Large Bayesian VARs for Binary, Censored and Count Variables

Joshua C.C. Chan
Purdue University

Davide Pettenuzzo
Brandeis University

Michael Pfarrhofer
WU Vienna

Aubrey Poon
University of Kent

Dan Zhu
Monash University

September 2026

We extend the vector autoregression (VAR) to model binary, censored, count and unrestricted variables jointly, treating missing observations of any type in the same way. Each restricted observation is linked to a latent Gaussian variable, so that the standard machinery for VARs, such as shrinkage priors, stochastic volatility and outlier components, applies directly to the VAR for the latent variables. The latent variables have a truncated normal conditional distribution with a banded precision matrix, from which they are drawn, jointly or equation by equation, using Hamiltonian Monte Carlo, and the sampler scales well to high dimensions. A simulation study shows that treating restricted variables as continuous distorts the levels and scales of the system and worsens forecasts of the restricted outcomes in most designs. In an application to 25 US macro-financial variables, we estimate a business cycle indicator, the intensity of banking distress and shadow rates. We use the model for scenario analysis with restrictions on the recession probability, bank failures and interest rates at the effective lower bound.

JEL: C11, C32, C34, C35, C53, E32
KEYWORDS: Bayesian VAR, binary and censored variables, count data, effective lower bound, shadow rate, bank failures, conditional forecasting

Contact: Michael Pfarrhofer (michael.pfarrhofer@wu.ac.at). This paper merges our previous work "Modeling and Forecasting Count Data with Bayesian Vector Autoregressions" (Pettenuzzo, Poon and Zhu; `SSRN 5285954`) and "Large Bayesian VARs for Binary and Censored Variables" (Chan and Pfarrhofer; `arXiv 2506.01422`). Codes and replication files: github.com/mpfarrho/protoc-var.

1. INTRODUCTION

Since the seminal work of Sims (1980) and Doan *et al.* (1984), vector autoregressions (VARs) have become the workhorse models for macroeconomic forecasting and structural analysis. Because most macroeconomic variables are continuous, the literature has focused on real-valued time series. Many important macroeconomic and financial indicators, however, have restricted supports. Nominal interest rates are bounded below by their effective lower bound (ELB); recessions, financial crises and sovereign defaults are binary events; and bank failures, corporate bankruptcies and loan defaults are counts, as are many other outcomes in finance (Cohn *et al.*, 2022). Standard VARs are ill-suited for such series, and the growing use of large datasets, which typically contain several of them, makes the issue more acute (Bańbura *et al.*, 2010; Koop, 2013).

We develop an econometric framework that extends the VAR to a mix of binary, censored, count and unrestricted variables, with missing observations of any type handled in the same way. The approach is based on data augmentation: each restricted observation is linked to a real-valued latent variable, so that a binary variable records the sign of its latent variable as in a probit model, a censored variable is observed only above a threshold as in a tobit model, and a count variable records the interval between consecutive thresholds in which its latent variable falls, as in an ordered probit model. The thresholds for the counts are given by a link function from the Box–Cox family, which contains unit intervals and logarithmic thresholds as special cases. Given the latent variables, the machinery developed for standard VARs applies directly. In particular, the framework accommodates adaptive hierarchical shrinkage priors (Giannone *et al.*, 2015), which are essential in high dimensions, and stochastic volatility and outlier components, which have been shown to be empirically important for macroeconomic and financial variables (Cogley and Sargent, 2005; Clark, 2011; Carriero *et al.*, 2019; Chan, 2023; Carriero *et al.*, 2024).[1]

[1] Bayesian data augmentation for binary and censored variables goes back to Chib (1992) and Albert and Chib (1993); see Anceschi *et al.* (2023) for a review.

Sampling the latent variables is challenging in a high-dimensional dynamic setting for two reasons. First, the latent variables are subject to one inequality restriction for each restricted observation, so that their conditional distribution is a truncated normal in a dimension of the order of the sample size, for which accept-reject sampling almost always rejects. Second, they are serially dependent through the VAR, so that a Gibbs sampler that draws them one at a time mixes poorly. Building on Chan *et al.* (2023), we show that the precision matrix of this truncated normal distribution is banded, and we sample from it, jointly or equation by equation, with the Zigzag Hamiltonian Monte Carlo (HMC) sampler of Nishimura *et al.* (2020) and Nishimura *et al.* (2025). A timing study shows that the sampler is feasible for VARs with as many as 30 variables, 750 observations and nine restricted variables.

In a Monte Carlo study, we compare the proposed model with a Gaussian VAR fitted directly to the observed outcomes across 24 designs. The proposed model forecasts the observed count variables better in every design, the binary and censored variables in most designs, and the continuous variables equally well. Treating the restricted variables as continuous distorts the intercepts and the innovation covariance matrix.

We then apply the framework to 25 monthly US macro-financial variables: the NBER recession indicator, the number of bank failures, six interest rates that are subject to censoring at the ELB, and 17 unrestricted variables, two of which are unavailable early in the sample. The latent variables are interpretable: a business cycle indicator underlying the NBER recession dates, the intensity of banking distress underlying the bank failure count, and shadow rates that summarize the stance of monetary policy at the ELB. Finally, we use the model for conditional forecasting and scenario analysis in which scenarios can be formulated for the restricted variables as well as for the continuous variables: a target for the recession probability on impact, the bank failure path of the Global Financial Crisis as the expected path, a tightening of financial conditions, the 2026 supervisory stress test scenarios, and oil price paths calibrated to the 2026 Iran war.

Our paper contributes to the literature on VARs with restricted variables. Dueker (2005) develops the qualitative VAR, which adds a single binary variable to a standard VAR and is used for forecasting and scenario analysis in Dueker and Nelson (2006), Dueker and Assenmacher-Wesche (2010) and McCracken *et al.* (2022). Poon and Zhu (2024) model several binary indicators jointly with a multinomial logit model. For censored variables, Johannsen and Mertens (2021) propose an unobserved components model for interest rates at the ELB, and Carriero *et al.* (2025) model interest rates as censored observations of latent variables within a VAR. Count time series have mostly been modeled one series at a time, by integer-valued autoregressions (Al-Osh and Alzaid, 1987) or by latent Gaussian transformations (Jia *et al.*, 2023); see Davis *et al.* (2021) and Fokianos (2024) for reviews. Multivariate count models and copula models for mixed data introduce the dependence across series separately from the univariate margins (Heinen and Rengifo, 2007; Berry and West, 2020; Zito and Kowal, 2025). Relative to this literature, our framework handles multiple variables of each type and missing values within one latent Gaussian VAR, and we develop a posterior sampler that scales to the dimensions of applied work.

The paper is organized as follows. Section 2 introduces the VAR for binary, censored, count and unrestricted variables, as well as the posterior sampler for the latent variables. Section 3 first documents the computational cost of the samplers across 24 designs and then reports the forecast accuracy of the proposed model relative to a Gaussian VAR fitted to the observed outcomes. Section 4 applies the framework to US macro-financial data, reports the estimates of the latent states and conducts the scenario analysis. Section 5 concludes.

2. ECONOMETRIC FRAMEWORK

This section introduces a VAR for a mix of binary, censored, count and unrestricted variables and discusses the normalizations that identify it. It then derives the conditional distribution of the latent variables, outlines the sampler that draws from it, and states the priors.

2.1. VAR for Mixed Variables

To set the stage, let $\widetilde{\boldsymbol{y}}_t = (\widetilde{\boldsymbol{y}}'_{b,t}, \widetilde{\boldsymbol{y}}'_{c,t}, \widetilde{\boldsymbol{y}}'_{z,t}, \boldsymbol{y}'_{u,t})'$, for $t = 1, \ldots, T$, be an $n \times 1$ vector, with $n = n_b + n_c + n_z + n_u$, consisting of:

- n_b binary variables $\widetilde{y}_{b,it} \in \{0, 1\}$ in $\widetilde{\boldsymbol{y}}_{b,t} = (\widetilde{y}_{b,1t}, \ldots, \widetilde{y}_{b,n_bt})'$;
- n_c censored variables $\widetilde{y}_{c,it} \in [l_i, \infty)$ in $\widetilde{\boldsymbol{y}}_{c,t} = (\widetilde{y}_{c,1t}, \ldots, \widetilde{y}_{c,n_ct})'$ with known thresholds l_i (for ease of exposition we take $l_i = 0$; the proposed framework can be easily adapted to cases with non-zero thresholds, or when observations are censored from above);
- n_z count variables with support in the set of non-negative integers, $\widetilde{y}_{z,it} \in \mathbb{Z}_{\geq 0}$, in $\widetilde{\boldsymbol{y}}_{z,t} = (\widetilde{y}_{z,1t}, \ldots, \widetilde{y}_{z,n_zt})'$;
- n_u unrestricted (continuous) variables $y_{u,it} \in \mathbb{R}$ in $\boldsymbol{y}_{u,t} = (y_{u,1t}, \ldots, y_{u,n_ut})'$.

