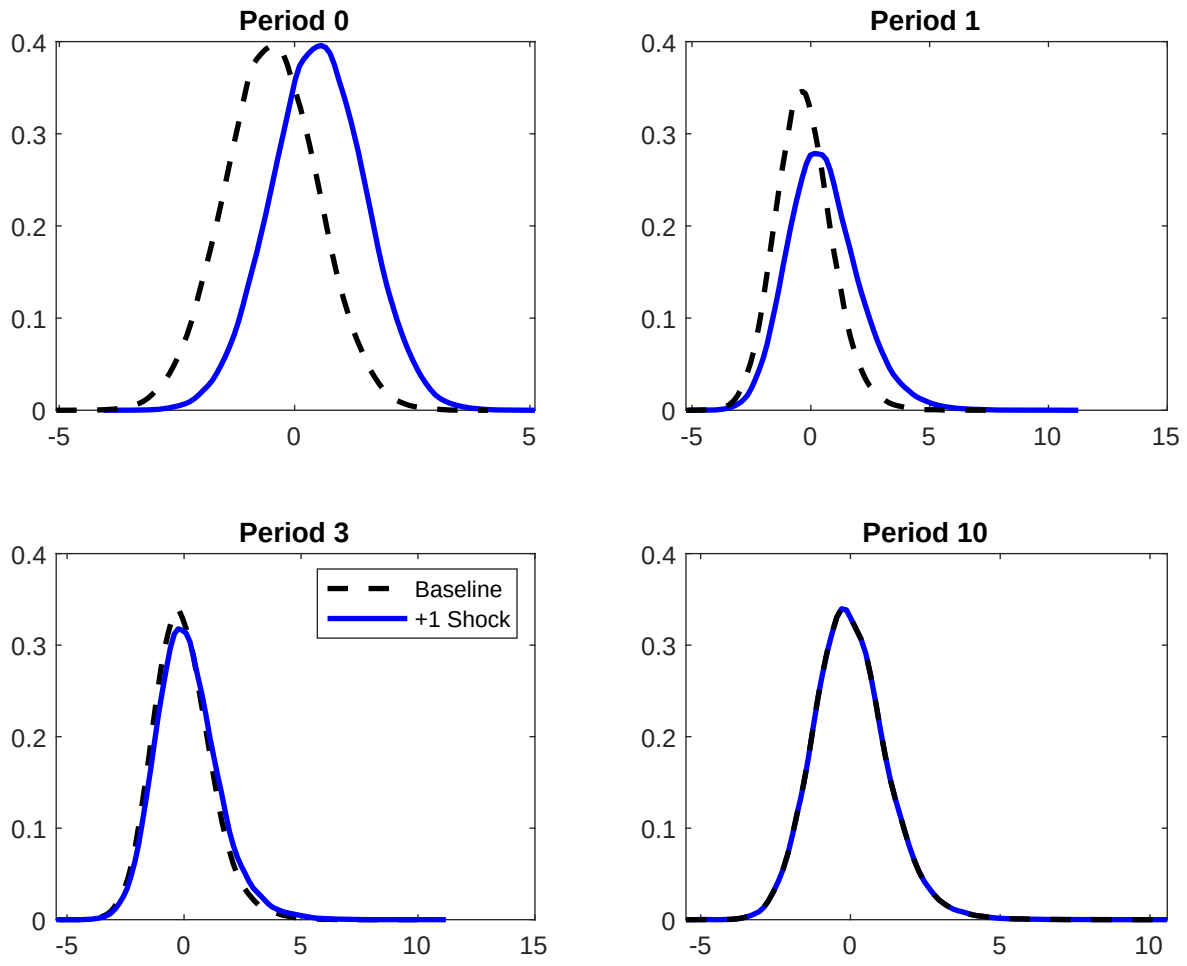
Figure 12 in Appendix Section C shows the distribution of the factor produced by a long simulation of the model. The distribution produced by the NLDF model is not normal, as evidenced by its positive skew. Its Kelley skewness, which measures the share of the distance from the 90th percentile to the 10th percentile that is above the median versus below the median, is 0.07.[6] By contrast, the distribution produced by the first-order component is normally distributed and therefore has a Kelley skewness of 0.

Focusing only on the unconditional distribution, however, masks important dynamics of the distributions following disturbances. In Figure 2, we illustrate this feature of our model by simu-

[6]The formula for Kelley skewness is $\frac{Q90+Q10-2*Q50}{Q90-Q10}$, where Q is the quantile of the distribution.

lating the predictive distributions at various horizons following the same one standard deviation shock that we began discussing in the top row of Figure 1.[7] The blue distribution is the baseline density – the one that characterizes the possible outcomes if we simulate the model from the initial conditions. The dashed black line is the density that realizes if we had a positive one standard deviation shock at period 0. Our model implies that such a shock leads to a positive shift in the distribution on impact. Crucially, in period 1, the distribution *widens out*. The differences in the distributions persist through period 3, and by period 10, the effects of the shock are largely gone.

Figure 2: Dynamics of the Distribution of the Demeaned Overall Factor in Response to a Positive Shock



NOTE: The blue line is the distribution after a positive shock, and the dashed black line is the distribution without a shock. Period 0 is the period of the shock.

[7] We generate these distributions by simulating $100,000$ paths from the initial condition. In the baseline case, we take draws from the data generating process. In the "$+1$ Shock" case, we add a one standard deviation shock to the impact period's innovations from the baseline. After the impact period, we use the exact same draws of the innovations in both scenarios.

Figure 3 shows the IRFs of the first- and higher-order moments to the shock. For the standard deviation and tail risk responses, we compute the IRFs as differences in the standard deviation and shortfall and longrise of the $+1$ Shock and Baseline distributions. The standard deviation of the distribution increases, peaking in the first period after the shock. These effects lead to a large increase in the 5% longrise, as a rise in the mean and increase in the standard deviation greatly increases the upside tail risk. The 5% shortfall also increases but by much less, as the mean increase is counteracted by the increased standard deviation. Therefore, through movements in the higher-order moments of the distribution, the model can generate distinct asymmetries in the movements of the upper and lower tails of the distribution, in line with the stylized facts documented by Adrian et al. (2019). These IRFs also reinforce that our model can generate time-varying volatility through the nonlinear dynamics.

Figures 13 and 14 in Appendix Section C show the corresponding distributional responses to a one standard deviation negative shock. A negative shock lowers the mean and decreases the standard deviation of the distribution. These factors together again generate a more persistent negative movement in the longrise but less persistent effects on the shortfall. Comparing the positive and negative responses, we see that the asymmetries we document for the first moment also carry over to higher moments. The declines in the standard deviation and tail risk are smaller in magnitude when compared with the increases in those features of the distribution following a positive shock.

3.3 Variants to Nonlinear Dynamic Factor Model

Before we move to the estimation section, we briefly discuss potential ways in which the model could be extended to study data that demand a richer factor structure.

Beyond a Second-Order Representation

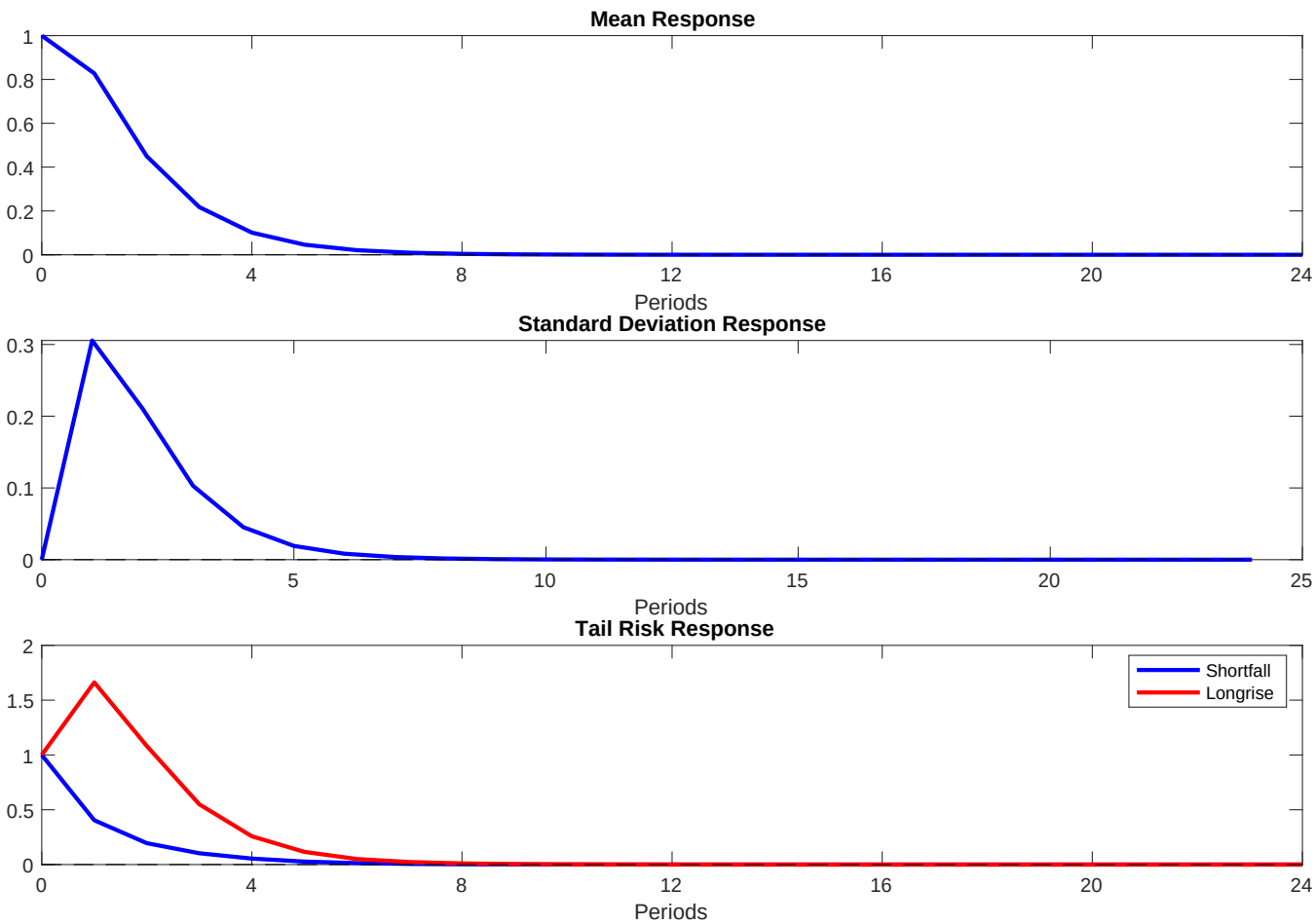
Our choice of the second-order pruned representation for the factor dynamics is based on its being parsimonious and on macroeconomists' familiarity with perturbation methods. But our exposition is general enough that one can use, for example, projection methods to approximate the functions $\mathcal{G}$ and $\mathcal{H}$. This alternative can capture richer nonlinearities that monomials cannot model. Let $\Psi_i(\cdot)$ denote the Chebyshev polynomial of degree i. Then the nonlinear state equation can be approximated by

$$f_t = \sum_{i=0}^{n} \theta_i \Psi_i(\widetilde{f}_{t-1}) + \sigma \nu_t.$$

Here, θ_i are parameters to be estimated and $\widetilde{f}_{t-1}$ is a transformation of the original $t-1$ factor such that it is bounded between -1 and 1.[8] However, this option comes at the cost of potentially more complex likelihood when estimating the model.

[8]This transformation is necessary because Chebyshev polynomials are defined in the interval $[-1, 1]$.

Figure 3: Impulse Response Functions of the Mean, Standard Deviation, and Tail Risk of the Demeaned Overall Factor in Response to a Positive Shock



NOTE: In the third panel, the blue line shows the response of the 5% shortfall, and the red line shows the response of the 5% longrise.

Multidimensional State

One can, in theory, easily expand our model to accommodate more factors. Below, we still use two observables, but we add an additional factor.

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \underset{2\times 2}{G} \begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} + \underset{[2\times 2]}{\eta} \underset{[2\times 1]}{\epsilon_t},$$

$$\begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} = \mathcal{H}\left(\begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} \right) + \underset{2\times 2}{\Sigma} \nu_t.$$

Here, the function $\mathcal{H}(\cdot)$ is the nonlinear map between the factors yesterday and the factors today. If one extends the pruned representation from above to the two-factor case, the results from Proposition 2.1 carry over. Specifically, two factors must be named – i.e., their loadings in two of the observable equations must be set to 1.

4 Macro and Finance Applications Revisited

Now we use the nonlinear factor model to study the role of exogenous and internal forces in the dynamics of macro and financial series. We have two empirical applications. First, we estimate a shadow rate model motivated by the work of Wu and Xia (2016). Second, we estimate a nonlinear credit cycle factor.

4.1 Shadow Rate

Short-term yields have recently hit their lower bound constraints, prompting modifications to existing yield curve models to account for this behavior. An important advancement came with models that considered an ELB constraint on short-term yields and thereby allowed the latent yield curve factor to turn negative. Notably, Wu and Xia (2016) showed that these shadow rate models delivered a meaningful measure of monetary policy conditions even when the short rate was stuck near zero.

The usual assumption made in shadow rate models is that the yield factor dynamics are linear and do not change upon entering the ELB. Implicitly, this assumption presumes that the economy does not undergo structural changes due to constrained monetary policy, which is at odds with some of the theoretical literature (Fernandez-Villaverde et al., 2015, Aruoba et al., 2017b) and supported by others (Wu and Zhang, 2019, Bernanke, 2020). Some empirical studies have investigated whether structural changes in the economy occurred as a result of the ELB constraining policy (Swanson and Williams, 2014, Debortoli et al., 2019, Wu and Zhang, 2019, Aruoba et al., 2021). Among these works, one of particular relevance for our purposes is Swanson and Williams (2014), who investigate if the behavior of yields at longer-term maturities in response to news changed during the ELB.

Our work contributes to this debate by investigating whether there is evidence of nonlinearities in the factor dynamics of a shadow rate model. We simultaneously model ELB restrictions on the yields in the measurement equation and nonlinear factor dynamics. If there were structural changes in the yield curve movements due to the ELB, our model would be able to capture them through the second-order factor. Another contribution of our work is that we model the yields in first differences instead of levels and show how to do so while accounting for the ELB (Onatski and Wang, 2021, Crump and Gospodinov, 2022).

Empirical Set-up. We estimate our model on the first differences of one-month-forward rate data at 3-month, 6-month, 1-year, 2-year, 3-year, 5-year, and 10-year maturities from February 1990 to September 2019, following Wu and Xia (2016) to construct the data.[9] We impose a 0.3% ELB on the series.

We consider the following model of differences in forward rates:

$$\Delta forward_t^h = m_h + \begin{cases} G_h(c + f_t^f + f_t^s) + \eta^h \varepsilon_t^h & \text{if } \widehat{S}_t^h >= 0.3 \\ -m_h + \eta^h \varepsilon_t^h & \text{otherwise} \end{cases} \tag{22}$$

where $c = -\frac{1}{2}\frac{h_{xx}\sigma^2}{(1-h_x)(1-h_x^2)}$, $\widehat{S}_t^h = \sum_{\tau=2}^{t}\left(m_h + G_h\left(c + f_\tau^f + f_\tau^s\right)\right) + forward_1^h$, $\Delta forward_t^h = forward_t^h - forward_{t-1}^h$, and index h stands for the maturity. We model the latent factor according to our second-order dynamics:

$$\begin{aligned} f_t^f &= h_x f_{t-1}^f + \sigma \nu_t \\ f_t^s &= h_x f_{t-1}^s + \frac{1}{2} h_{xx}\left(f_{t-1}^f\right)^2 \end{aligned} \tag{23}$$

We normalize the factor loading G_1 on the three-month rate to be 1. We also allow for yield-specific constants m_h that capture any differences in the average forward rate changes across maturities.

The model has two key nonlinearities. First, there is an ELB constraint in the measurement equation, which removes any influence of the factor if it predicts a rate lower than 0.3% in levels. Instead, the measurement equation is set to $\Delta forward_t^h = \eta^h \epsilon_t^h$. At the ELB, the observed change in the forward rate should be 0%, as the level is stuck at 0.3%. The measurement error picks up a residual difference. Second, the latent factor dynamics are allowed to be nonlinear.

We estimate the nonlinear model using Metropolis Hastings combined with the bootstrap particle filter.[10] We use $500,000$ particles in the particle filter. We take $510,000$ draws from the posterior distribution with a burn in of $210,000$. We construct the posterior distributions for our results by taking every 100th parameter draw from the remaining $300,000$ draws.

As a comparison, we also estimate a linear version of the model in which we do not model the ELB in the measurement equation and impose linear factor dynamics.[11] Insofar as the nonlinearities we consider are empirically important, the nonlinear model should fit the data better than the linear model does. If the h_{xx} term in the nonlinear model is estimated to be insignificant, then we interpret that as evidence against state-dependent dynamics at the ELB.

Historical Estimates of the Forward Rate Factor. We first discuss the historical filtered estimates of the forward rate factor implied by our nonlinear model and compare it with a linear

[9] Further details about the data construction can be found in Appendix Section D.

[10] The estimation details and prior distributions can be found in Appendix Sections A and D.

[11] We estimate this model with the Metropolis Hastings algorithm and the Kalman Filter. Further details can be found in Appendix Section D.

model that does not account for the ELB restriction and imposes only linear factor dynamics. The top panel in Figure 4 shows the filtered estimates of the factor produced by our nonlinear model. The factor captures the contours of historical yield curve movements. For example, we catch the rapid drop in short- and long-term rates in the early 1990s, which was then followed by a tightening cycle in monetary policy during the middle of the decade. From 1995 through 1996, the 10-year forward rate dropped more than 2 percentage points, which our factor captures. As we move on to the 2000s, our estimated factor reflects the easing cycle in the early 2000s followed by tightening beginning in 2004. The factor then rapidly drops entering into the GFC. In the middle of the GFC, short-term interest rates hit their ELB, which is highlighted by the gray area in the figure. Because our model can account for an ELB in the measurement equation, however, it still estimates variability in the factor, chiefly the continued decline in long-term rates. The lack of variation of shorter-term maturities leads to a widening of the uncertainty in the estimated factor as more forward rates hit their ELB. Finally, as we exit the ELB, the factor captures the rise in short-term forward rates.

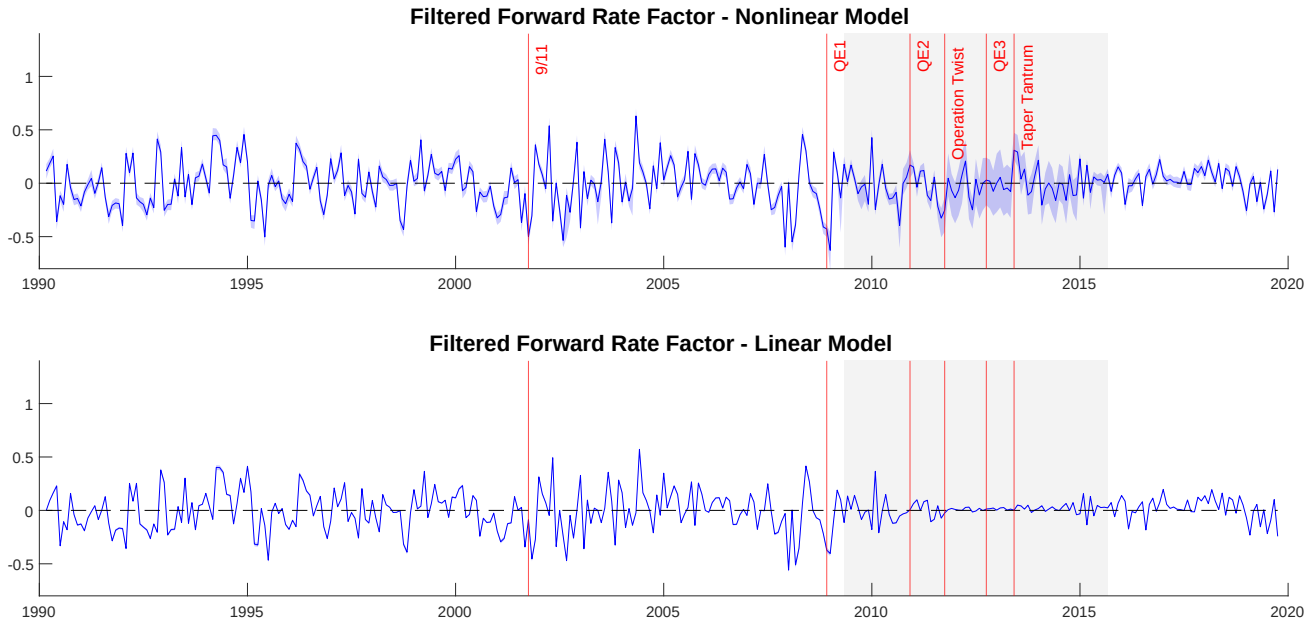
It is worth emphasizing that our factor matches some important events in recent monetary policy (the red lines in Figure 4). For example, there is a sharp decline in the factor in the aftermath of 9/11. A similar sudden drop is observed around the announcement of the Fed's first round of quantitative easing (QE1). In contrast, the Taper Tantrum of 2013 coincides with a rise in the factor. Interestingly, our estimated factor shows that QE3 did not result in a change of the stance of monetary policy.

The bottom panel shows the corresponding filtered estimates from a linear model. Outside of the highlighted ELB period in gray, the two models estimate similar factors, foreshadowing the limited role played by the second-order factor dynamics. During the ELB period, however, the factor estimates diverge. The linear model is constrained by the absence of variation in short-term forward rates, so they were stuck at 0, while long-term rates continued to vary. Indeed, these fluctuations in the longer maturity rates inform the dynamics of our factor during the zero lower bound episodes. On balance, the linear model estimates little variation during the ELB period, thereby sacrificing fit to the long-term yields.

Figure 5 shows the model implications for the filtered level of the shadow $3-$month and $10-$year rates in red. We calculate these values by $\widehat{forward}_t^{h,sh} = \sum_{\tau=2}^{t} m_h + G_h \left(c + f_t^f + f_t^s\right) + forward_1^h$. We ignore the ELB restrictions placed on the forward rate, and so these estimates are best interpreted as shadow rates (hence the sh in the superscript). In the top row, we also show the Wu and Xia (2016) shadow federal funds rate in blue for comparison.

As our nonlinear model allows the shadow rate to go negative, we capture remarkably similar dynamics to Wu and Xia (2016). Namely, our shadow three-month rate continues to trend down into 2014 and 2015 before lifting off in early 2016. Wu and Xia (2016) interpret the decline of the shadow rates during this period as evidence of the effectiveness of unconventional monetary policy, the effects of which can be seen in the decline of longer-term unconstrained rates. The influence

Figure 4: Filtered Estimates of the Forward Rate Factor from the Nonlinear and Linear Models



NOTE: Filtered estimates of the forward rate factor estimated by the nonlinear model (top panel) and linear model (bottom panel) and 80% credible sets (blue shaded area). The gray shaded areas denote the periods in which the 3-month forward rate is at the effective lower bound. QE1, QE2, and QE3 are the first, second, and third rounds of quantitative easing by the Fed.

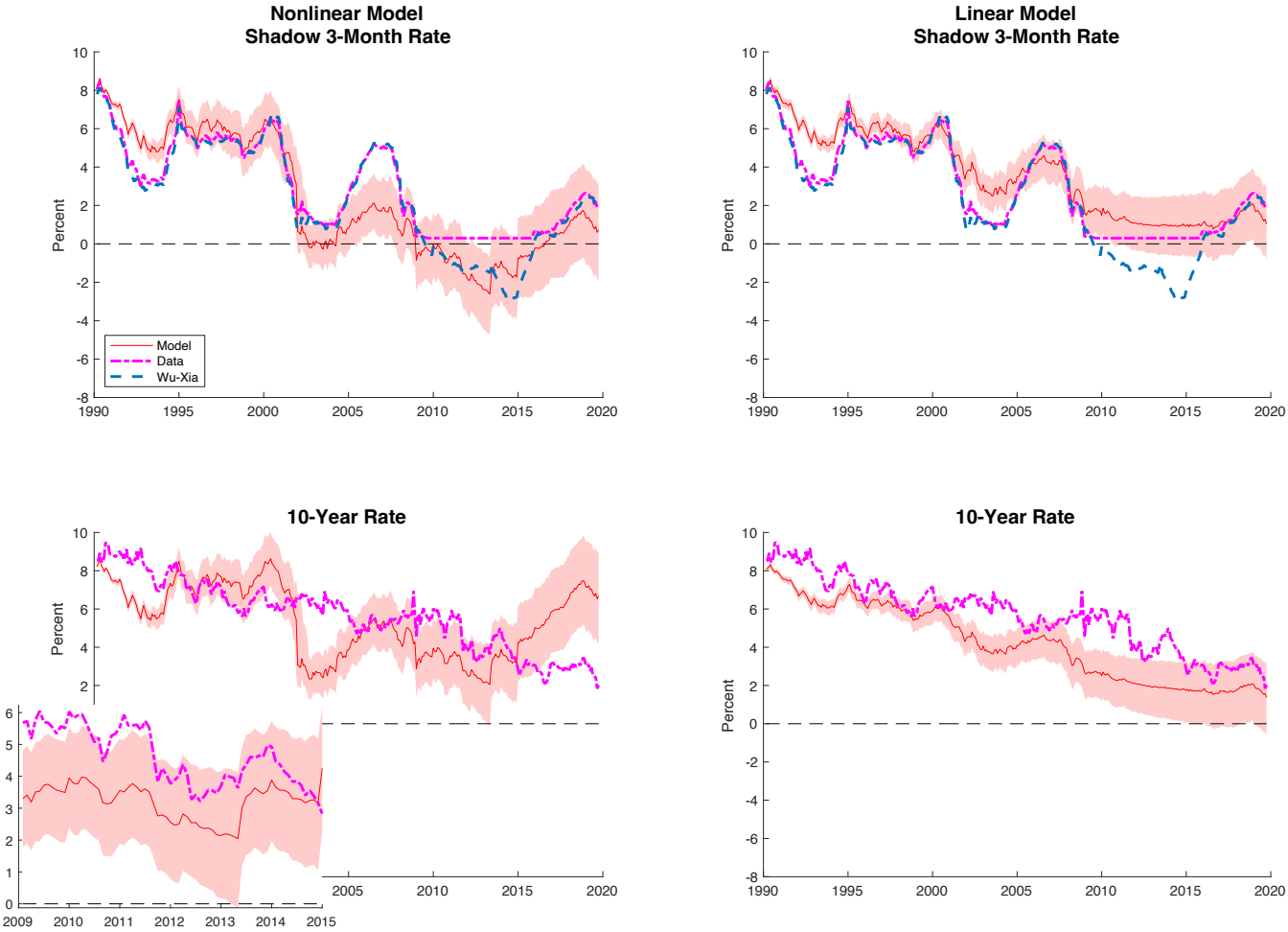
of the longer-maturity rates on the shadow rate between 2009 and 2015 can be seen in the inset at the bottom-left panel in Figure 5. Both the 10-year and the shadow rates feature a downward trend and have broadly similar dynamics. In contrast, the linear model's estimates do not capture any of these movements in the ELB period, and we see a largely flat prediction of the three-month rate from the factor.

Importance of the Nonlinear Components. How important are the nonlinear additions to the model? From the filtered estimates, we see clear evidence that the ELB restriction tangibly changes the factor estimates. Moreover, a marginal likelihood comparison between the two models shows that the nonlinear model is heavily favored by the data at 640 versus 575 log points.[12] Taken together, these results suggest that allowing for nonlinearities is important to understand yield curve dynamics.

The central question in our investigation is whether the ELB produced structural changes in yield curve dynamics and therefore changes in the behavior of longer-term forward rates uncon-

[12]We compute the marginal likelihood using the modified harmonic mean (Geweke, 1999). We use a truncation parameter of 0.95. The results are similar for truncation parameters of 0.5 and 0.75.

Figure 5: Filtered Estimates of the Shadow 3-Month Rate and 10-Year Rate from the Nonlinear and Linear Models



NOTE: Filtered estimates of the shadow $3-$month forward rate (top row) and the $10-$year rate (bottom row), by the nonlinear model (first column) and linear model (second column). The red lines denote the filtered estimates with 80% credible sets (red shaded area). The dot dashed magenta line is the observed data. The dashed blue line is the shadow federal funds rate estimated by Wu and Xia (2016).

strained by the ELB. This question can be answered by examining whether the statistical gains from the nonlinear model are primarily due to the ELB constraint on the measurement equation, the second-order factor dynamics, or both. Table 2 in Section D of the Appendix shows the 80% credible sets of parameter estimates. There, we can see that the credible sets for h_{xx} – the key parameter that governs the second-order factor – ranges from -0.01 to 0.47. These estimates contain 0 and, at best, can be characterized as marginally significant, suggesting that higher-order factor dynamics play a limited role. We can also estimate a version of the model in which we maintain the ELB restriction in Equation 22, but we impose linear factor dynamics. This version of the model

produces filtered factors similar to the fully nonlinear model and fits the data slightly better in a marginal likelihood sense (642 versus 640 log points).