Due to the restricted supports of the subvectors $\widetilde{\boldsymbol{y}}_{b,t}$, $\widetilde{\boldsymbol{y}}_{c,t}$ and $\widetilde{\boldsymbol{y}}_{z,t}$, it is inappropriate to model $\widetilde{\boldsymbol{y}}_t$ using a VAR with errors drawn from a continuous distribution. A standard approach to handling discrete and censored variables is via data augmentation: introduce latent variables $\boldsymbol{y}_{b,t} \in \mathbb{R}^{n_b}$, $\boldsymbol{y}_{c,t} \in \mathbb{R}^{n_c}$ and $\boldsymbol{y}_{z,t} \in \mathbb{R}^{n_z}$ that are linked to, respectively, the binary observations $\widetilde{\boldsymbol{y}}_{b,t}$, the censored observations $\widetilde{\boldsymbol{y}}_{c,t}$, and the count variables $\widetilde{\boldsymbol{y}}_{z,t}$ through

$$\widetilde{y}_{b,it} = \mathbb{I}(y_{b,it} > 0), \quad \text{for } i = 1, \ldots, n_b, \tag{1}$$

$$\widetilde{y}_{c,it} = y_{c,it}\mathbb{I}(y_{c,it} > 0) = \max(y_{c,it}, 0), \quad \text{for } i = 1, \ldots, n_c, \tag{2}$$

$$\widetilde{y}_{z,it} = \sum_{k=1}^{\infty} \mathbb{I}\left(y_{z,it} \geq g_i(k)\right), \quad \text{for } i = 1, \ldots, n_z, \tag{3}$$

where $\mathbb{I}(E)$ is an indicator function that takes the value 1 if the event E is true and 0 otherwise, and g_i is a strictly increasing function on the positive integers with $g_i(1) = 0$. The link functions (1) and (2), respectively, give the standard probit and tobit models. Under (3), the observed count is the number of thresholds $g_i(1) < g_i(2) < \cdots$ that lie below the latent variable: $\widetilde{y}_{z,it} = 0$ if $y_{z,it} < 0$, and $\widetilde{y}_{z,it} = k$ if $g_i(k) \leq y_{z,it} < g_i(k+1)$ for

$k \geq 1$. Equation (3) is therefore an ordered probit model with fixed thresholds, extended to an unbounded support. We use the Box–Cox family, $g_i(k) = (k^{\psi_i} - 1)/\psi_i$, $\psi_i \in [0, 1]$, with $g_i(k) = \log k$ for $\psi_i = 0$. For $\psi_i = 1$ the thresholds are the non-negative integers, each count corresponds to an interval of unit length, and (3) reduces to $\widetilde{y}_{z,it} = \max(\lceil y_{z,it} \rceil, 0)$ except at the integer values of $y_{z,it}$, which occur with probability zero; for $\psi_i = 0$ the count k corresponds to the interval $[\log k, \log(k+1))$ of width $\log(1 + 1/k) \approx 1/k$, and (3) reduces to $\widetilde{y}_{z,it} = \lfloor \exp(y_{z,it}) \rfloor$. For every ψ_i a zero count corresponds to the half-line $y_{z,it} < 0$.

The scale of g_i is a normalization: multiplying g_i and $y_{z,it}$ by the same positive constant leaves the distribution of the counts unchanged, in the same way that the unit error variance is a normalization in a probit model. The spacing of the thresholds, together with the Gaussian latent process introduced below, restricts the conditional distribution of the count, which is a discretized normal: $\mathbb{P}(\widetilde{y}_{z,it} = k) = \Phi((g_i(k+1) - \mu_{it})/\sigma_{it}) - \Phi((g_i(k) - \mu_{it})/\sigma_{it})$ for $k \geq 1$ and $\mathbb{P}(\widetilde{y}_{z,it} = 0) = \Phi(-\mu_{it}/\sigma_{it})$, where μ_{it} and σ_{it}^2 are the conditional mean and variance of $y_{z,it}$ and Φ is the standard normal distribution function. For $\psi_i = 1$, the thresholds are equally spaced. Away from the zero boundary, symmetry and dispersion depend on the latent mean's position relative to the threshold grid. When the latent standard deviation is sufficiently large relative to the unit spacing, these discretization effects are small, which gives an approximately symmetric distribution whose variance is approximately independent of the latent mean, holding the latent variance fixed. For $\psi_i = 0$ the distribution is right-skewed and its standard deviation is proportional to its mean for large counts. For every ψ_i the zero cell collects all the mass below zero. The choice of ψ_i affects the sampler only through the interval bounds of the count observations.

Poisson and negative binomial regressions are the standard tools for a single count series. In the Poisson distribution the conditional variance equals the conditional mean and in the negative binomial distribution it is a quadratic function of the mean, whereas under (3) the location and the scale of the conditional distribution are not tied to each other. Its shape is governed by one fixed parameter rather than by one threshold for each

distinct count, as in an ordered probit model with free thresholds (see also Kowal and Canale, 2020). Multivariate count models specify univariate margins for the counts and introduce the dependence across series separately, through a copula (Heinen and Rengifo, 2007) or through shared latent factors (Berry and West, 2020), as do copula models for mixed data (Zito and Kowal, 2025); in (1) to (3) the margins and the dependence between all variables are instead implied jointly by the Gaussian VAR for the latent vector we discuss next, with a common lag structure and contemporaneous correlations.

Finally, we define $\boldsymbol{y}_t = (\boldsymbol{y}'_{b,t}, \boldsymbol{y}'_{c,t}, \boldsymbol{y}'_{z,t}, \boldsymbol{y}'_{u,t})'$ which collects the latent variables associated with the binary, censored and count variables, as well as the observed continuous variables. This vector can be modeled as a VAR(P) process:

$$\boldsymbol{y}_t = \boldsymbol{A}'\boldsymbol{x}_t + \boldsymbol{u}_t, \quad \boldsymbol{u}_t = \boldsymbol{A}_0^{-1}\boldsymbol{O}_t\boldsymbol{H}_t^{1/2}\boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim \mathcal{N}(\boldsymbol{0}_n, \boldsymbol{I}_n), \tag{4}$$

where $\boldsymbol{x}_t = (\boldsymbol{y}'_{t-1}, \ldots, \boldsymbol{y}'_{t-P}, 1)'$ and $\boldsymbol{A} = (\boldsymbol{A}_1, \ldots, \boldsymbol{A}_P, \boldsymbol{a})'$ stacks the lag-specific $n \times n$ coefficient matrices $\boldsymbol{A}_p$ for $p = 1, \ldots, P$, and $\boldsymbol{a}$ is a vector of intercepts. The regressor vector $\boldsymbol{x}_t$ can also include deterministic terms or exogenous variables, which enter the derivations below through the intercept term; we omit them to simplify the notation. The matrix $\boldsymbol{A}_0$ is unrestricted and invertible. The diagonal matrix $\boldsymbol{H}_t^{1/2} = \text{diag}(\exp(h_{1t}/2), \ldots, \exp(h_{nt}/2))$ contains the standard deviations of the structural shocks, whose log-volatilities h_{it} follow the processes specified in the next section. The outlier component $\boldsymbol{O}_t = \text{diag}(o_{1t}, \ldots, o_{nt})$ consists of scale factors that absorb large idiosyncratic shocks and outliers, as in Carriero *et al.* (2024). Conditional on $\boldsymbol{H}_t$ and $\boldsymbol{O}_t$ the innovations are Gaussian, so that unconditionally they follow a scale mixture of normals. The reduced-form covariance matrix is $\text{Var}(\boldsymbol{u}_t) = \boldsymbol{\Sigma}_t = \boldsymbol{A}_0^{-1}\boldsymbol{O}_t\boldsymbol{H}_t\boldsymbol{O}_t\boldsymbol{A}_0^{-1\prime}$.