Our empirical evidence then is in favor of the idea that the ELB mainly was a restriction on the behavior of short-term yields. There is little evidence of nonlinearities in the factor dynamics, at least using our model. Therefore, dynamics of the factor continued to propagate linearly as in unconstrained times.

4.2 Nonlinear Credit Cycle

Since the GFC, economists have once again taken a close look at the importance of credit growth for macroeconomic fluctuations (Schularick and Taylor, 2012). Excessive credit buildups often precede financial crises, and leverage can further amplify shocks. Moreover, it is not enough to focus on one credit sector, but instead a broad monitoring framework is needed (Adrian et al., 2015). For instance, Mian et al. (2017) emphasize the importance of household debt to GDP as a predictor of lower GDP growth and higher unemployment worldwide. Corporate leverage may lead to distorted investment decisions due to debt overhang effects (Gomes et al., 2016). Financial-sector leverage can amplify shocks via the financial accelerator and binding borrowing constraints (Bernanke et al., 1999, Gertler and Karadi, 2011). Finally, as discussed in Jorda et al. (2016), high levels of public debt tend to prolong the pain of private-sector deleveraging. Taking center stage in these studies is the importance of credit growth.

Our second application investigates the importance of a common component in real credit growth in the United States across the nonfinancial business, household, financial, and public sectors from 1952:Q1 through 2021:Q4. Credit growth across different sectors may move together because of common factors such as changes in risk appetite, financial technology, or structural reforms. Moreover, economic theory suggests the potential importance of nonlinearities in determining the dynamics of credit growth. Minsky (1977) describes an economy that may experience a rapid contraction in credit after a long boom with speculative lending as expectations rapidly change. Bordalo et al. (2021) formalize these dynamics in a model with diagnostic expectations. Several papers highlight the role of occasionally binding borrowing constraints in modeling U.S. business and credit cycles (Brunnermeier and Sannikov, 2014, Guerrieri and Iacoviello, 2017). We view our model as one avenue to check how important the nonlinear dynamics are in the data without needing to resort to a fully specified structural model.

We estimate a one-factor version of our nonlinear factor model (Equation 3). We use a particle Gibbs sampling algorithm with 100 particles and take 1.5 million draws from the posterior distribution, burning in the first $600,000$. To form our posterior distribution, we take every 300th draw for a total of $3,000$ draws. For posterior distributions of IRFs and distributional moments, which require heavier computation, we use $1,000$ draws of the parameters. Further details about the estimation, including the prior specification, can be found in Sections A and E of the Appendix.

Credit Growth Dynamics. Figure 15 in Appendix Section E shows the data that we use to estimate the model, which are normalized U.S. real credit growth in the nonfinancial business, household, financial, and government sectors. The data come from the Statistical Release Z.1 "Financial Accounts of the United States" data provided by the Federal Reserve Board.[13]

Private credit growth generally increases during expansions and declines during recessions, although the troughs of nonfinancial business and financial credit growth lag the troughs of recessions. The three private credit growth series are fairly positively correlated, ranging from 0.4 to 0.5. In the nonfinancial business and household sectors, credit growth exhibits an important asymmetry, with expansions marked by steady, strong growth and recessions associated with sharp, violent declines. These dynamics have implications for higher-order moments, with the Kelly skewness of nonfinancial business credit growth at -0.23 and household credit growth at -0.15. Financial credit growth experienced rapid declines in the GFC but overall has a skewness close to 0.

Conversely, government credit growth is mildly negatively correlated to the three other series because it has increased in recent recessions. The series has a distinct positive skew due to several large spikes in public debt.

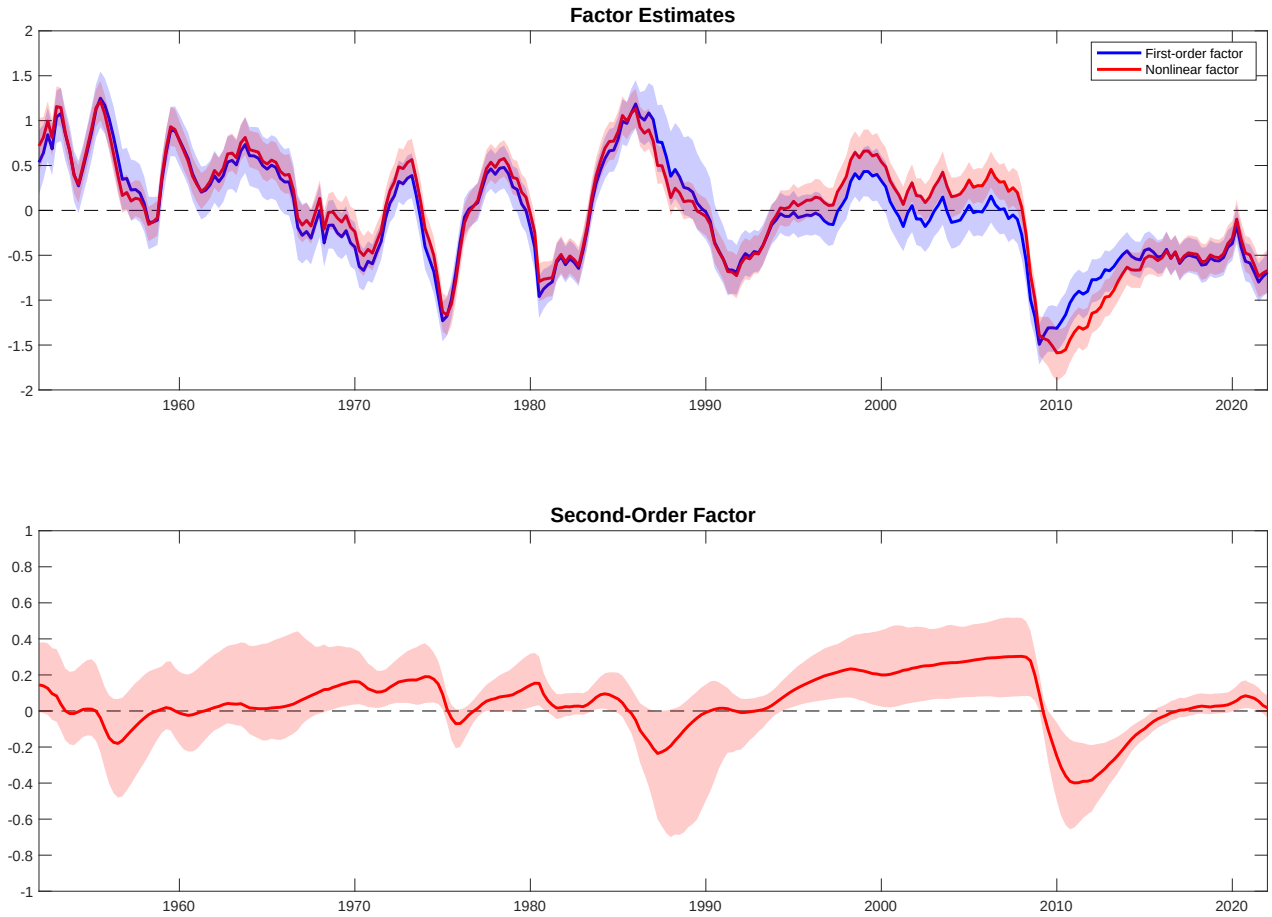
Historical Credit Cycle Estimates. Our estimates provide evidence of a nonlinear factor that we call the credit cycle. For identification purposes, we fix the factor loading for nonfinancial business credit growth at 1. The factor positively loads onto the household and financial sectors, with 100% of draws above 0 in both cases. Indeed, the posterior median of the factor loading on household credit is 1.3, with nearly all draws above 1, while the posterior median of the loading on financial credit is around 1. The factor, therefore, is heavily informed by the common cyclical co-movement of the three private credit growth series. The factor, however, also plays a role in understanding the public credit growth dynamics. It has a factor loading of -0.2 on the government credit growth series, with nearly all draws less than 0. Therefore, the factor broadly captures the correlation dynamics we documented in the data.

Figure 6 shows the smoothed factor estimates. In the top panel, the red line is the posterior median of the nonlinear factor estimates along with the 68% credible bands.[14] The credit cycle factor was strong throughout the 1960s before the recession in 1969. It then rebounded before collapsing again during the mid-1970s recession, with similar dynamics repeating again in the late 1970s to early 1980s. The frequency of the credit cycle lengthened afterward, with a robust expansion in the 1980s before declining again in the late 1980s and early 1990s with the savings and loan crises. Following that episode was a prolonged expansion through the 1990s and 2000s before the collapse in the GFC. The recovery from the GFC was especially slow, with the credit factor still below its mean even more than 10 years after the recession. This finding points to a secular stagnation in financial markets, that is, the GFC resulted in a significant and permanent

[13] The details of the data construction can also be found in Section E of the Appendix.

[14] We show 68% bands instead of 80% in the previous application because we are using quarterly as opposed to monthly data (e.g. Stock and Watson (2016)).

Figure 6: Estimated Credit Cycle and the Contribution of the Second-Order Factor



NOTE: Smoothed factor estimates produced by the nonlinear dynamic factor model. The top panel shows the demeaned factor estimates, with the red line being the estimate of the nonlinear factor and the blue line the estimate of the first-order factor. The shaded areas denote 68% credible sets. The bottom panel shows the demeaned second-order factor with 68% credible sets.

change in financial markets. Indeed, we can see that the stagnation captured by the factor arises from the dynamics of credit in the household and financial sectors.

The blue line and shaded areas are the corresponding movements of the first-order factor only for comparison purposes. These are the counterfactual estimates of the factor if we had set h_{xx} to 0 across all of the draws, holding all else equal. The bottom panel shows the estimates of the second-order factor adjusted to have 0 mean. The nonlinear component of the model was significantly positive starting in the 1970s, providing a boost to credit growth. It then declined to negative territory in the late 1980s during the savings and loan crisis. The factor again turned positive for a 15-year stretch beginning in the early 1990s until the GFC, when the second-order factor swung heavily negative. This negative swing contributed to the sluggish recovery of credit growth post-crisis.

Table 1: Unconditional Moments Implied by the Nonlinear Dynamic Factor Model and the Linear Factor Only

	Skewness	5% Shortfall	95% Longrise	Corr$\left(f^s, \left(f^f\right)^2\right)$	Variance Decomp
Nonlinear	-0.18	-1.91	1.22	-0.53	9.69
	$[-0.27, -0.07]$	$[-2.59, -1.38]$	$[1.08, 1.37]$	$[-0.54, -0.51]$	$[1.41, 25.78]$
Linear Only	-0.00	-1.34	1.33	–	–
	$[-0.00, 0.00]$	$[-1.49, -1.20]$	$[-1.48, -1.19]$	–	–

NOTE: The table shows the Kelly skewness, 5th/95th shortfall and longrise, and the variance decomposition showing the percentage of unconditional variation implied by the second order factor. "Nonlinear" refers to the full model while "Linear Only" refers to a counterfactual in which $h_{xx} = 0$ for all of the draws, keeping everything else the same. The headline number is the posterior median while the numbers in brackets are the $16/84$ credible sets.

Importance of the Nonlinear Factor. How important is the nonlinear factor when modeling the credit cycle? We answer this question in three ways. First, we look at the unconditional distribution of the nonlinear model compared with a counterfactual one with only the first-order factor active. Second, we investigate the state-dependent effects of shocks conditioning on three periods: the credit boom in the mid-2000s, the bust in the late 2000s and early 2010s, and a mixed case in the late 1980s. We find the last period listed particularly interesting, as it had a positive first-order and overall credit factor but a negative second-order factor. This finding is in contrast to the first two periods, in which the first- and second-order factors had the same signs. Finally, we look at the standard deviation and tail risk effects of shocks. It is important to reemphasize that in a linear DFM, shocks do not have state-dependent nor higher-order moment effects.

Unconditional Distribution

A key implication of the nonlinear model is that the unconditional distribution of the factor is not normally distributed, even though the exogenous innovations to the system are. This divergence from normality does not occur if we ignore the second-order component. In examining the credit growth data, we saw some evidence of asymmetries. These features of the data inform the estimation of the nonlinear model. Table 1 shows that the nonlinear model generates a negative Kelly skewness, with mass below the median of the distribution covering nearly 60% of the total distance from the 10th to the 90th percentiles. The credible sets of the skewness estimates are wide, reflecting the difficulty in pinning down the magnitude of the higher-order moments. However, the evidence indicates that the skewness is negative at the 68th percentile credible sets, as seen in the table. This behavior continues to be the case at the 80th percentile sets as well. The second row of the table shows the corresponding estimates for the linear-only model. With a linear process and Gaussian shocks, the model cannot generate any skewness.

The next two columns in Table 1 show the estimates of the lower and upper tails of the distri-

bution. As a reference, the mean of the factor by assumption is 0. The nonlinear model generates a distribution that has higher probability on large declines in the credit cycle as opposed to large increases. This asymmetric tail behavior is consistent with the negative skewness previously discussed. In comparison, the linear model generates symmetric tail behavior.

Underlying the skewness and tail risk behavior of the model is a strong correlation between the level and volatility components of the nonlinear factor. The second-to-last column of Table 1 shows the model-implied correlation between the second-order factor, which enters into the level of the nonlinear factor, and the square of the first-order factor, which determines the conditional volatility of the innovations to the second-order factor. This correlation is -0.53, which suggests that the conditional volatility of the credit cycle increases as the credit cycle declines. This result is again consistent with the idea that credit expansions are smoother than credit contractions. Moreover, it can also generate the negative skewness and long lower tails coupled with short upper tails we see.

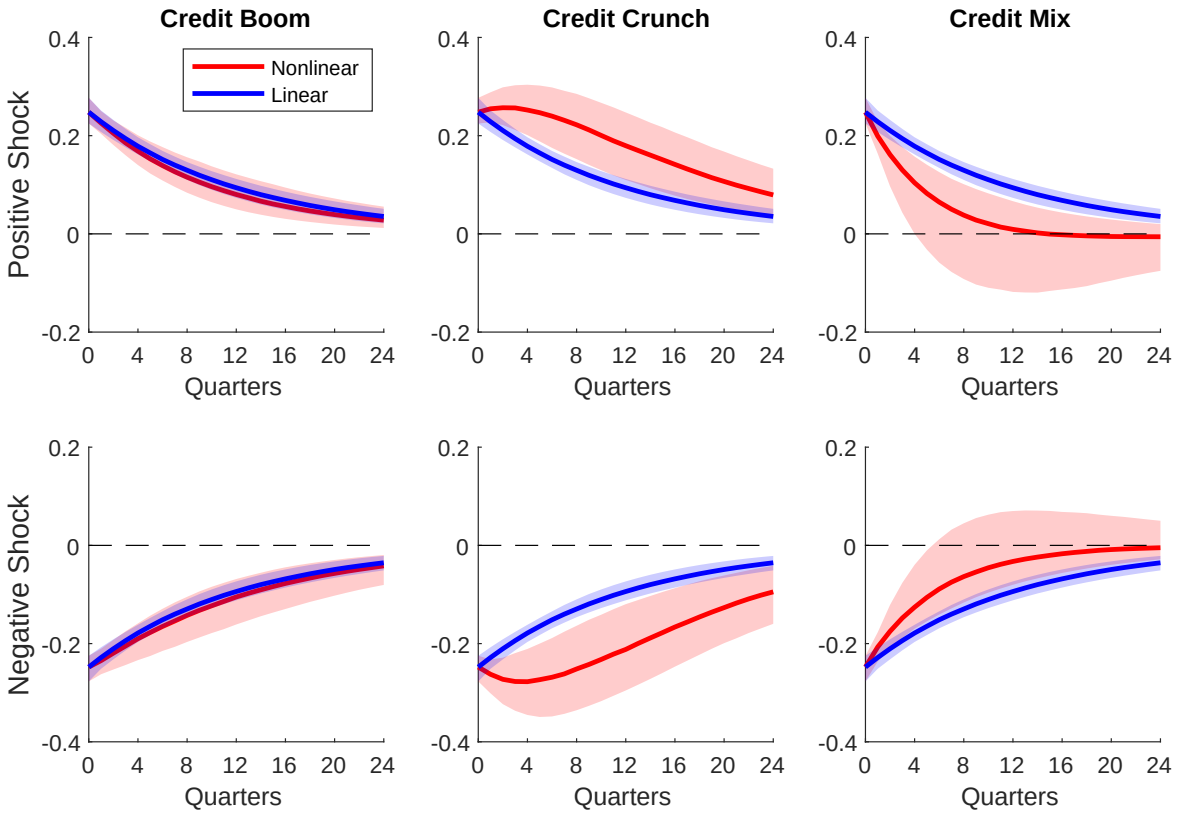
In addition to examining deviations from normality, we can compute the unconditional variance decomposition of the overall factor into its linear and nonlinear components. If the second-order factor's share of overall fluctuations is high, then it is further evidence that nonlinearities play an important role in the credit cycle. The last column shows this variance decomposition for the second-order factor. Its median estimate is around 10%, indicating a secondary, although still quantitatively relevant, role. Similar to the results before, its credible set is wide.

State-Dependent and Asymmetric Effects of Shocks

Two key aspects of the nonlinear credit cycle are state-dependent and asymmetric responses to shocks. Figure 7 shows the responses to one standard deviation positive and negative shocks to the credit cycle factor. The red line and shaded areas are the responses from the nonlinear model, while the blue line and shaded areas are the responses from the linear model. The first row shows the effects of a positive shock, whereas the second row shows the effects of a negative shock. The columns condition on the smoothed state estimates of three different periods: a boom period in the mid-2000s, the bust after the GFC in 2010, and a mixed case leading into the early 1990s recession.

During the credit boom period, where both the first- and second-order factors were positive, the expected path of the credit factor actually behaves similarly to the linear-only model. The persistence of the first-order factor, governed by the h_x parameter, has a posterior mean of 0.92, with the nonlinear factor showing similar intertemporal dynamics. As we move to the credit crunch period in 2010, the first- and second-order factors both were negative, which generates a response to the shock that is more persistent and with a slight hump shape in the initial quarters. There is a change in the conditional volatility of the shock when compared with the credit boom period as well, with the magnitudes of the responses to the same-sized shock larger in the quarters after its realization. These findings are consistent with the unconditional distribution results, which found a negative relationship between the level of the credit factor and its conditional volatility. The final column of the figure shows a mixed period before the early 1990s recession. The smoothed

Figure 7: State-Dependent Impulse Response Functions in Three Periods



NOTE: The red lines denote the responses of the overall factor following a positive shock (top row) and negative shock (bottom row), while the blue lines denote the responses of the linear component of the model. The first column conditions on a credit boom period in the mid-2000s, the second column conditions on a credit bust period in 2010, and the third column conditions on a mixed case before the early 1990s recession. The shaded areas denote 68% credible sets.

first-order factor was positive but declining, while the second-order factor became negative. This combination of states leads to a response to the credit factor shock that dies out more quickly compared with both the credit boom and crunch states. This response is true both for a positive and negative shock.

In summary, the nonlinear model exhibits evidence of state dependence in the responses to a shock. The results are in line with theoretical predictions as well. When the credit cycle is strong, a credit factor shock behaves approximately linearly. These times correspond to periods of slack borrowing constraints and easy credit (Guerrieri and Iacoviello, 2017). Times immediately after credit crunches generate amplification and persistence as borrowing constraints tighten. Our empirical results suggest that dynamics in the data are consistent with these theories.

We also comment briefly on the asymmetry in the responses to positive versus negative shocks. Across all of the periods, evidence indicates that a negative shock generates a larger and more

persistent response when compared with a positive shock. Negative shocks lead to a response approximately 10% larger in magnitude when compared with positive ones.[15]

Finally, we search for evidence of size dependencies in the response to a shock. We identify two historical episodes in which the model estimates large shocks: 1980:Q2 and 2008:Q2. Then, we ask whether a two standard deviation shock generates a different response when compared with two times a one standard deviation shock during these times. We find little evidence of this mechanism at play for either a positive or negative shock.

Higher-Order Moment Effects of Shocks

Figure 8 shows the higher-order moment effects of shocks. For these results, we condition on the credit crunch state, although many of the qualitative features we discuss apply to the other times as well.

A positive shock leads to an increase in the mean and a decline in the volatility of the credit factor predictive distribution, as seen on the first two columns of the figure. As we move to the last column, we see the effects that these shocks have on the tail risk of the predictive distribution. The shortfall increases more than the longrise does because the increase in mean and decrease in volatility both lead to the lower tail of the distribution shifting leftward. By contrast, these effects partially cancel each other out on the upper end of the distribution, generating the more muted response.

The bottom row shows the response to a negative shock. The responses flip in sign, with the shock generating an increase in the volatility. Both the shortfall and longrise decline, with the decline in the shortfall still greater than the decline in the longrise. Taken together, these results suggest that a credit cycle shock produces larger moves in downside risk relative to upside risk. An adverse credit cycle shock lowers the factor on average, and it also increases the risk of particularly large declines due to an increase in volatility. In contrast, a positive shock increases the factor, on average, and further decreases the risk of large declines due to a decline in volatility.

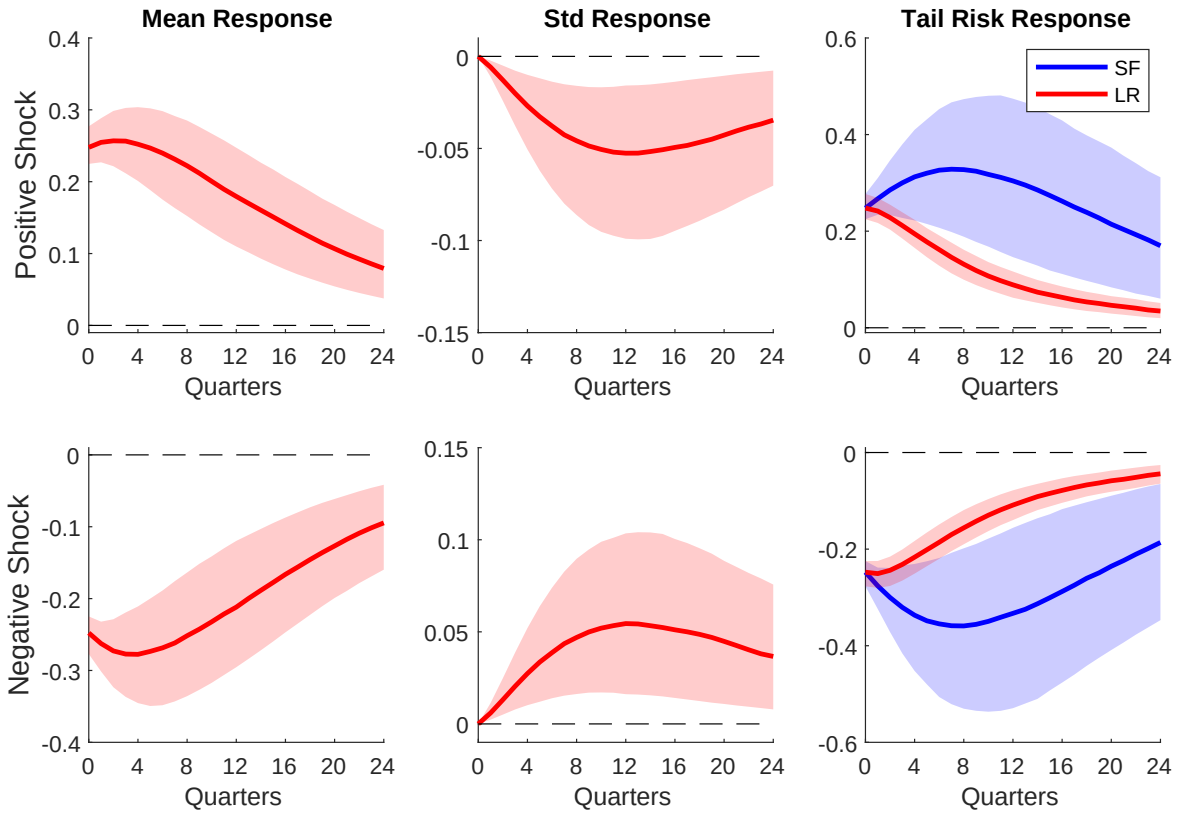
5 Conclusion

We propose a parsimonious NLDF model that is built around a *pruned* second-order factor equation. In this model, the propagation of shocks is asymmetric, state dependent, and size dependent, and stationarity is guaranteed by construction. The application of the particle filter to evaluate the likelihood and extract the factor allows us to augment the nonlinear factor motion with nonlinearities in the measurement equation, which makes the model applicable to macroeconomic environments in which variables can be constrained.

We investigate the properties of the model and illustrate the nonlinear measurement equation, estimating the *shadow rate* model à la Wu and Xia (2016) with a measurement equation that specifies

[15] In Section E of the Appendix, we present results on the difference in the magnitudes between positive and negative shocks.

Figure 8: Impulse Response Functions of the Mean, Standard Deviation, and Tail Risk During the Credit Crunch



NOTE: Impulse response functions of the mean, standard deviation, and 5% shortfall and longrise of the demeaned overall factor produced by the nonlinear dynamic factor model during the credit crunch period in 2010. The responses to a positive shock are shown in the top row, and the responses to a negative shock are shown in the bottom row. In the third column, the blue lines denote the shortfall (SF) response, while the red lines denote the longrise (LR) response. The shaded areas are 68% credible sets.

an ELB on U.S. data. We show how the extracted *shadow rate* factor and thus conclusions regarding the monetary conditions differ between models: one with a nonlinear measurement equation and second-order factor dynamics and another that is a standard linear factor model. The former predicts an easing of monetary conditions during the ELB period, while the latter does not provide evidence of such.

Our credit cycle application emphasizes the importance of a second-order component when measuring the credit cycle. This nonlinearity leads to state-dependent IRFs and changes in the higher-order moments in response to shocks.

Our work can be expanded in several directions. We mentioned already the multidimensional factor in the main text. Another fruitful avenue is to use the NLDF model with a VAR in the same fashion as the factor-augmented vector autoregression model (Stock and Watson, 2016).

References

T. Adrian, D. Covitz, and N. Liang. Financial Stability Monitoring. *Annual Review of Financial Economics*, 7(1):357–395, December 2015. doi: 10.1146/annurev-financial. URL https://ideas.repec.org/a/anr/refeco/v7y2015p357-395.html.

T. Adrian, N. Boyarchenko, and D. Giannone. Vulnerable Growth. *American Economic Review*, 109(4):1263–1289, April 2019. URL https://ideas.repec.org/a/aea/aecrev/v109y2019i4p1263-89.html.

M. M. Andreasen, J. Fernandez-Villaverde, and J. F. Rubio-Ramirez. The Pruned State-Space System for Non-Linear DSGE Models: Theory and Empirical Applications. *The Review of Economic Studies*, 85(1):1–49, 06 2017.

B. Aruoba and F. Diebold. Real-time macroeconomic monitoring: Real activity, inflation, and interactions. *American Economic Review Papers and Proceedings*, 100:10–24, 2010.

B. Aruoba, L. Bocola, and F. Schorfheide. Assessing dsge model nonlinearities. *Journal of Economic Dynamics and Control*, 83:34–54, 2017a.

S. B. Aruoba, P. Cuba-Borda, and F. Schorfheide. Macroeconomic Dynamics Near the ZLB: A Tale of Two Countries. *The Review of Economic Studies*, 85(1):87–118, 04 2017b. ISSN 0034-6527. doi: 10.1093/restud/rdx027. URL https://doi.org/10.1093/restud/rdx027.

S. B. Aruoba, M. Mlikota, F. Schorfheide, and S. Villalvazo. Svars with occasionally-binding constraints. *Journal of Econometrics*, 2021. ISSN 0304-4076. doi: https://doi.org/10.1016/j.jeconom.2021.07.013. URL https://www.sciencedirect.com/science/article/pii/S0304407621002487.

S. R. Baker, N. Bloom, and S. J. Davis. Measuring economic policy uncertainty*. *The Quarterly Journal of Economics*, 131(4):1593–1636, 2016. doi: 10.1093/qje/qjw024. URL http://dx.doi.org/10.1093/qje/qjw024.

M. Banbura and M. Modugno. Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data. *Journal of Applied Econometrics*, 29(1):133–160, 2014.

G. Benigno, P. Benigno, and S. Nistico. Second-order approximation of dynamic models with time-varying risk. *Journal of Economic Dynamics and Control*, 37(7):1231–1247, 2013.

B. S. Bernanke. The new tools of monetary policy. *American Economic Review*, 110(4):943–83, April 2020. doi: 10.1257/aer.110.4.943. URL https://www.aeaweb.org/articles?id=10.1257/aer.110.4.943.

B. S. Bernanke, M. Gertler, and S. Gilchrist. The financial accelerator in a quantitative business cycle framework. In J. B. Taylor and M. Woodford, editors, *Handbook of Macroeconomics*, volume 1 of *Handbook of Macroeconomics*, chapter 21, pages 1341–1393. Elsevier, December 1999. URL https://ideas.repec.org/h/eee/macchp/1-21.html.

N. Bloom. The Impact of Uncertainty Shocks. *Econometrica*, 77(3):623–685, May 2009. URL https://ideas.repec.org/a/ecm/emetrp/v77y2009i3p623-685.html.

P. Bordalo, N. Gennaioli, A. Shleifer, and S. J. Terry. Real Credit Cycles. NBER Working Papers 28416, National Bureau of Economic Research, Inc, Jan. 2021. URL https://ideas.repec.org/p/nbr/nberwo/28416.html.