To summarize, the proposed framework incorporates multivariate probit, tobit and ordered probit models to construct a VAR, which we call the ProToC-VAR for its probit, tobit and count components. It generalizes a variety of VARs previously considered in the

literature. When $n_b = n_c = n_z = 0$, it reduces to the conventional VAR with continuous variables; when $n_c = n_z = 0$, it becomes a multivariate generalization of the Qual VAR of Dueker (2005) that contains a single binary variable; when $n_b = n_c = 0$, it reduces to a VAR for count and continuous variables; when $n_b = n_z = 0$, it becomes what has been termed a simple shadow rate VAR by Carriero *et al.* (2025) in an application to infer interest rates at the effective lower bound (ELB).

2.2. Normalizations

There are two identification issues in this model. First, since $\boldsymbol{A}_0$ is unrestricted, the scale of the volatilities in $\boldsymbol{H}_t$ is not identified separately from the scale of $\boldsymbol{A}_0$. We follow Chan *et al.* (2024) and fix the unconditional mean of the log-volatilities at zero through the state equations $h_{it} = \phi_i h_{i,t-1} + u_{h,it}$, $u_{h,it} \sim \mathcal{N}(0, \omega_i^2)$, where $|\phi_i| < 1$ and the initial condition is $h_{i0} \sim \mathcal{N}(0, \omega_i^2/(1-\phi_i^2))$. The presence of stochastic volatility identifies $\boldsymbol{A}_0$ up to permutations and sign changes of its rows. All objects reported in this paper, namely the reduced-form covariance matrices $\boldsymbol{\Sigma}_t$, the latent states, the forecasts and the generalized impulse responses, are invariant to this labeling. For diagnostic purposes, the stored draws of $\boldsymbol{A}_0$ are relabeled so that the diagonal elements are positive and the row permutation is closest to the identity.

Second, since the binary observations $\widetilde{\boldsymbol{y}}_{b,t}$ identify only the signs of the corresponding latent variables, the diagonal elements of $\boldsymbol{\Sigma}_t$ that correspond to the binary variables are not identified. In the spirit of Dueker (2005), we normalize each draw of $\boldsymbol{A}_0$ so that the diagonal elements of the time-average innovation covariance matrix that correspond to the binary variables equal one. This matrix excludes the outlier component $\boldsymbol{O}_t$, $\bar{\boldsymbol{\Sigma}} = \sum_{t=1}^{T} \boldsymbol{A}_0^{-1} \boldsymbol{H}_t \boldsymbol{A}_0^{-1\prime}/T,$, so that the unit-scale restriction is a statement about normal times.

In each MCMC iteration, we sample $\boldsymbol{A}_0$, the log-volatilities and the outlier component from their unrestricted conditional posteriors. We then compute, for each binary variable i, the time-average innovation variance: $v_i = [\bar{\boldsymbol{\Sigma}}]_{ii} = \sum_{j=1}^{n} \left(\boldsymbol{A}_0^{-1}\right)_{ij}^2 \sum_{t=1}^{T} \exp(h_{jt})/T$, and form $\boldsymbol{A}_0^* = \boldsymbol{A}_0 \boldsymbol{D}$, where $\boldsymbol{D} = \text{diag}(d_1, \ldots, d_n)$ with $d_i = \sqrt{v_i}$ for the binary variables and

$d_i = 1$ otherwise. Since $(\boldsymbol{A}_0^*)^{-1} = \boldsymbol{D}^{-1}\boldsymbol{A}_0^{-1}$, the rescaled draw satisfies the normalization exactly, while the volatility paths h_{jt} are unaffected. Rescaling $\boldsymbol{A}_0$ in this way is the change that a division of the latent variable $y_{b,it}$ by $\sqrt{v_i} > 0$ would imply for $\boldsymbol{A}_0$. Such a division preserves signs, so the likelihood of the observed binary data is unaffected by the scale that the normalization fixes. Only $\boldsymbol{A}_0$ is rescaled in this step. The latent states and the VAR coefficients keep their current values and are redrawn conditional on the normalized draw in the subsequent steps. The precision matrix $\boldsymbol{\Sigma}_t^{-1} = \boldsymbol{A}_0^{*\prime}\boldsymbol{H}_t^{-1}\boldsymbol{O}_t^{-2}\boldsymbol{A}_0^*$ used for sampling the latent states still includes the outlier component. We replace $\boldsymbol{A}_0$ by $\boldsymbol{A}_0^*$, so that all remaining MCMC steps condition on the normalized draw.

The rescaling is a deterministic step inserted between the conditional draws, as in Dueker (2005). The resulting scheme is a draw-then-rescale device rather than a Gibbs sampler for the normalized model in the strict sense, but the normalized draws are the objects that the likelihood identifies, and the simulation study in Section 3.3 evaluates the parameter estimates on this normalized scale.[2]

2.3. Sampling the Latent States

It is useful to divide $\boldsymbol{y}_t$ into a subvector $\boldsymbol{y}_{\mathtt{l},t}$ of latent elements and a subvector $\boldsymbol{y}_{\mathtt{o},t}$ of observed elements (with subscripts `l` and `o`). Latent elements arise in two ways: from the data augmentation for the binary, censored and count variables, and from missing observations of any variable type, for example at the beginning or the end of the sample because of an unbalanced panel or a release delay. We partition $\boldsymbol{y}_t$ accordingly into three groups: the

[2] In a preliminary analysis, we consider several alternatives. Normalizing the stored draws ex post leaves the unidentified scale free to drift along the chain, which destabilizes the latent-state draws numerically. Fixing the diagonal elements of $\boldsymbol{A}_0$ that correspond to the binary variables through tight priors leaves the marginal variance of the latent index unpinned, which complicates the interpretation of its scale, and causes convergence problems in our experiments. Using $\boldsymbol{A}_0^*$ only in the step that draws the latent states, so that the parameter draws remain unrestricted, imposes the normalization only approximately and performs worse than the exact rescaling of every draw.

missing values $\boldsymbol{y}_{\mathtt{m},t}$ (labeled $\mathtt{m}$, irrespective of variable type); the restricted values $\boldsymbol{y}_{\mathtt{r},t}$ (labeled $\mathtt{r}$), which are the latent variables underlying the observed binary and count variables and the censored observations; and the observed unrestricted values $\boldsymbol{y}_{\mathtt{o},t}$ (labeled $\mathtt{o}$), which are the observed continuous variables and the uncensored observations of the censored variables. Let $n_t^{\mathtt{m}}$, $n_t^{\mathtt{r}}$ and $n_t^{\mathtt{o}}$ denote the numbers of elements in the three groups, and let $\boldsymbol{S}_{\mathtt{m},t}$, $\boldsymbol{S}_{\mathtt{r},t}$ and $\boldsymbol{S}_{\mathtt{o},t}$ be the $n \times n_t^{\mathtt{m}}$, $n \times n_t^{\mathtt{r}}$ and $n \times n_t^{\mathtt{o}}$ selection matrices with $\boldsymbol{S}'_{\mathtt{m},t}\boldsymbol{y}_t = \boldsymbol{y}_{\mathtt{m},t}$, $\boldsymbol{S}'_{\mathtt{r},t}\boldsymbol{y}_t = \boldsymbol{y}_{\mathtt{r},t}$ and $\boldsymbol{S}'_{\mathtt{o},t}\boldsymbol{y}_t = \boldsymbol{y}_{\mathtt{o},t}$. The matrix $\boldsymbol{S}_t = [\boldsymbol{S}_{\mathtt{m},t}, \boldsymbol{S}_{\mathtt{r},t}, \boldsymbol{S}_{\mathtt{o},t}]$ is an $n \times n$ permutation matrix, so that $\boldsymbol{S}_t\boldsymbol{S}'_t = \boldsymbol{I}_n$ and $\boldsymbol{S}'_t\boldsymbol{y}_t = (\boldsymbol{y}'_{\mathtt{m},t}, \boldsymbol{y}'_{\mathtt{r},t}, \boldsymbol{y}'_{\mathtt{o},t})'$. Premultiplying the latter by $\boldsymbol{S}_t$ gives