M. K. Brunnermeier and Y. Sannikov. A Macroeconomic Model with a Financial Sector. *American Economic Review*, 104(2):379–421, February 2014. URL https://ideas.repec.org/a/aea/aecrev/v104y2014i2p379-421.html.

M. Carrasco and B. Rossi. In-sample inference and forecasting in misspecified factor models. *Journal of Business and Economic Statistics*, 3:313–338, 2016.

M. Chauvet. An econometric characterization of business cycles dynamics with factor structure and regime switching. *International Econometric Review*, 39(4):969–996, 1998.

M. Chen, I. Fernández-Val, and M. Weidner. Nonlinear factor models for network and panel data. *Journal of Econometrics*, 220(2):296 – 324, 2021. ISSN 0304-4076. doi: https://doi.org/10.1016/j.jeconom.2020.04.004. URL http://www.sciencedirect.com/science/article/pii/S0304407620301238. Annals Issue: Celebrating 40 Years of Panel Data Analysis: Past, Present and Future.

X. Cheng, Z. Liao, and F. Schorfheide. Shrinkage estimation of high-dimensional factor models with structural instabilities. *Review of Economic Studies*, 83:1151–1543, 2016.

R. K. Crump and N. Gospodinov. On the Factor Structure of Bond Returns. *Econometrica*, 90(1):295–314, January 2022. doi: 10.3982/ECTA17943. URL https://ideas.repec.org/a/wly/emetrp/v90y2022i1p295-314.html.

D. Debortoli, J. Gali, and L. Gambetti. On the Empirical (Ir)relevance of the Zero Lower Bound Constraint. In *NBER Macroeconomics Annual 2019, volume 34*, NBER Chapters, pages 141–170. National Bureau of Economic Research, Inc, January 2019. URL https://ideas.repec.org/h/nbr/nberch/14241.html.

M. Del Negro and C. Otrok. Dynamic factor models with time-varying parameters: measuring changes in international business cycles. Staff Reports 326, Federal Reserve Bank of New York, 2008. URL https://ideas.repec.org/p/fip/fednsr/326.html.

J. Fernández-Villaverde and J. F. Rubio-Ramírez. Estimating macroeconomic models: A likelihood approach. *The Review of Economic Studies*, 74(4):1059–1087, 2007.

J. Fernandez-Villaverde, G. Gordon, P. Guerron-Quintana, and J. F. Rubio-Ramirez. Nonlinear adventures at the zero lower bound. *Journal of Economic Dynamics and Control*, 57:182–204, 2015. ISSN 0165-1889.

J. Fernández-Villaverde, P. Guerrón-Quintana, K. Kuester, and J. Rubio-Ramírez. Fiscal volatility shocks and economic activity. *American Economic Review*, 105(11):3352–84, 2015a.

J. Fernández-Villaverde, P. Guerrón-Quintana, and J. F. Rubio-Ramírez. Estimating dynamic equilibrium models with stochastic volatility. *Journal of Econometrics*, 185(1):216–229, 2015b.

J. Fernandez-Villaverde, J. Rubio-Ramirez, and F. Schorfheide. Solution and estimation methods for dsge models. In J. Taylor and H. Uhlig, editors, *Handbook of Macroeconomics*, chapter 9, pages 557–724. North-Holland, 2016.

M. Gertler and P. Karadi. A model of unconventional monetary policy. *Journal of Monetary Economics*, 58(1):17–34, January 2011. URL https://ideas.repec.org/a/eee/moneco/v58y2011i1p17-34.html.

J. Geweke. Using simulation methods for bayesian econometric models: inference, development,and communication. *Econometric Reviews*, 18(1):1–73, 1999. doi: 10.1080/07474939908800428. URL https://ideas.repec.org/a/taf/emetrv/v18y1999i1p1-73.html.

S. Gilchrist and E. Zakrajsek. Credit spreads and business cycle fluctuations. *American Economic Review*, 102(4):1692–1720, June 2012.

J. Gomes, U. Jermann, and L. Schmid. Sticky Leverage. *American Economic Review*, 106(12):3800–3828, December 2016. URL https://ideas.repec.org/a/aea/aecrev/v106y2016i12p3800-3828.html.

S. Goncalves, A. M. Herrera, L. Kilian, and E. Pesavento. Impulse response analysis for structural dynamic models with nonlinear regressors. *Journal of Econometrics*, 2021. ISSN 0304-4076.

Y. Gorodnichenko and S. Ng. Level and volatility factors in macroeconomic data. *Journal of Monetary Economics*, 91:52–68, 2017.

C. Gourieroux, A. Monfort, and J.-P. Renne. Identification and Estimation in Non-Fundamental Structural VARMA Models. *The Review of Economic Studies*, 05 2019.

L. Guerrieri and M. Iacoviello. Collateral constraints and macroeconomic asymmetries. *Journal of Monetary Economics*, 90(C):28–49, 2017. doi: 10.1016/j.jmoneco.2017.06. URL https://ideas.repec.org/a/eee/moneco/v90y2017icp28-49.html.

R. S. Gurkaynak, B. Sack, and J. H. Wright. The U.S. Treasury yield curve: 1961 to the present. *Journal of Monetary Economics*, 54(8):2291–2304, November 2007. URL https://ideas.repec.org/a/eee/moneco/v54y2007i8p2291-2304.html.

C. Gust, E. Herbst, D. López-Salido, and M. E. Smith. The empirical implications of the interest-rate lower bound. *American Economic Review*, 107(7):1971–2006, 2017.

J. T. Hwang. Multiplicative errors-in-variables models with applications to recent data released by the u.s. department of energy. *Journal of the American Statistical Association*, 81(395):680–688, 1986.

O. Jorda, M. Schularick, and A. M. Taylor. Sovereigns Versus Banks: Credit, Crises, And Consequences. *Journal of the European Economic Association*, 14(1):45–79, February 2016. URL https://ideas.repec.org/a/bla/jeurec/v14y2016i1p45-79.html.

A. Justiniano and G. E. Primiceri. The time-varying volatility of macroeconomic fluctuations. *American Economic Review*, 98(3):604–41, 2008.

J. Kim, S. Kim, E. Schaumburg, and C. A. Sims. Calculating and using second-order accurate solutions of discrete time dynamic equilibrium models. *Journal of Economic Dynamics and Control*, 32(11):3397–3414, 2008. ISSN 0165-1889. doi: https://doi.org/10.1016/j.jedc.2008.02.003. URL https://www.sciencedirect.com/science/article/pii/S0165188908000316.

G. Koop, M. H. Pesaran, and S. M. Potter. Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics*, 74(1):119–147, September 1996. URL https://ideas.repec.org/a/eee/econom/v74y1996i1p119-147.html.

F. Lindsten, M. I. Jordan, and T. B. Schön. Particle gibbs with ancestor sampling. *Journal of Machine Learning Research*, 15(63):2145–2184, 2014. URL http://jmlr.org/papers/v15/lindsten14a.html.

S. C. Ludvigson, S. Ma, and S. Ng. Uncertainty and business cycles: Exogenous impulse or endogenous response? *American Economic Journal: Macroeconomics*, 13(4):369–410, October 2021.

A. Mian, A. Sufi, and E. Verner. Household Debt and Business Cycles Worldwide. *The Quarterly Journal of Economics*, 132(4):1755–1817, 2017. URL https://ideas.repec.org/a/oup/qjecon/v132y2017i4p1755-1817..html.

H. P. Minsky. The Financial Instability Hypothesis: An Interpretation of Keynes and an Alternative to Standard Theory. *Challenge*, 20(1):20–27, March 1977. doi: 10.1080/05775132.1977.114. URL https://ideas.repec.org/a/mes/challe/v20y1977i1p20-27.html.

A. Onatski and C. Wang. Spurious Factor Analysis. *Econometrica*, 89(2):591–614, March 2021. doi: 10.3982/ECTA16703. URL https://ideas.repec.org/a/wly/emetrp/v89y2021i2p591-614.html.

S. Särkkä. *Bayesian filtering and smoothing*, volume 3. Cambridge University Press, 2013.

M. Schularick and A. M. Taylor. Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870-2008. *American Economic Review*, 102(2):1029–1061, April 2012. URL https://ideas.repec.org/a/aea/aecrev/v102y2012i2p1029-61.html.

M. Shintani. Nonlinear forecasting analysis using diffusion indexes: An application to japan. *Journal of Money, Credit, and Banking*, 37(3):517–538, 2005.

J. Stock and M. Watson. Factor models and structural vector autoregressions in macroeconomics. In J. Taylor and H. Uhlig, editors, *Handbook of Macroeconomics*, chapter 8, pages 415–526. North-Holland, 2016.

E. T. Swanson and J. C. Williams. Measuring the Effect of the Zero Lower Bound on Medium- and Longer-Term Interest Rates. *American Economic Review*, 104(10):3154–3185, October 2014. URL https://ideas.repec.org/a/aea/aecrev/v104y2014i10p3154-85.html.

J. C. Wu and F. D. Xia. Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound. *Journal of Money, Credit and Banking*, 48(2-3):253–291, March 2016. URL https://ideas.repec.org/a/wly/jmoncb/v48y2016i2-3p253-291.html.

J. C. Wu and J. Zhang. A shadow rate New Keynesian model. *Journal of Economic Dynamics and Control*, 107(C):1–1, 2019. doi: 10.1016/j.jedc.2019.10372. URL https://ideas.repec.org/a/eee/dyncon/v107y2019ic7.html.

Appendices

A Estimation Algorithms

In this section of the appendix, we provide further details about the two estimation algorithms that we use. The first one is a Metropolis Hastings algorithm using the particle filter. The second is a particle Gibbs sampling algorithm. We refer to our benchmark NLDF model, shown again in Equation 24 for convenience.

$$
\begin{aligned}
&y_t = G f_t + \eta \epsilon_t \\
&\begin{cases} f_t = c + f_t^f + f_t^s \\ f_t^f = h_x f_{t-1}^f + \sigma \nu_t \\ f_t^s = h_x f_{t-1}^s + \frac{1}{2} h_{xx} \left(f_{t-1}^f \right)^2 . \end{cases}
\end{aligned} \tag{24}
$$

Here, we assume that $c = -\frac{1}{2} \frac{h_{xx}\sigma^2}{(1-h_x)(1-h_x^2)}$.

A.1 Metropolis Hastings with Bootstrap Particle Filter

Our Metropolis Hastings algorithm is as follows:

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

   - In practice, we break up the proposals into three blocks: Block 1 (factor equation) $\Theta_1^{prop} = \left\{ h_x^{prop}, h_{xx}^{prop}, \sigma^{2,prop} \right\}$; Block 2 (measurement equation loadings) $\Theta_2^{prop} = \{ G^{prop} \}$; and Block 3 (measurement equation variances) $\Theta_3^{prop} = \{ \eta^{prop} \}$. For each block, we take 50 draws, holding the parameters in the other blocks 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 previous 250 draws within the block had an acceptance rate less than 10%.

2. Evaluate the likelihood of the proposed parameters using the bootstrap particle filter (Särkkä, 2013).

   - **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: $f_0^{f,(j)}, f_0^{s,(j)}, f_1^{s,(j)}$. Note that $f_1^{s,(j)}$ is a function of $f_0^{f,(j)}, f_0^{s,(j)}$, so it is known. We set $w_t^{(j)} = 1$ for all particles.

     For $t = 1, ..., T$:

   - **Prediction step:**

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

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

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

     (b) Calculate weights:

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

   - **Update step:**

     (a) Define normalized weights: $\widetilde{w}_t^{(j)} = \frac{\omega_t^{(j)} w_{t-1}^{(j)}}{\frac{1}{N}\sum \omega_t^{(j)} w_{t-1}^{(j)}}$.

     (b) Resample from multinomial distribution $\left\{\omega_t^{(j)}, \widetilde{w}_t^{(j)}\right\}$ and set $w_t^{(j)} = 1$.

   - **Compute conditional likelihood:**

     $$p(y_t | Y_{1:t-1}) \approx \frac{1}{N} \sum_{i=1}^{N} \omega_t^{(j)} w_{t-1}^{(j)}. \tag{25}$$

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

3. We accept the proposal with probability

   $$prob = \max\left\{\frac{p(y|\Theta_i^{prop}, \Theta_{-i}^{curr}) g(\Theta_i^{prop}, \Theta_{-i}^{curr})}{p(y|\Theta_i^{curr}, \Theta_{-i}^{curr}) g(\Theta_i^{curr}, \Theta_{-i}^{curr})}, 1\right\} \tag{26}$$

   where $g(.)$ is the prior distribution.

A.2 Gibbs Sampling with Particle Smoother

Our Gibbs sampling algorithm is as follows:

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

2. Draw h_x, h_{xx} given σ, G, η, f_t^f, f_t^s, and y_t.

   We use a random walk Metropolis step to draw h_x and h_{xx}. Given the current accepted draw of h_x and h_{xx}, our proposal is as follows:

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

   We throw away draws that violate the stationarity conditions $h_x^{prop} > 1$.

   Given proposed h_x^{prop} and h_{xx}^{prop}, we calculate its likelihood. The new parameters change c and f_t^s.

   We update

$$c^{prop} = -\frac{1}{2} \frac{h_{xx}^{prop} \sigma^2}{(1 - h_x^{prop}) \left(1 - (h_x^{prop})^2\right)}$$

   and

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

   We initialize $f_0^{s,prop} = f_0^s$.

   We then form the likelihood of the proposal, which can be calculated in two parts. The first is based on the measurement equation and the second is from the transition equation of the first-order factor:

$$\begin{aligned} y_t - G\left(c^{prop} + f_t^f + f_t^{s,prop}\right) &= \eta \epsilon_t \\ f_t^f - h_x^{prop} f_{t-1}^f &= \sigma \nu_t. \end{aligned} \tag{27}$$

   We accept the proposal with probability:

$$prob = \max\left\{ \frac{\prod_{t=1}^T p\left(y_t | c^{prop}, G, \eta, f_t^f, f_t^{s,prop}\right) p_{trans}\left(f_t^f | h_x^{prop}, \sigma, f_{t-1}^f\right) g\left(h_x^{prop}, h_{xx}^{prop}\right)}{\prod_{t=1}^T p\left(y_t | c^{curr}, G, \eta, f_t^f, f_t^s\right) p_{trans}\left(f_t^f | h_x^{curr}, \sigma, f_{t-1}^f\right) g\left(h_x^{curr}, h_{xx}^{curr}\right)}, 1 \right\} \tag{28}$$

   where $p(y_t|.)$ denotes the likelihood from the measurement equation, $p_{trans}(f_t^f|.)$ denotes the likelihood from the transition equation, and $g(.)$ is the prior distribution.

3. Draw σ^2 given G, η, f_t^f, f_t^s, and y_t.

   We draw σ^2 using a random walk Metropolis step. Given the current accepted draw of σ^2, our proposal is as follows:

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

   We throw out draws that are negative.

   Given the proposed $\sigma^{2,prop}$, we calculate its likelihood. The new parameters change c.

   We update

$$c^{prop} = -\frac{1}{2}\frac{h_{xx}\sigma^{2,prop}}{(1-h_x)(1-h_x^2)}. \tag{29}$$

   We then form the likelihood of the proposal, which can be calculated in two parts. The first is based on the measurement equation, and the second is from the transition equation of the first-order factor:

$$\begin{aligned} y_t - G\left(c^{prop} + f_t^f + f_t^s\right) &= \eta\epsilon_t \\ f_t^f - h_x f_{t-1}^f &= \sigma^{prop}\nu_t. \end{aligned} \tag{30}$$

   We accept the proposal with probability:

$$prob = \max\left\{\frac{\prod_{t=1}^T p\left(y_t|c^{prop}, G, \eta, f_t^f, f_t^s\right) p_{trans}\left(f_t^f|h_x, \sigma^{2,prop}, f_{t-1}^f\right) g\left(\sigma^{2,prop}\right)}{\prod_{t=1}^T p\left(y_t|c^{curr}, G, \eta, f_t^f, f_t^s\right) p_{trans}\left(f_t^f|h_x, \sigma^{2,curr}, f_{t-1}^f\right) g\left(\sigma^{2,curr}\right)}, 1\right\} \tag{31}$$

   where $p(y_t|.)$ denotes the likelihood from the measurement equation, $p_{trans}(f_t^f|.)$ denotes the likelihood from the transition equation, and $g(.)$ is the prior distribution.

4. 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. We discuss our implementation of the sampler here, but further details of the algorithm can be found in Lindsten et al. (2014).

   - **Initialize particle smoother:** For particles $j = 1, ..., N-1$. 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: $f_0^{f,(j)}, f_0^{s,(j)}, f_1^{s,(j)}$. 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 $*$ denotes 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:

$$\begin{aligned} f_t^{f,(j)} &= h_x f_{t-1}^{f,(a_t^{(j)})} + \sigma \nu_t \\ f_{t+1}^{s,(j)} &= h_x f_t^{s,a_t^{(j)}} + \frac{1}{2} h_{xx} \left(f_t^{f,(j)} \right)^2 \end{aligned} \tag{32}$$

- **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:

$$w_t^{aux,(j)} = w_{t-1}^{(j)} p(y_t|f_t^{f,(N)}, f_t^{s,(j)}) g(f_t^{f,(N)}|f_{t-1}^{f,(j)}) p(y_{t+1}|f_{t+1}^{f,(N)}, f_{t+1}^{s,(N')}) g(f_{t+1}^{f,(N)}|f_t^{f,(N)}). \tag{33}$$

  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

$$f_{t+1}^{s,(N')} = h_x f_t^{s,(j)} + \frac{1}{2} h_{xx} \left(f_t^{f,(N)} \right)^2. \tag{34}$$

  Note that this formula comes from Equation 23 in Lindsten et al. (2014) with lag $= 2$. Our model is a degenerate state-space model discussed in Section 7.2 of that paper. We can view our model alternatively as a non-Markovian model with one factor f_t^f. See the associated discussion there.
- **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:

$$f_{t+1}^{s,(N)} = h_x f_t^{s,(a_t^{(N)})} + \frac{1}{2} h_{xx} \left(f_t^{f,(N)} \right)^2. \tag{35}$$

- **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 $j = 1, ..., N$.

Note that for the $t = T$, we do not have to update f^s_{t+1} because it is the end of the sample. When computing the auxiliary weights for the fixed particle, we also do not consider the $T + 1$ likelihood.

- **Sample selected states:** Sample $*$ according to w_T. Set $f^{f,*}_t, f^{s,*}_t$ equal to the sampled state.

B Monte Carlo Results

To better understand the estimation of our model, we turn to a Monte Carlo experiment. Here, we show that if the true data-generating process is the NLDF model, our estimation strategy successfully recovers all parameters. We assume that the underlying model is our benchmark NLDF model with the following parameters: $c = 0, h_x = 0.85$, $h_{xx} = 2.15$, $\sigma = 0.18$, diag$(\eta) = [0.54, 0.06, 0.79, 1.08, 0.39]$, and $G = [1, 0.17, 1.5, 2.21, 0.56]$. We generate 50 series of length $T = 1000$, starting from $f_0^f = 0, f_0^s = 0$.

With the synthetic data in hand, we then estimate the linear factor model and our benchmark NLDF model with a linear measurement equation. The models are estimated using the Metropolis Hastings and particle filter procedure detailed in Section 2.4 with $200,000$ Markov Chain Monte Carlo draws. We assume flat priors for all of the parameters.

The parameter estimates converge to the true values under correct specification. As seen in the left panel in Figure 9, the log likelihood is higher for the nonlinear model (vertical axis) than it is for the linear one (horizontal axis) across all simulations. The average difference between the log likelihoods in the nonlinear and linear models is 80 points; the difference can be as low as 48 points and as high as 127 points. Correspondingly, the mean square errors of the factors are smaller in the nonlinear factor version (right panel in Figure 9).[16]

We report the estimates of the state equation's parameters in Figures 10 and 11. Whereas the nonlinear factor model's estimate for h_x (y-axis) is clustered around its true value, the linear estimate (x-axis) is about 14% more persistent. This over-persistence is compensated for with a downward bias estimate of the factor innovation volatility. This compensation is needed so the factor delivers second moments consistent with the data. In contrast, the volatility estimate from the NLDF model is around the true value. Furthermore, the second-order component (h_{xx}) is estimated close to its true value.

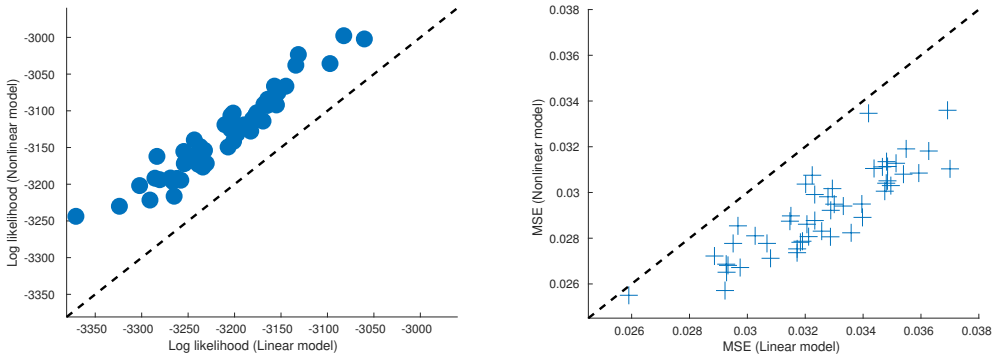


Figure 9: Performance of Estimated Linear and Nonlinear Models on Simulated Data

[16]The mean square error is defined as $\frac{\sum_{t=1}^{T=1000}(\hat{f}_{t|t}-f_t)^2}{T}$, where $\hat{f}_{t|t}$ is the factor filtered from the estimated model (linear or nonlinear), and f_t is the true simulated factor.

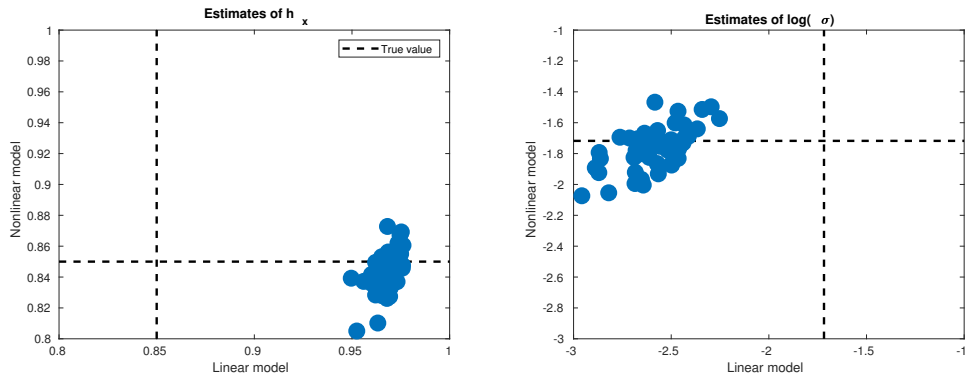


Figure 10: Estimation Bias of Linear and Nonlinear Models on Simulated Data

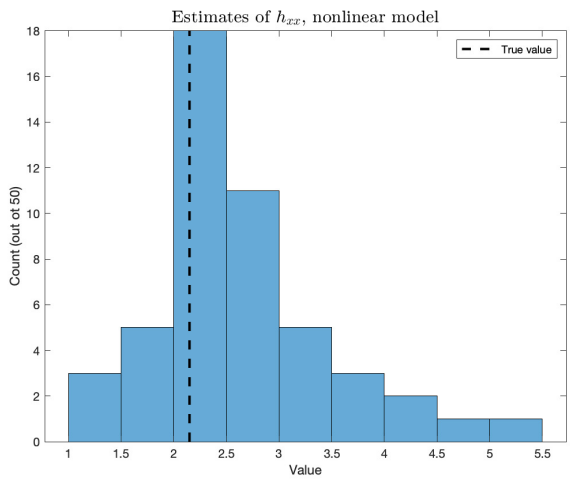


Figure 11: Performance of Nonlinear Model in Estimation of h_{xx} on Simulated Data

C Further Simulation Results

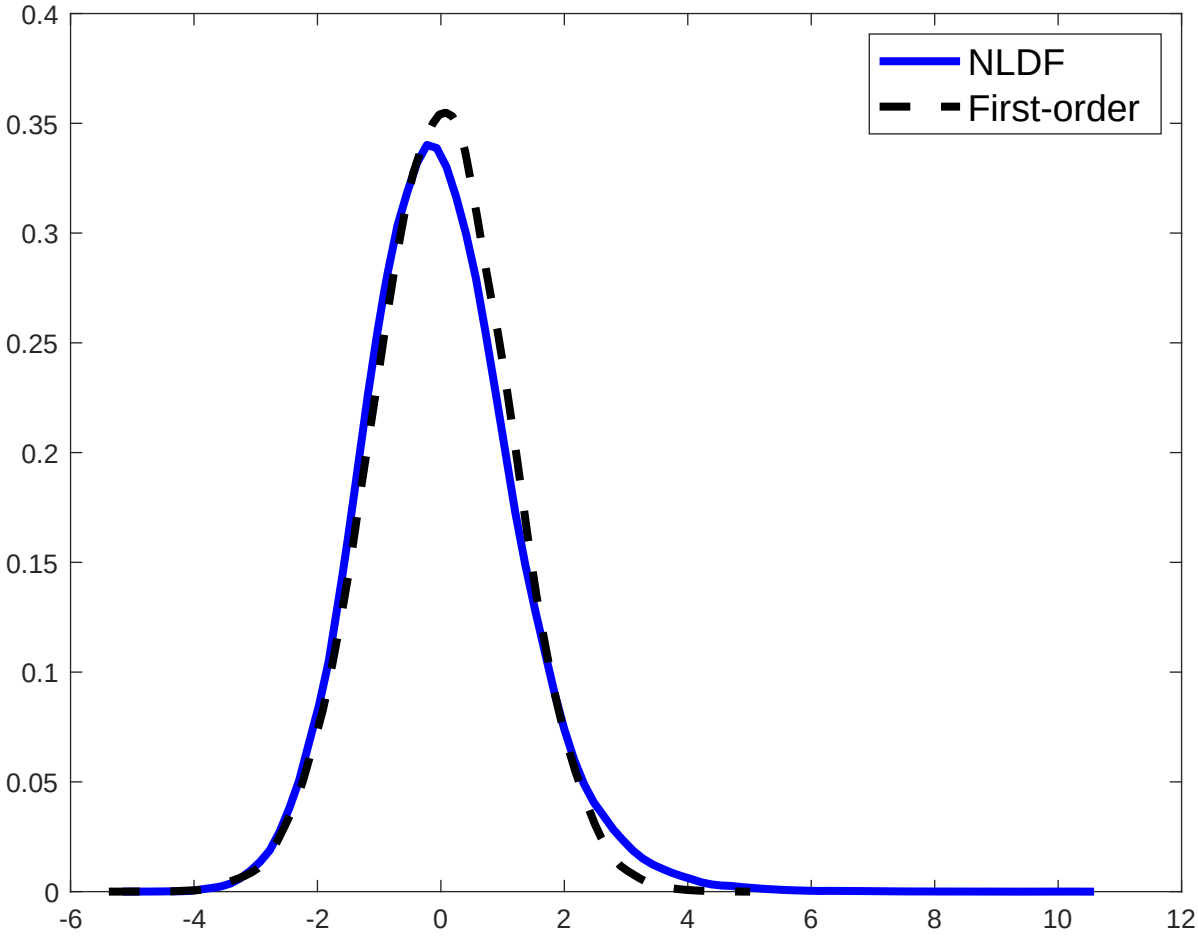


Figure 12: Unconditional distribution of the overall factor in a calibrated model. The blue line denotes the unconditional distribution of the demeaned overall factor in the nonlinear dynamic factor model. The dashed black line denotes the unconditional distribution of the first-order factor.