$$\boldsymbol{y}_t = \boldsymbol{S}_{\mathtt{m},t}\boldsymbol{y}_{\mathtt{m},t} + \boldsymbol{S}_{\mathtt{r},t}\boldsymbol{y}_{\mathtt{r},t} + \boldsymbol{S}_{\mathtt{o},t}\boldsymbol{y}_{\mathtt{o},t} = \boldsymbol{S}_{\mathtt{l},t}\boldsymbol{y}_{\mathtt{l},t} + \boldsymbol{S}_{\mathtt{o},t}\boldsymbol{y}_{\mathtt{o},t}, \tag{5}$$

where $\boldsymbol{S}_{\mathtt{l},t} = [\boldsymbol{S}_{\mathtt{m},t}, \boldsymbol{S}_{\mathtt{r},t}]$ and $\boldsymbol{y}_{\mathtt{l},t} = (\boldsymbol{y}'_{\mathtt{m},t}, \boldsymbol{y}'_{\mathtt{r},t})'$ has $n_t^{\mathtt{l}} = n_t^{\mathtt{m}} + n_t^{\mathtt{r}}$ elements. The total number of latent elements is $N^{\mathtt{l}} = \sum_{t=-P+1}^{T} n_t^{\mathtt{l}}$, where the presample values $t = -P+1, \ldots, 0$ are treated in the same way as the in-sample values.

The next step is to derive the joint density of $\boldsymbol{y} = (\boldsymbol{y}'_{-P+1}, \ldots, \boldsymbol{y}'_0, \boldsymbol{y}'_1, \ldots, \boldsymbol{y}'_T)'$, which is of dimension $(T+P)n \times 1$. Let $\boldsymbol{y}_{\mathtt{l}}$ and $\boldsymbol{y}_{\mathtt{o}}$ denote, respectively, the vectors of all latent variables and observed data. Furthermore, let bdiag$(\cdot)$ represent the operator that constructs a block diagonal matrix from its inputs. We define the selection matrices $\boldsymbol{S}_{\mathtt{l}} = \text{bdiag}(\boldsymbol{S}_{\mathtt{l},(-P+1)}, \ldots, \boldsymbol{S}_{\mathtt{l},T})$ and $\boldsymbol{S}_{\mathtt{o}} = \text{bdiag}(\boldsymbol{S}_{\mathtt{o},(-P+1)}, \ldots, \boldsymbol{S}_{\mathtt{o},T})$, so that $\boldsymbol{S}'_{\mathtt{l}}\boldsymbol{y} = \boldsymbol{y}_{\mathtt{l}}$ and $\boldsymbol{S}'_{\mathtt{o}}\boldsymbol{y} = \boldsymbol{y}_{\mathtt{o}}$. Stacking (5) over $t = -P+1, \ldots, T$ gives

$$\boldsymbol{y} = \boldsymbol{S}_{\mathtt{l}}\boldsymbol{y}_{\mathtt{l}} + \boldsymbol{S}_{\mathtt{o}}\boldsymbol{y}_{\mathtt{o}}. \tag{6}$$

Following Chan *et al.* (2023), we encode the VAR dynamics in the $Tn \times (T+P)n$ matrix

$$\widetilde{\boldsymbol{M}} = \begin{bmatrix} -\boldsymbol{A}_P & \cdots & -\boldsymbol{A}_1 & \boldsymbol{I}_n & \boldsymbol{0}_{n\times n} & \cdots & \cdots & \boldsymbol{0}_{n\times n} \\ \boldsymbol{0}_{n\times n} & -\boldsymbol{A}_P & \cdots & -\boldsymbol{A}_1 & \boldsymbol{I}_n & \boldsymbol{0}_{n\times n} & \cdots & \boldsymbol{0}_{n\times n} \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \ddots & \vdots \\ \boldsymbol{0}_{n\times n} & \cdots & \boldsymbol{0}_{n\times n} & -\boldsymbol{A}_P & \cdots & -\boldsymbol{A}_1 & \boldsymbol{I}_n & \boldsymbol{0}_{n\times n} \\ \boldsymbol{0}_{n\times n} & \cdots & \cdots & \boldsymbol{0}_{n\times n} & -\boldsymbol{A}_P & \cdots & -\boldsymbol{A}_1 & \boldsymbol{I}_n \end{bmatrix}.$$

Moreover, $\widetilde{\boldsymbol{\Omega}} = \text{bdiag}(\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_T)$ is the $Tn \times Tn$ error covariance matrix, $\widetilde{\boldsymbol{m}} = (\boldsymbol{\iota}_T \otimes \boldsymbol{a})$ is the Tn-vector of intercepts, $\boldsymbol{\iota}_T$ is a T-vector of ones, and $\otimes$ is the Kronecker product. The presample vector $(\boldsymbol{y}'_{-P+1}, \ldots, \boldsymbol{y}'_0)'$ has prior mean $\boldsymbol{m}_0 = \boldsymbol{0}_{nP}$ and prior covariance $\boldsymbol{\Omega}_0 = \boldsymbol{I}_{nP}$. We accordingly define $\boldsymbol{M}_0 = [\boldsymbol{I}_{nP}, \boldsymbol{0}_{nP\times Tn}]$ to obtain $\boldsymbol{M} = [\boldsymbol{M}'_0, \widetilde{\boldsymbol{M}}']'$, $\boldsymbol{\Omega} = \text{bdiag}(\boldsymbol{\Omega}_0, \widetilde{\boldsymbol{\Omega}})$, $\boldsymbol{m} = (\boldsymbol{m}'_0, \widetilde{\boldsymbol{m}}')'$, so that (4) and the prior on the initial conditions can be written as

$$\boldsymbol{M}\boldsymbol{y} = \boldsymbol{m} + \boldsymbol{e}, \quad \boldsymbol{e} \sim \mathcal{N}(\boldsymbol{0}_{(T+P)n}, \boldsymbol{\Omega}). \tag{7}$$

Substituting (6) into (7), we have $\boldsymbol{M}\boldsymbol{S}_1\boldsymbol{y}_1 + \boldsymbol{M}\boldsymbol{S}_\circ\boldsymbol{y}_\circ = \boldsymbol{m} + \boldsymbol{e}$. We now derive the distribution of the latent vector $\boldsymbol{y}_1$ conditional on the observed vector $\boldsymbol{y}_\circ$ and on the parameters. Defining $\boldsymbol{G}_1 = \boldsymbol{M}\boldsymbol{S}_1$ and $\boldsymbol{G}_\circ = \boldsymbol{M}\boldsymbol{S}_\circ$, the exponent of the joint density implied by (7) is $-\frac{1}{2}(\boldsymbol{G}_1\boldsymbol{y}_1 + \boldsymbol{G}_\circ\boldsymbol{y}_\circ - \boldsymbol{m})'\boldsymbol{\Omega}^{-1}(\boldsymbol{G}_1\boldsymbol{y}_1 + \boldsymbol{G}_\circ\boldsymbol{y}_\circ - \boldsymbol{m})$, which is quadratic in $\boldsymbol{y}_1$. Completing the square, as in Chan *et al.* (2023), gives $\boldsymbol{y}_1 \mid (\boldsymbol{y}_\circ, \bullet) \sim \mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{V}_1)$ with

$$\boldsymbol{V}_1 = (\boldsymbol{G}'_1\boldsymbol{\Omega}^{-1}\boldsymbol{G}_1)^{-1}, \quad \boldsymbol{\mu}_1 = \boldsymbol{V}_1\boldsymbol{G}'_1\boldsymbol{\Omega}^{-1}(\boldsymbol{m} - \boldsymbol{G}_\circ\boldsymbol{y}_\circ). \tag{8}$$

Since $\boldsymbol{M}$ is block banded with bandwidth P and $\boldsymbol{\Omega}$ is block diagonal, the precision matrix $\boldsymbol{V}_1^{-1} = \boldsymbol{G}'_1\boldsymbol{\Omega}^{-1}\boldsymbol{G}_1$ is banded when the latent elements are ordered chronologically. All computations below exploit this structure through sparse and band matrix routines.