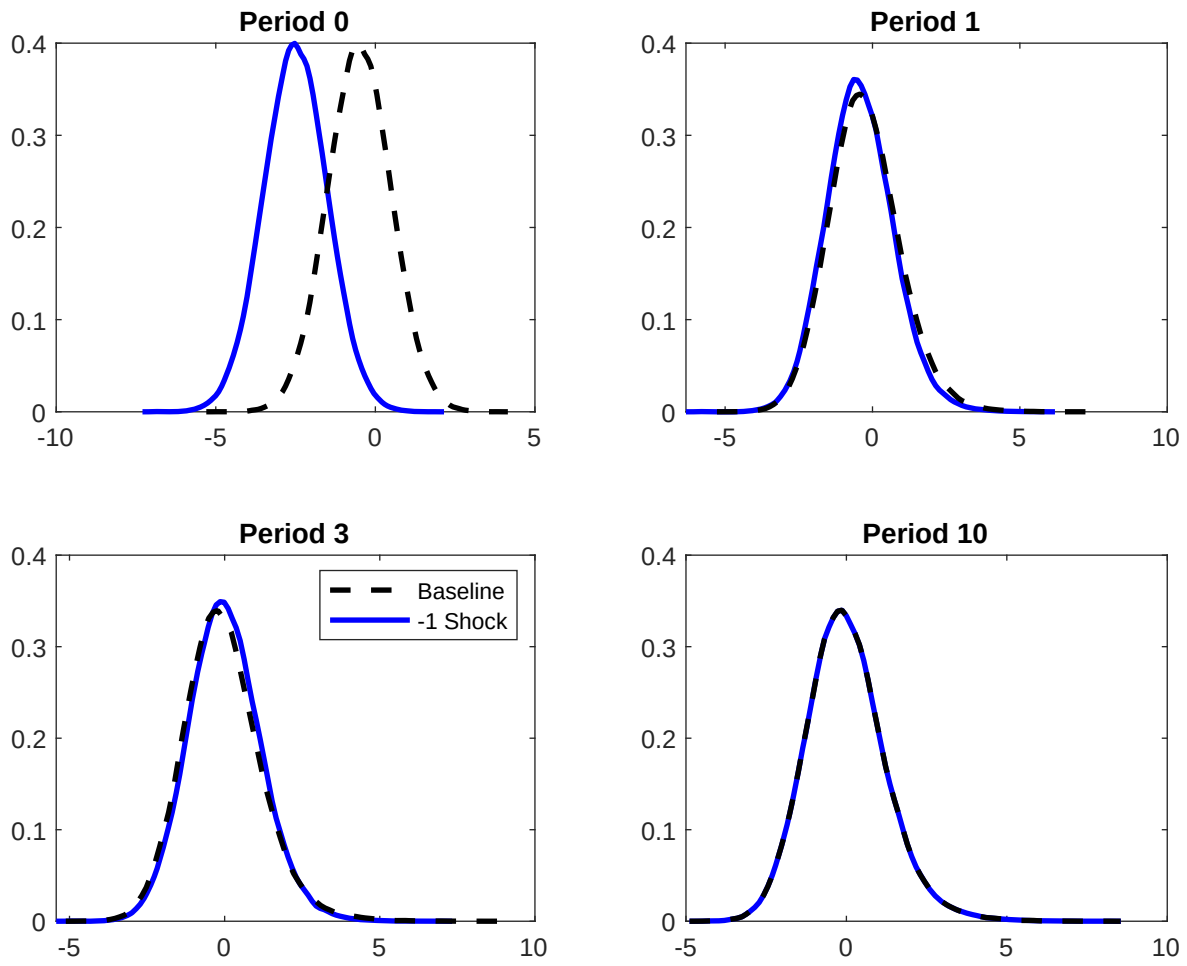


Figure 13: Dynamics of the distribution of the demeaned overall factor at various periods after a shock in a calibrated model. The blue line is the distribution after a negative shock, and the dashed black line is the distribution without a shock. Period 0 is the period of the shock.

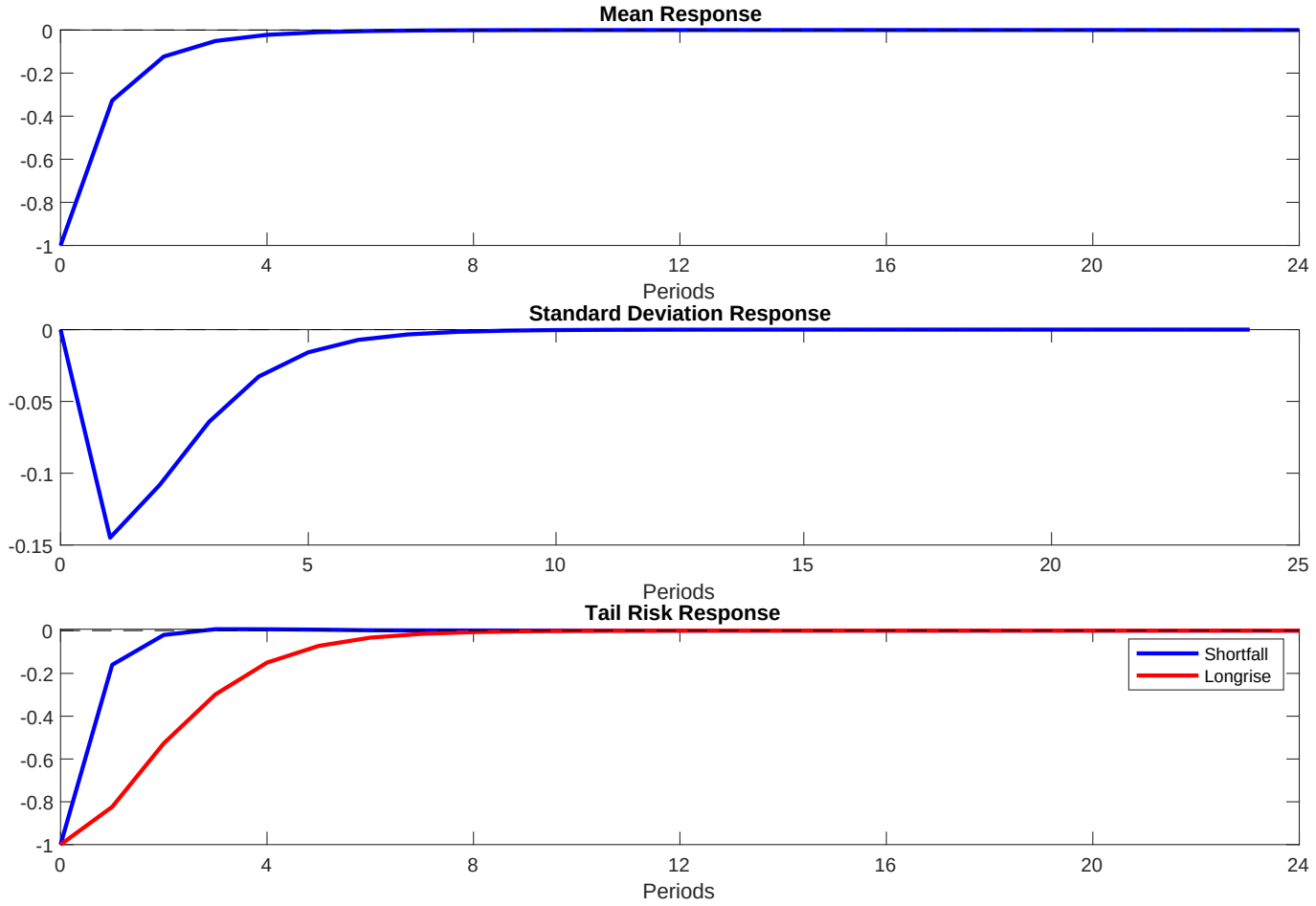


Figure 14: Impulse response functions of the mean, standard deviation, and tail risk of the demeaned overall factor in response to a negative shock at period 0 in a calibrated model. In the third panel, the blue line shows the response of the shortfall, and the red line shows the response of the longrise.

D Shadow Rate: Additional Results

D.1 Details on Data Construction

We followed the approach of Wu and Xia (2016) in constructing the one-month forward rates for seven maturities based on the nominal yield curve data from Gurkaynak et al. (2007). We use the code provided by Wu and Xia (2016). The end-of-month monthly data spans the period from January 1990 to September 2019, and the maturities used are the same as in the original paper: 3 and 6 months and 1, 2, 5, 7, and 10 years.

We download the Wu and Xia (2016) shadow rate from Cynthia Wu's website.

D.2 Tailoring the Estimation to the Shadow Rate Model

We discuss here how we modify the estimation presented in Appendix Section A.1 to account for the nonlinear measurement equation in Equation 22, reproduced below for convenience.

$$\Delta forward_t^h = m_h + \begin{cases} G_h(c + f_t^f + f_t^s) + \eta^h \varepsilon_t^h & \text{if } \widehat{S}_t^h >= 0.3 \\ -m_h + \eta^h \varepsilon_t^h & \text{otherwise} \end{cases} \tag{36}$$

where $c = -\frac{1}{2}\frac{h_{xx}\sigma^2}{(1-h_x)(1-h_x^2)}$, $\widehat{S}_t^h = \sum_{\tau=2}^{t}\left(m_h + G_h\left(c + f_\tau^f + f_\tau^s\right)\right) + forward_1^h$, $\Delta forward_t^h = forward_t^h - forward_{t-1}^h$, and index h stands for the maturity. We model the latent factor according to our second-order dynamics:

$$\begin{aligned} f_t^f &= h_x f_{t-1}^f + \sigma \nu_t \\ f_t^s &= h_x f_{t-1}^s + \frac{1}{2} h_{xx}\left(f_{t-1}^f\right)^2. \end{aligned} \tag{37}$$

Relative to the benchmark NLDF with a linear measurement equation, there are three main differences. First, we have forward-rate-specific constants m_h that capture the long-run mean of each series. This is a straightforward addition to the Metropolis Hastings algorithm, and we add a block to the estimation procedure. Second, we have to keep track of $\widehat{S}_t$ in the particle filter, which is the sum of the entire path of the particle. To account for this sum, we add an additional component to the particle called

$$\widehat{S}_t^{(j)} = \left(c + f_t^{f,(j)} + f_t^{s,(j)}\right) + \widehat{S}_{t-1}^{(j)}.$$

The conversion from $\widehat{S}_t^{(j)}$ to $\widehat{S}_t^{h,(j)}$ for each maturity h is straightforward from their respective formulas.

Third, the measurement equation that we use to evaluate the weight of the particle in the particle filter changes depending on whether $\widehat{S}_t^{h,(j)}$ is greater than or less than 0.3. This feature affects the prediction step in our algorithm.

D.3 Estimation of the Linear Model

We also estimate a linear version of the model that removes the nonlinearity in the measurement equation and only allows for first-order factor dynamics. We use the same Metropolis Hastings scheme as laid out in Appendix Section A.1, with two differences. First, we have the extra parameters m_h that we estimate as an additional block. Second, we use the Kalman Filter instead of the particle filter to estimate the model, as it is now a linear Gaussian state-space model.

D.4 Parameter Estimates

Table 2: Parameter Estimates

	Prior	NLDFM and ELB	Linear factor and ELB	Linear
h_x	$N(0.5, 1)$	0.192	0.183	0.188
		(0.132 , 0.249)	(0.118 , 0.242)	(0.126 , 0.261)
h_{xx}	$N(0, 5)$	0.272	0.000	0.000
		(-0.010 , 0.474)	(0.000 , 0.000)	(0.000 , 0.000)
σ^2	$IW(v = 4, \eta = 1)$	0.050	0.055	0.031
		(0.045 , 0.056)	(0.048 , 0.062)	(0.026 , 0.035)
G_1	$N(0, 5)$	1.000	1.000	1.000
		(1.000 , 1.000)	(1.000 , 1.000)	(1.000 , 1.000)
G_2	$N(0, 5)$	1.171	1.110	1.290
		(1.121 , 1.218)	(1.059 , 1.174)	(1.209 , 1.383)
G_3	$N(0, 5)$	1.410	1.354	1.623
		(1.356 , 1.464)	(1.306 , 1.422)	(1.542 , 1.735)
G_4	$N(0, 5)$	1.466	1.385	1.704
		(1.418 , 1.513)	(1.300 , 1.483)	(1.607 , 1.827)
G_5	$N(0, 5)$	1.022	0.947	1.123
		(0.944 , 1.099)	(0.860 , 1.034)	(1.011 , 1.248)
G_6	$N(0, 5)$	0.779	0.775	0.809
		(0.708 , 0.871)	(0.682 , 0.878)	(0.686 , 0.920)
G_7	$N(0, 5)$	0.635	0.616	0.575
		(0.580 , 0.706)	(0.537 , 0.712)	(0.451 , 0.685)
η_1^2	$IW\left(v = 4, \eta = \frac{1}{5} Std\left(\Delta forward^1\right)\right)$	0.015	0.016	0.019
		(0.013 , 0.016)	(0.014 , 0.017)	(0.018 , 0.021)
η_2^2	$IW\left(v = 4, \eta = \frac{1}{5} Std\left(\Delta forward^2\right)\right)$	0.008	0.008	0.010
		(0.007 , 0.009)	(0.008 , 0.009)	(0.009 , 0.011)
η_3^2	$IW\left(v = 4, \eta = \frac{1}{5} Std\left(\Delta forward^3\right)\right)$	0.002	0.001	0.002
		(0.001 , 0.002)	(0.001 , 0.002)	(0.002 , 0.003)
η_4^2	$IW\left(v = 4, \eta = \frac{1}{5} Std\left(\Delta forward^4\right)\right)$	0.013	0.013	0.015
		(0.012 , 0.015)	(0.012 , 0.015)	(0.014 , 0.017)
η_5^2	$IW\left(v = 4, \eta = \frac{1}{5} Std\left(\Delta forward^5\right)\right)$	0.043	0.045	0.055
		(0.039 , 0.047)	(0.041 , 0.051)	(0.050 , 0.061)
η_6^2	$IW\left(v = 4, \eta = \frac{1}{5} Std\left(\Delta forward^6\right)\right)$	0.057	0.058	0.070
		(0.052 , 0.064)	(0.052 , 0.064)	(0.065 , 0.078)
η_7^2	$IW\left(v = 4, \eta = \frac{1}{5} Std\left(\Delta forward^7\right)\right)$	0.063	0.064	0.071
		(0.058 , 0.068)	(0.059 , 0.072)	(0.065 , 0.079)

Median values of the posterior are reported. 10% and 90% are shown in brackets. Log likelihoods are reported at the mode. NLDFM is nonlinear dynamic factor model. ELB is effective lower bound.

	Prior	NLDFM and ELB	Linear factor and ELB	Linear
m_1	$N(-0.019, 1)$	-0.019	-0.020	-0.019
		(-0.021 , -0.017)	(-0.022 , -0.017)	(-0.024 , -0.014)
m_2	$N(-0.019, 1)$	-0.021	-0.020	-0.019
		(-0.023 , -0.019)	(-0.022 , -0.018)	(-0.024 , -0.015)
m_3	$N(-0.019, 1)$	-0.022	-0.021	-0.019
		(-0.025 , -0.020)	(-0.024 , -0.019)	(-0.024 , -0.015)
m_4	$N(-0.019, 1)$	-0.022	-0.020	-0.019
		(-0.025 , -0.019)	(-0.024 , -0.018)	(-0.024 , -0.015)
m_5	$N(-0.019, 1)$	-0.010	-0.008	-0.019
		(-0.012 , -0.008)	(-0.011 , -0.006)	(-0.024 , -0.014)
m_6	$N(-0.019, 1)$	-0.007	-0.007	-0.019
		(-0.009 , -0.005)	(-0.009 , -0.005)	(-0.025 , -0.014)
m_7	$N(-0.019, 1)$	-0.005	-0.005	-0.019
		(-0.007 , -0.004)	(-0.007 , -0.004)	(-0.024 , -0.014)
LL		723.731	723.742	636.693

Median values of the posterior are reported. 10% and 90% are shown in brackets. Log likelihoods are reported at the mode. NLDFM is nonlinear dynamic factor model. ELB is effective lower bound.

E Nonlinear Credit Cycle: Additional Results

E.1 Details on Data Construction

Our data extend from 1952:Q1 through 2021:Q4 at a quarterly frequency. Our data are from the Statistical Release Z.1 "Financial Accounts of the United States" and were downloaded from the Federal Reserve Bank of St. Louis, Federal Reserve Economic Data. These data are not seasonally adjusted, and we seasonally adjust them using the Census X-13 Seasonal Adjustment procedure implemented in Eviews 12. We deflate the seasonally adjusted data by the seasonally adjusted GDP deflator to turn them into real values.

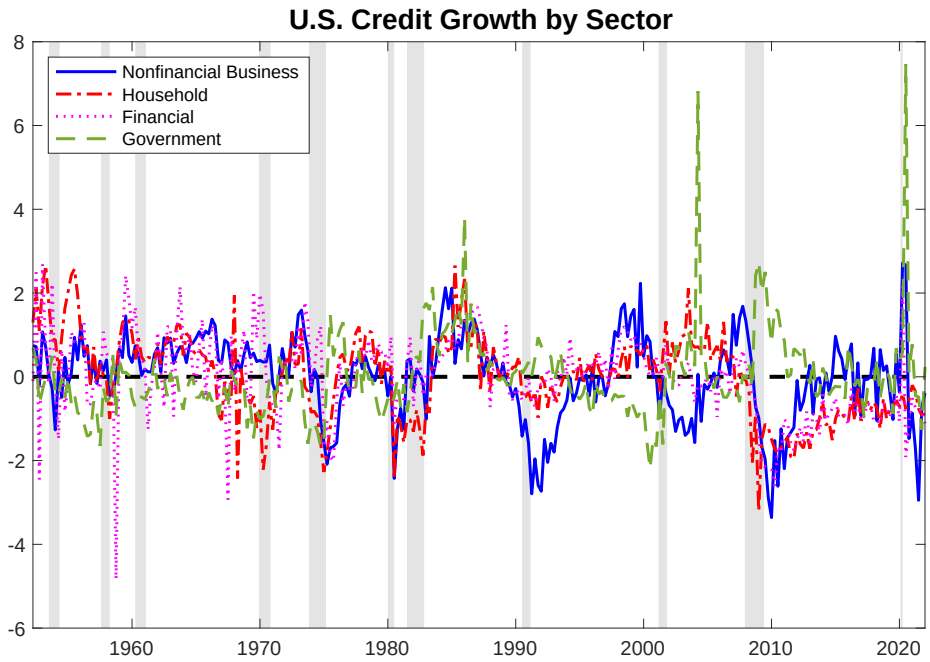


Figure 15: Normalized real credit growth by sector in the United States: 1952:Q1-2021:Q4 with National Bureau of Economic Research recession shading.

For nonfinancial business debt, we use the category Nonfinancial Business, Debt Securities and Loans, Liability, Level (BOGZ1FL144104005Q). For household debt, we use the category Households and Nonprofit Organizations, Debt Securities and Loans, Liability, Level (TCMILBSHNO). For financial-sector debt, we use the category Domestic Financial Sectors, Debt Securities and Loans, Liability, Level (TCMDODFS). Finally, for government debt, we sum the categories Federal Government, Debt Securities and Loans, Liability, Level (FGTCMDODNS) and State and Local Governments, Debt Securities and Loans, Liability, Level (SLGTCMDODNS). We seasonally adjust the federal and state and local government debt separately before summing them up.

E.2 Difference Between Positive and Negative Shocks

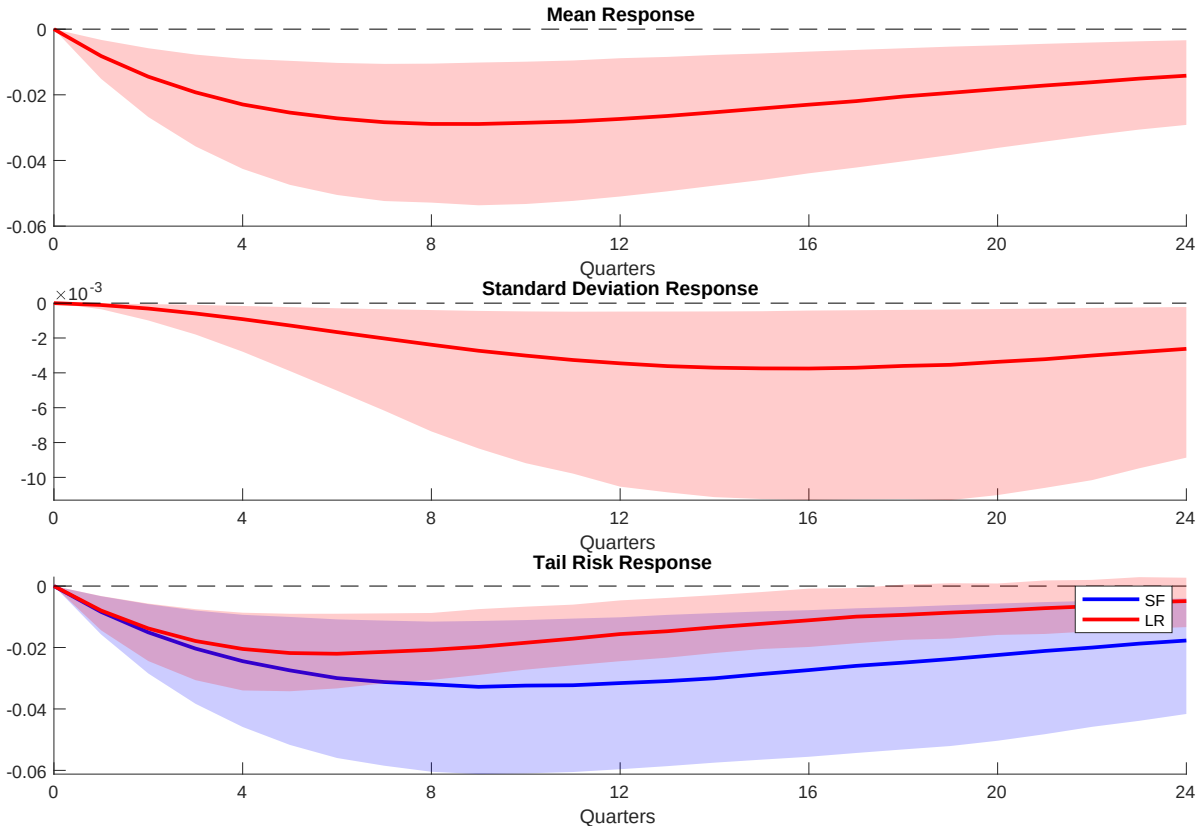


Figure 16: Credit Boom State (mid-2000s): Draw-by-draw differences between positive and negative shocks on the mean, standard deviation, and tail risk responses. Shaded areas denote 68% credible sets. SF is shortfall and LR is longrise.

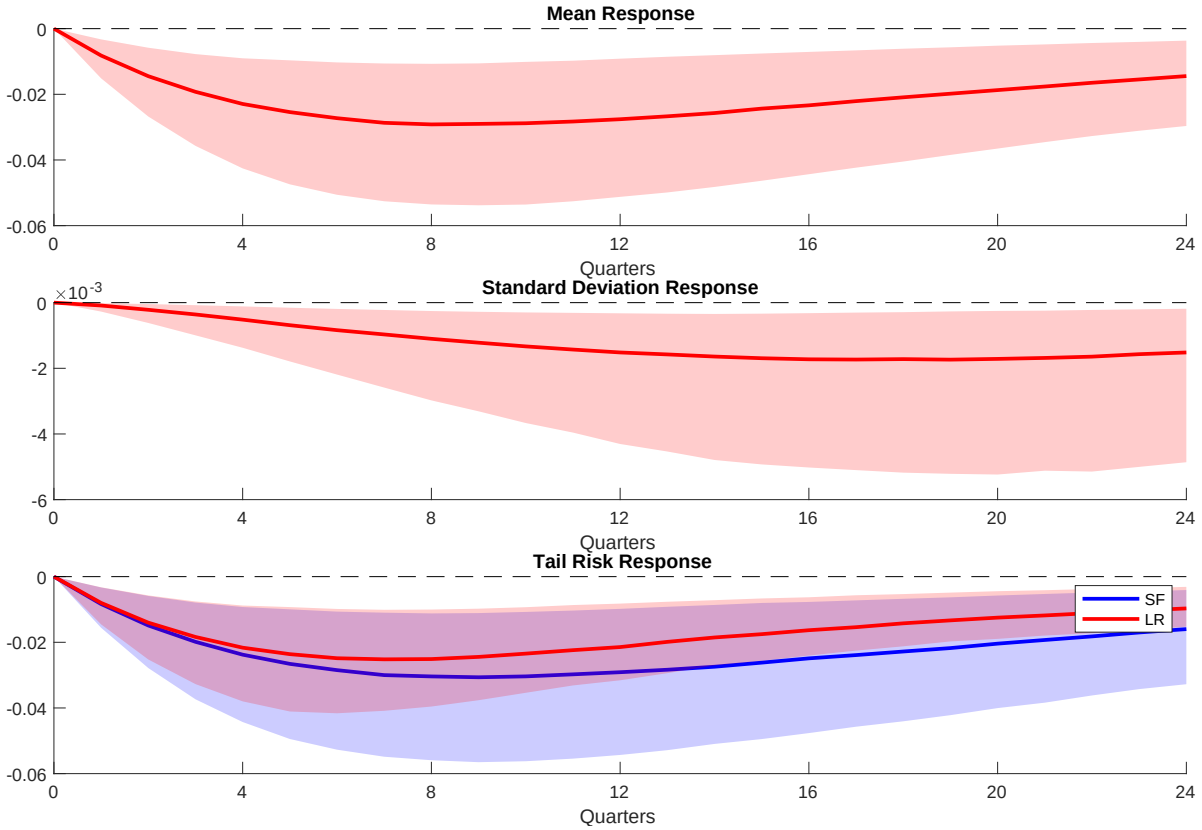


Figure 17: Credit Crunch State (2010): Draw-by-draw differences between positive and negative shocks on the mean, standard deviation, and tail risk responses. Shaded areas denote 68% credible sets. SF is shortfall and LR is longrise.

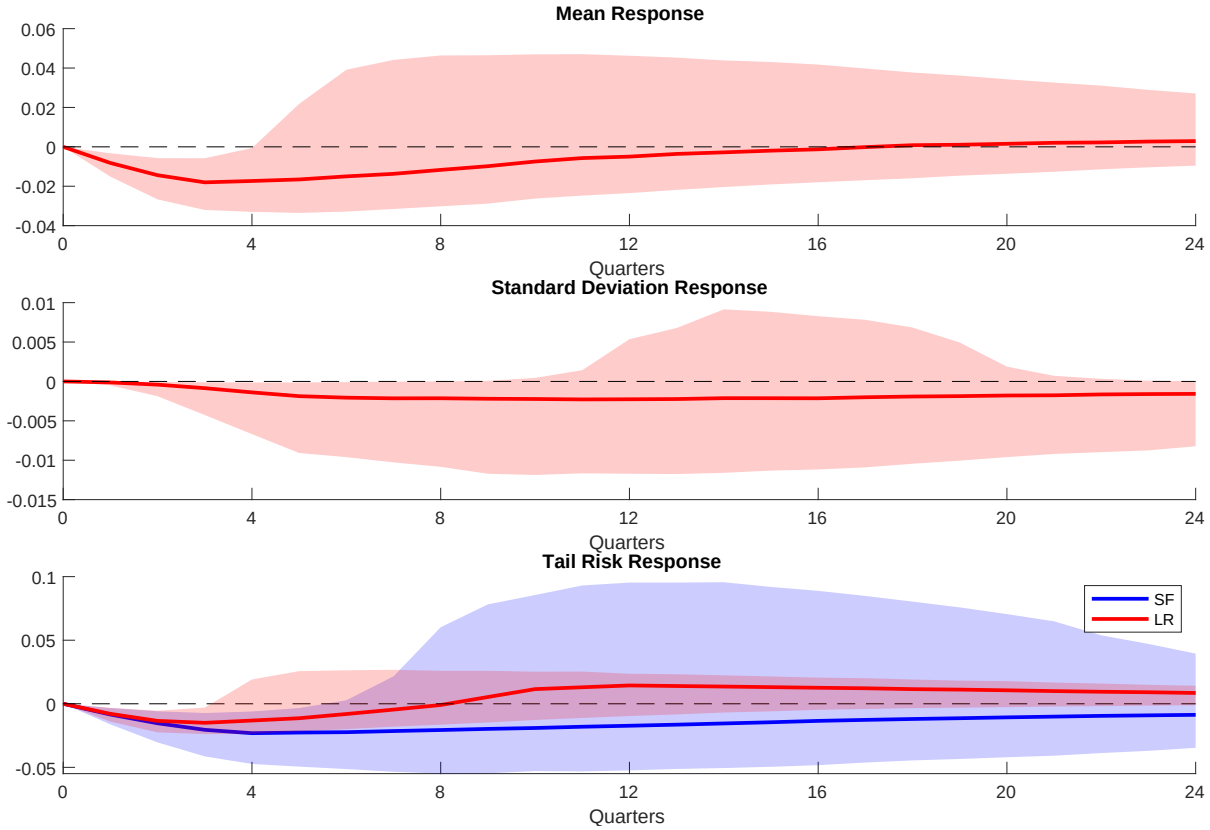


Figure 18: Credit Mix State (before the early 1990s recession): Draw-by-draw differences between positive and negative shocks on the mean, standard deviation, and tail risk responses. Shaded areas denote 68% credible sets. SF is shortfall and LR is longrise.

E.3 Parameter Estimates

Table 3: Parameter Estimates

	Prior	NLDFM
h_x	$N(0.5, 1)$	0.922
		(0.901 , 0.938)
h_{xx}	$N(0, 5)$	-0.130
		(-0.223 , -0.053)
σ^2	$IW(v = 4, \eta = 1)$	0.062
		(0.050 , 0.077)
G_1	$N(0, 5)$	1.000
		(1.000 , 1.000)
G_2	$N(0, 5)$	1.318
		(1.162 , 1.470)
G_3	$N(0, 5)$	0.983
		(0.881 , 1.100)
G_4	$N(0, 5)$	-0.220
		(-0.327 , -0.116)
η_1^2	$IW\ (v = 4, \eta = 1)$	0.626
		(0.551 , 0.698)
η_2^2	$IW\ (v = 4, \eta = 1)$	0.282
		(0.234 , 0.347)
η_3^2	$IW\ (v = 4, \eta = 1)$	0.592
		(0.537 , 0.651)
η_4^2	$IW\ (v = 4, \eta = 1)$	0.968
		(0.888 , 1.05)

Median values of the posterior are reported. 16% and 84% are shown in brackets. NLDFM is nonlinear dynamic factor model.