Conditioning in addition on the observed binary, censored and count variables $\widetilde{\boldsymbol{y}}_b$, $\widetilde{\boldsymbol{y}}_c$ and $\widetilde{\boldsymbol{y}}_z$, we obtain $p(\boldsymbol{y}_1 \mid \widetilde{\boldsymbol{y}}_b, \widetilde{\boldsymbol{y}}_c, \widetilde{\boldsymbol{y}}_z, \boldsymbol{y}_\circ, \bullet)$, which is a truncated multivariate normal distribution:

$$\boldsymbol{y}_1 \mid (\widetilde{\boldsymbol{y}}_b, \widetilde{\boldsymbol{y}}_c, \widetilde{\boldsymbol{y}}_z, \boldsymbol{y}_\circ, \bullet) \sim \mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{V}_1), \quad \text{subject to} \quad \boldsymbol{l} \leq \boldsymbol{y}_1 \leq \boldsymbol{u}. \tag{9}$$

The lower bound $\boldsymbol{l} = (l_1, \ldots, l_{N^1})'$ and upper bound $\boldsymbol{u} = (u_1, \ldots, u_{N^1})'$ are determined by the measurements in (1) to (3). In case an observation is missing, irrespective of the variable type, we have $(l_j, u_j) = (-\infty, \infty)$. The remaining cases are as follows:

- Binary variables: $\widetilde{y}_{b,it} = 1$ gives $(l_j, u_j) = (0, \infty)$ and $\widetilde{y}_{b,it} = 0$ gives $(l_j, u_j) = (-\infty, 0)$;
- Censored variables: $\widetilde{y}_{c,it} = 0$ gives $(l_j, u_j) = (-\infty, 0)$;
- Count variables: $\widetilde{y}_{z,it} = 0$ gives $(l_j, u_j) = (-\infty, 0)$ and $\widetilde{y}_{z,it} = k > 0$ gives $(l_j, u_j) = (g_i(k), g_i(k+1))$, which are the unit intervals $(k-1, k)$ for $\psi_i = 1$ and the intervals $(\log k, \log(k+1))$ for $\psi_i = 0$.

Blocked updates. Drawing all latent elements jointly from (9) is possible but slow when n and T are large. Alternatively, the latent elements can be updated equation by equation, conditional on the current values of all other elements of $\boldsymbol{y}$. For variable i, let $\boldsymbol{S}_1^{(i)}$ be the selection matrix that picks the latent elements of variable i out of $\boldsymbol{y}$ and $\boldsymbol{S}_\circ^{(i)}$ the selection matrix that picks all remaining elements, latent or observed, so that $\boldsymbol{y} = \boldsymbol{S}_1^{(i)}\boldsymbol{y}_1^{(i)} + \boldsymbol{S}_\circ^{(i)}\boldsymbol{y}_\circ^{(i)}$. The conditional distribution of $\boldsymbol{y}_1^{(i)}$ follows from (8) and (9) with $\boldsymbol{S}_1$ and $\boldsymbol{S}_\circ$ replaced by $\boldsymbol{S}_1^{(i)}$ and $\boldsymbol{S}_\circ^{(i)}$. Its precision matrix has dimension at most $T + P$ and bandwidth P, so that the operations on it are linear in T, and one sweep of the sampler cycles through the blocks $i = 1, \ldots, n$. Blocks of several variables can be formed in the same way.

Sampling from the truncated normal distribution. Draws from (9), jointly or by blocks, can be generated by Gibbs sampling one element at a time, by the minimax ex-

ponential tilting (MET) method of Botev (2017), or by Hamiltonian Monte Carlo (HMC). Gibbs sampling mixes slowly when the elements are strongly correlated, as consecutive values of a latent variable are here. MET produces independent draws by accept-reject sampling, but its acceptance rate falls with the dimension of the constrained vector, which here is of order T for a single variable, and it is infeasible for the larger designs considered in later sections. We therefore use HMC.

HMC borrows an idea from physics (Bishop, 2006; Neal, 2011). It treats the current draw as the position of a particle and the negative log density as its potential energy, a landscape whose valleys are the regions of high probability. Each step gives the particle a random momentum, and hence kinetic energy, and moves it under Hamiltonian dynamics, which conserve the total energy. The particle rolls toward the valleys, exchanging potential for kinetic energy on the way, and its position after a fixed travel time is the new draw. This draw can be far from the previous one, so that consecutive draws are much less dependent than in a Gibbs sampler. Because the total energy is fixed, the particle can climb only as far as its kinetic energy allows, so it stays in the regions of high probability while it travels.

Pakman and Paninski (2014) show that for a truncated normal target the Hamiltonian dynamics can be solved exactly: the potential energy is the quadratic form $\frac{1}{2}(\boldsymbol{y}_1 - \boldsymbol{\mu}_1)'\boldsymbol{V}_1^{-1}(\boldsymbol{y}_1 - \boldsymbol{\mu}_1)$, the truncation bounds act as reflecting walls, and a trajectory is followed for a fixed travel time and reflected whenever a coordinate reaches its bound. We use the Zigzag variant of Nishimura *et al.* (2020) and Nishimura *et al.* (2025), in which the momenta have Laplace instead of Gaussian distributions, so that the trajectories are piecewise linear and the reflection times are available in closed form; each step then requires only products of the banded precision matrix with a vector, and, because the dynamics are solved exactly, the accept-reject step that corrects the discretization error of HMC in general is unnecessary. Zhang *et al.* (2026) compare these samplers, and we use their implementation.

The only tuning parameter is the travel time, which we set during the burn-in period from the smallest eigenvalue of the precision matrix of the block and keep fixed afterwards.[3] HMC trades some autocorrelation of the draws for a large reduction in computing time. In Section 3 we compare the computational cost of the samplers and report the effective sample sizes of the latent states.

2.4. Priors

We assume a global-local shrinkage prior on the VAR lag coefficients. Collecting these coefficients in $\boldsymbol{\beta}$, we specify the horseshoe prior (Carvalho *et al.*, 2010):

$$\beta_j \mid \tau, \lambda_j \sim \mathcal{N}(0, \tau^2\lambda_j^2), \qquad \tau \sim \mathcal{C}^+(0,1), \qquad \lambda_j \sim \mathcal{C}^+(0,1), \qquad j = 1, \dots, n^2P.$$

The common global scale shrinks coefficients towards zero, while coefficient-specific local scales allow important coefficients to escape strong shrinkage. The intercepts are excluded from hierarchical shrinkage and assigned independent $\mathcal{N}(0,1)$ priors.

For the stochastic volatility specification, we assign independent standard normal priors to the elements of $\boldsymbol{A}_0$. The unconditional means of the log-volatilities are fixed at zero, leaving the overall scale to $\boldsymbol{A}_0$. For the persistence and innovation variance parameters, we specify $(\phi_i + 1)/2 \sim \mathcal{B}(5, 1.5)$ and $\omega_i^2 \sim \mathcal{G}(1/2, 5)$, where the Gamma distribution uses the shape-rate parameterization; the latter is equivalent to $\omega_i \sim \mathcal{N}(0, 0.1)$. Initial log-volatilities follow their stationary distributions.

When the outlier component is included, each structural shock has a separate scale factor o_{it}, following Carriero *et al.* (2024). Conditional on its outlier probability $\mathfrak{p}_i$, this factor equals one with probability $1-\mathfrak{p}_i$ and otherwise follows a discrete uniform distribution

[3] The travel time is seeded at $\sqrt{2}/\sqrt{\lambda_{\min}}$, where $\lambda_{\min}$ is the smallest eigenvalue, computed by the Lanczos method. During the burn-in period we refresh the eigenvalue estimate every ten sweeps with two steps of a warm-started inverse power iteration and the associated Rayleigh quotient, clamp the implied travel time to the interval $[0.01, 10]$, and use the average of the last five travel times.

over nine equally spaced points between 2 and 20. We specify $\mathfrak{p}_i \sim \mathcal{B}(2.5, 2.5(T-1))$, implying approximately one outlier per structural shock over the estimation sample a priori. The outlier factors are estimated, so the data decide which months receive a larger shock variance and by how much. A large factor reduces the influence of that month on the VAR coefficients and on the persistent volatility path, while the observations of the month remain in the likelihood and continue to enter the conditional means of later months through the lags.

3. SIMULATION EVIDENCE

In this section we first describe the Monte Carlo designs, which are common to the two exercises that follow. We then compare the computational cost of the posterior samplers for the latent variables across the designs. Finally, we compare the proposed model with a Gaussian Bayesian VAR (BVAR) fitted directly to the observed data, in terms of the accuracy of one-step-ahead forecasts and of the distortion of the parameter estimates that results from treating the restricted variables as continuous.

3.1. Simulation Design

We generate data from a latent VAR(4) with stochastic volatility. Lag coefficients are drawn at random, with large own-lag coefficients and small cross-variable coefficients at the first lag and sparse higher lags whose scale decays with the lag, and are rescaled so that the companion matrix has a spectral radius drawn uniformly from $[0.85, 0.98]$. Innovations have covariance $\boldsymbol{A}_0^{-1}\boldsymbol{H}_t\boldsymbol{A}_0^{-1\prime}$, where $\boldsymbol{A}_0$ is generated at random and the log-volatilities h_{it} follow independent Gaussian AR(1) processes with mean zero, persistence 0.95 and innovation variance 0.05. The system is then standardized so that, at the mean log-volatility, each latent variable has unit unconditional variance. The intercepts are set so that the unconditional means of the latent variables equal $\Phi^{-1}(0.3)$, 1, $\Phi^{-1}(0.7)$ and zero for the binary, count, censored and

continuous variables, respectively. The binary, count and censored observations are obtained from the latent variables through the measurement equations (1) to (3), with $\psi_i = 1$ for the count variables, which are therefore generated and estimated with unit intervals; the continuous variables are observed directly.

We consider sample sizes $T \in \{100, 250, 750\}$, system dimensions $n \in \{10, 20, 30\}$, and $r \in \{1, 2, 3\}$ variables of each of the binary, count and censored types, retaining designs with at least two unrestricted continuous variables, which excludes $(n, r) = (10, 3)$ and leaves 24 designs. After discarding 500 initial observations, we retain T observations for estimation and one for forecast evaluation. Datasets are redrawn until each binary series has a proportion of ones strictly between 0.1 and 0.5, each count series contains at least five distinct values, and each censored series has more than 10% zeros and more than 50% positive observations. These checks use all $T + 1$ retained observations, including the holdout; the results therefore describe the simulation population selected by these restrictions.

The designs are used in two exercises. The timing exercise in Section 3.2 times single runs of the proposed model, with 1,000 burn-in and 1,000 retained draws, on one dataset per design and volatility specification; the data for the homoskedastic specification are generated with constant log-volatilities. The Monte Carlo study in Section 3.3 uses 500 replications per design. On each dataset it estimates the proposed model, with stochastic volatility but without the outlier component and with the equation-by-equation HMC update, and a Gaussian BVAR with the same volatility specification fitted directly to the observed outcomes, with 9,000 burn-in draws and 2,000 retained draws after thinning by three, and it compares the two models on the replications that both completed. Both models have $P = 4$ lags.

3.2. Computational Cost

Table 1 reports the computation time per 100 iterations of the full MCMC sweep, including initialization, for each design and for four variants of the latent-state update: minimax exponential tilting (MET) and Zigzag HMC, each with joint and equation-by-equation updates,

under the stochastic volatility specification and the homoskedastic specification defined in Appendix B. Each run has a budget of one hour for its 2,000 iterations, and blank cells mark runs that did not complete within the budget.

Table 1: Computation time per 100 MCMC iterations, in seconds, for the MET and Zigzag HMC samplers with joint and equation-by-equation updates across the simulation designs.

			Homoskedastic				Stochastic volatility			
			MET		HMC		MET		HMC	
T	n	r	joint	eq.	joint	eq.	joint	eq.	joint	eq.
100	10	1	14.8	4.28	**1.01**	1.16	13.0	5.41	**1.65**	1.86
		2		7.92	1.78	**1.69**		8.94	3.00	**2.48**
	20	1	9.43	5.13	**2.38**	2.66	9.36	6.33	**3.76**	4.11
		2	108	8.34	**3.16**	3.50		11.9	**4.49**	4.79
		3		11.4	4.18	**4.15**		15.5	6.72	**5.72**
	30	1	10.8	8.18	**5.24**	5.78	14.0	10.6	**7.40**	7.97
		2	39.3	11.6	**6.27**	6.90	42.6	13.9	**8.50**	9.27
		3		15.2	**7.79**	7.95		17.3	**10.1**	10.5
250	10	1		109	2.46	**1.82**		26.7	3.62	**2.95**
		2		66.6	8.98	**2.87**			23.8	**4.37**
	20	1	89.7	20.3	3.91	**3.87**	67.8	19.1	5.77	**5.74**
		2			8.39	**5.70**			10.1	**7.66**
		3			20.9	**7.57**		161	29.5	**9.31**
	30	1	107	25.7	7.96	**7.88**	106	31.0	**10.7**	10.8
		2		38.8	12.5	**10.7**		42.4	15.1	**13.2**
		3		52.6	19.7	**13.4**		59.4	24.3	**16.4**
750	10	1			14.7	**5.98**			17.6	**8.30**
		2			95.8	**11.2**			154	**15.3**
	20	1			15.1	**9.75**			18.0	**12.9**
		2			53.6	**17.5**			62.3	**22.2**
		3				**25.1**			164	**27.8**
	30	1			22.2	**17.7**			26.0	**22.4**
		2			53.7	**27.0**			62.3	**32.6**
		3			132	**35.8**			135	**42.4**

Notes: T: observations; n: variables; r: variables of each binary, count, and censored type. Each design includes at least two continuous variables. MET: minimax exponential tilting; HMC: Hamiltonian Monte Carlo; joint and eq.: joint and equation-by-equation updates. Bold marks the fastest method for each design and volatility specification. Entries are seconds per 100 iterations: the total time of each run, including initialization, divided by 20, since each run has 2,000 iterations (1,000 burn-in and 1,000 retained). Blank cells indicate that the run did not complete within one hour, so no comparable average is reported.

We make three observations. First, HMC is the fastest method in every design, and the equation-by-equation HMC update completes every design; in the largest design, with $T = 750$, $n = 30$ and nine restricted variables, it requires about 42 seconds per 100 iterations

under stochastic volatility. Second, MET does not scale: the joint MET update fails to complete within the budget in most designs with $T = 250$, and no MET variant completes any design with $T = 750$. Third, the joint HMC update is marginally faster in most of the small designs with $T = 100$, whereas the equation-by-equation update is faster in nearly every design with $T = 250$ and in every design with $T = 750$, and its advantage grows with T and with the number of restricted variables. Each of its blocks has dimension at most $T + P$ and bandwidth P, whereas both the dimension and the bandwidth of the joint update grow with the number of restricted variables.

We note that the reported times are not times per effective draw: MET produces independent draws from the conditional distribution of the latent states, whereas the HMC draws are autocorrelated, with inefficiency factors between 1.4 and 6.1 in the Monte Carlo study below (Table 2). In the designs with $T = 100$, where MET completes, the comparison per effective draw is closer and in some designs favors MET; in the larger designs MET does not complete within the budget, so HMC is the only feasible option.

3.3. Estimation and Forecast Accuracy

Binary observations identify the sign of the latent variable but leave its scale unrestricted. For evaluation, we apply the normalization of Section 2.2 to the data generating process (DGP) and to the posterior draws of both models, so that each binary latent equation has unit time-average innovation variance $v_i = [\bar{\boldsymbol{\Sigma}}]_{ii}$ over the estimation sample, excluding the initial four lags and the holdout. With $\boldsymbol{D} = \mathrm{diag}(d_1, \ldots, d_n)$, where $d_i = \sqrt{v_i}$ for the binary variables and $d_i = 1$ otherwise, the lag matrices, the intercepts and the covariance matrices are transformed as $\boldsymbol{D}^{-1}\boldsymbol{A}_p\boldsymbol{D}$, $\boldsymbol{D}^{-1}\boldsymbol{a}$ and $\boldsymbol{D}^{-1}\boldsymbol{\Sigma}_t\boldsymbol{D}^{-1}$, and the latent forecasts are premultiplied by $\boldsymbol{D}^{-1}$. This is the full transformation implied by dividing the binary latent variables by $\sqrt{v_i}$. The draws of the proposed model satisfy the normalization by construction, so that $\boldsymbol{D} = \boldsymbol{I}_n$ and they are unchanged; the transformation therefore matters for the DGP and for the draws of the BVAR, and it ensures that the losses of both models are free of the arbitrary

scale of the binary latent variables. This convention fixes the average innovation variances; the innovation variances at individual dates and the unconditional variances of the latent levels remain free. The fixed count thresholds and the observed continuous values anchor the remaining scales. Since the Gaussian BVAR is fitted directly to the observed outcomes, it is misspecified, and its parameter losses measure how well it recovers the parameters of the latent data generating process.

Within each replication, parameter losses are the mean absolute deviations of the elementwise posterior medians from the normalized truth, computed separately over all lag coefficients, intercepts, and unique elements of the time-average innovation covariance matrix. One-step-ahead forecasts are evaluated using the continuous ranked probability score (CRPS; Gneiting and Raftery, 2007), with the Brier score for the observed binary outcomes. For the BVAR, the predictive draws are mapped to the support of each outcome: they are clipped to $[0, 1]$ for the binary variables, rounded upward and bounded below by zero for the count variables, and censored at zero for the censored variables. Each forecast score is divided by the corresponding variable's training-sample standard deviation on the latent or observed scale, then averaged across variables within the reported group.

Table 2 reports ratios of median losses, ProToC-VAR relative to the BVAR, taking medians separately for each model and metric across replications. In panel A, and in the latent-scale forecast row of panel B, the losses of both models are measured against the latent data generating process, so the ratios quantify the distortion that results from treating the restricted variables as continuous; in the remaining rows of panel B, both models forecast the same observed outcomes, and ratios below one favor ProToC-VAR. The table also reports the median latent-state effective sample size (ESS) and inefficiency factor (IF) across ProToC-VAR replications. Within each replication, ESS is summarized by its median across latent states.

We note three patterns in Table 2. First, treating the restricted variables as continuous distorts the intercepts and the innovation covariance matrix in every design: the loss ratios

Table 2: Median simulation losses relative to the Gaussian BVAR and sampling diagnostics.

		$n = 10$		$n = 20$			$n = 30$		
T		$r = 1$	$r = 2$	$r = 1$	$r = 2$	$r = 3$	$r = 1$	$r = 2$	$r = 3$
A. Parameter estimation: relative MAE									
100	Dynamics	1.007	1.000	0.992	0.991	0.991	0.990	0.986	0.978
	Intercepts	0.546	0.492	0.652	0.546	0.523	0.730	0.606	0.567
	Covariance	0.857	0.782	0.942	0.911	0.905	0.980	0.958	0.950
250	Dynamics	0.964	0.943	0.983	0.978	0.986	0.998	0.988	0.989
	Intercepts	0.420	0.392	0.557	0.480	0.469	0.619	0.539	0.515
	Covariance	0.688	0.525	0.822	0.728	0.675	0.895	0.820	0.763
750	Dynamics	0.894	0.849	0.955	0.930	0.924	0.980	0.962	0.958
	Intercepts	0.306	0.287	0.432	0.398	0.388	0.526	0.461	0.456
	Covariance	0.493	0.331	0.645	0.485	0.397	0.740	0.595	0.504
B. Forecast evaluation: relative scores									
100	Latent: all	0.855	0.762	0.936	0.888	0.846	0.960	0.932	0.904
	Binary	0.845	0.915	0.625	0.860	0.892	0.628	0.852	0.906
	Count	0.813	0.793	0.867	0.782	0.822	0.820	0.768	0.810
	Censored	0.923	0.921	0.968	0.987	0.956	0.967	0.931	0.954
	Continuous	1.007	0.986	1.001	0.994	0.990	0.999	1.001	0.991
	Observed: all	0.958	0.907	0.986	0.977	0.964	0.989	0.984	0.971
250	Latent: all	0.827	0.730	0.915	0.865	0.837	0.940	0.909	0.871
	Binary	1.070	1.093	0.681	0.906	1.010	0.726	0.831	0.907
	Count	0.714	0.793	0.816	0.873	0.870	0.790	0.788	0.798
	Censored	0.946	0.973	0.981	0.927	0.997	0.989	0.926	0.972
	Continuous	1.011	0.991	0.995	0.994	1.000	0.995	0.991	1.005
	Observed: all	0.954	0.913	0.978	0.939	0.959	0.979	0.975	0.969
750	Latent: all	0.833	0.649	0.890	0.820	0.765	0.939	0.868	0.818
	Binary	0.830	1.100	0.785	0.862	0.949	0.727	0.820	0.903
	Count	0.917	0.729	0.790	0.730	0.746	0.730	0.765	0.733
	Censored	0.970	0.937	1.050	0.933	0.954	0.998	0.943	0.959
	Continuous	1.000	0.966	1.005	1.010	0.989	1.002	0.994	0.987
	Observed: all	0.955	0.899	0.980	0.935	0.926	0.989	0.976	0.936
ProToC-VAR sampling diagnostics									
100	ESS (IF)	1052 (1.9)	1019 (2.0)	588 (3.4)	582 (3.4)	558 (3.6)	338 (5.9)	337 (5.9)	327 (6.1)
250	ESS (IF)	1173 (1.7)	1095 (1.8)	775 (2.6)	730 (2.7)	727 (2.8)	505 (4.0)	485 (4.1)	499 (4.0)
750	ESS (IF)	1412 (1.4)	1258 (1.6)	1164 (1.7)	973 (2.1)	901 (2.2)	872 (2.3)	733 (2.7)	668 (3.0)

Notes: T: observations; n: variables; r: variables of each binary, count, and censored type. Each design includes at least two continuous variables. Entries in panels A and B are ratios of median losses across 500 replications, ProToC-VAR relative to the Gaussian BVAR fitted to the observed outcomes. In panel A and in the row Latent: all, the losses are measured against the latent data generating process, so the ratios quantify the distortion from treating the restricted variables as continuous; in the remaining rows of panel B, both models forecast the same observed outcomes and values below one favor ProToC-VAR. MAE: mean absolute error. Forecast rows refer to observed outcomes except for Latent: all. Forecast scores are divided by training-sample standard deviations: Brier scores for binary outcomes, continuous ranked probability scores otherwise. For each ProToC-VAR replication, ESS is the median latent-state effective sample size and IF (inefficiency factor) is 2000/ESS, based on 2000 saved draws. The table reports medians of both diagnostics across replications.

lie between 0.29 and 0.73 for the intercepts and between 0.33 and 0.98 for the covariance matrix, and they fall as T grows, because the estimation error of the proposed model shrinks while the distortion of the BVAR does not. Second, the lag coefficients are much less affected: their loss ratios are close to one at $T = 100$ and lie between 0.85 and 0.98 at $T = 750$. Third, in forecasting the observed outcomes, the proposed model improves on the BVAR for the count variables in every design and for the binary and censored variables in most designs, while the two models are on par for the continuous variables. Averaged over all observed variables (the row Observed: all), the proposed model is ahead in every design. Overall, these results show that the shortcut of treating restricted variables as continuous distorts the levels and the innovation scales of the latent system, and that modeling the observation mechanism improves the forecasts of the restricted outcomes at no cost for the continuous variables.

4. MODELING US MACRO-FINANCIAL DYNAMICS

We now apply the proposed model to 25 monthly US macro-financial variables, among them the NBER recession indicator, the number of bank failures and six interest rates that are subject to censoring at the effective lower bound. After describing the data, we examine the estimates of the latent variables underlying the restricted series, the imputed early observations of two variables and the time-varying volatilities. We then use the model for conditional forecasting and scenario analysis with conditioning assumptions on restricted and continuous variables alike.

4.1. Data

Our monthly dataset runs from 1960:01 to 2025:12 and is sourced from FRED-MD (McCracken and Ng, 2016) for most macroeconomic and financial variables. The binary variable `NBERREC` is the indicator of expansions and recessions (contractions) released by the NBER

Business Cycle Dating Committee. The count variable BANKFAIL is the number of bank failures, aggregated from the individual bank-level data on banks.data.fdic.gov. The censored variables are six interest rates at maturities from overnight to ten years, some of which have been at their ELB during the sample. Following Carriero *et al.* (2025), we treat all observations at or below 0.25 percent as constrained by the ELB and record them at the threshold, so that FEDFUNDS, TB3MS, TB6MS and GS1 each have at least one censored observation.

Our information set includes the key variables assessed by the NBER Business Cycle Dating Committee for dating recessions—non-farm payroll employment (PAYEMS), industrial production (INDPRO), real manufacturing and trade sales (CMRMTSPLx), real personal income excluding transfers (W875RX1)—and adds the federal funds rate (FEDFUNDS) and a term spread (TS10G1) as important financial predictors. This information set also mirrors and extends the one of McCracken *et al.* (2022).

The remaining unrestricted variables cover the labor market (the unemployment rate, and average weekly hours and hourly earnings in goods-producing industries), prices (the PCE and CPI price indices and the crude oil price), housing starts, and financial markets (the S&P 500 index, the US–UK exchange rate, Moody's Baa corporate bond yield, the excess bond premium of Gilchrist and Zakrajšek (2012) and the Chicago Fed national financial conditions index). The last two are available from 1973:01 and 1971:01 onwards, and their earlier values are treated as missing.

Employment, industrial production, sales, real income, hourly earnings, the price indices, the stock price index, the exchange rate and the oil price enter as monthly log-differences in percent; weekly hours and housing starts enter in logs; and the remaining variables enter in levels. The transformations follow McCracken *et al.* (2022) for the real activity and price variables and Carriero *et al.* (2025) for the interest rates, so that the information set is comparable with theirs. One could instead use the variables in levels, with the same shrinkage prior, since the measurement equations do not depend on how the continuous variables are transformed. We keep the growth-rate specification because the objects

reported below, the latent states and the conditional forecasts over a three-year horizon, depend on the short-run dynamics of the system. Appendix B lists all variables with their types and transformations.

4.2. Estimates of the Latent States

We estimate the model with $P = 12$ lags, stochastic volatility and the outlier component in all equations. The count variable uses the log thresholds ($\psi_i = 0$), the interest rates are censored at 0.25 percent, and we apply the normalizations of Section 2.2. Before estimation, the continuous and censored series are standardized using outlier-robust measures of location and scale, and the count series and its thresholds are divided by the standard deviation of $\log(k + 1/2)$ over the months with positive counts k, which leaves the zero threshold at zero; all estimates are transformed back to the original units. The latent states are sampled with the equation-by-equation Zigzag HMC update. We run the sampler for 65,000 iterations, discard the first 15,000 as burn-in, and retain every 10th of the remaining draws, which gives 5,000 posterior draws.

A key output of the proposed VAR is the set of latent states associated with the restricted variables. The state underlying the NBER recession indicator can be interpreted as an indicator of the business cycle and its turning points, the state underlying the bank failure count as the intensity of banking distress, and the states associated with the interest rates, or shadow rates, as a measure of the overall stance of monetary policy when the ELB is binding. The model also imputes the missing early observations of the two unrestricted variables that start after 1960, and it delivers the time-varying volatilities and the outlier component of all innovations.

Figure 1 presents the dynamic evolution of the business cycle indicator and the corresponding recession probabilities. The recession probability at date t is $\Phi(\mu_{b,t}/\sigma_{b,t})$, where $\mu_{b,t}$ is the conditional mean of the business cycle indicator given the lagged variables and $\sigma_{b,t}$ its innovation standard deviation at date t, including the outlier component; the figure

shows posterior quantiles of this probability across the draws. Our implementation gives a counter-cyclical indicator, which is negative throughout expansions and rises above zero in every NBER recession; a pro-cyclical indicator can be obtained by redefining the binary variable. As discussed in Dueker and Assenmacher-Wesche (2010), the distance of the indicator from zero measures how far the business cycle is from its next turning point. By this metric, the most severe recessions in the sample before 2020 are the 1973–75 recession and the Great Recession of 2007–09. The implied recession probabilities switch between values close to zero and one within a few months of the NBER turning points, and the credible sets are narrow relative to the movements of the indicator. The COVID-19 recession produces a spike in the indicator about six times as high as the earlier peaks, which the outlier component absorbs (see Figure 5 below); the right panels therefore show this period on its own scale. After the 2020 recession the indicator falls to the lowest value of the sample, which reflects the record growth rates of employment, industrial production, sales and income during the recovery, and it then rises gradually, with brief and uncertain increases in the recession probability in 2024 and 2025.

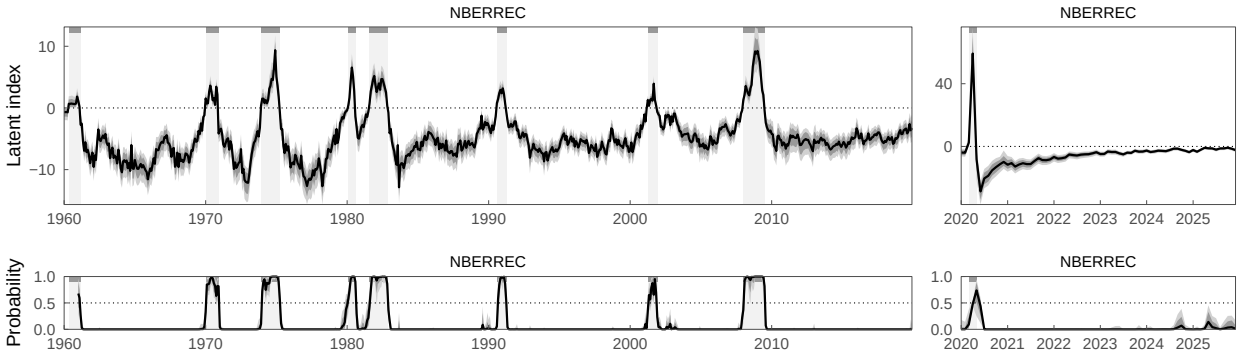

Figure 1: Estimates of the business cycle indicator underlying `NBERREC` (upper panels) and the implied recession probabilities (lower panels).

Notes: Posterior median alongside 68/90 percent credible sets. The indicator is counter-cyclical and rises above zero in recessions. Shaded areas indicate NBER recessions. The left panels cover 1960:01 to 2019:12, the right panels the period from 2020:01 onwards; the samples are plotted separately so that the COVID-19 observations do not compress the scale of the business cycle indicator for the remainder of the sample.

Figure 2 shows the intensity of banking distress underlying the bank failure count on the log scale, on which values below zero correspond to months without failures. Two episodes stand out: the savings and loan crisis, during which the intensity rises from 1982 to a peak in 1989 and returns to zero by 1994, and the Global Financial Crisis, with a rise from 2008 to a peak in 2009 and a return to zero by 2015. In calm periods the intensity fluctuates around zero and its lower credible band is wide, because a month without failures only bounds it from above. The COVID-19 months push the intensity far below zero, another outlier episode, after which it returns to values around zero, with brief excursions above zero at the bank failures of 2023 and 2024.

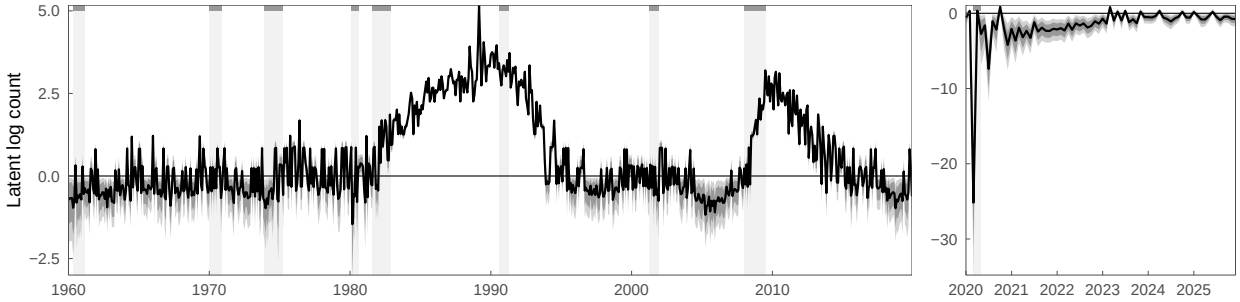

Figure 2: Estimates of the intensity of banking distress underlying the number of FDIC-insured bank failures (`BANKFAIL`).

Notes: Posterior median alongside 68/90 percent credible sets. The intensity is on the model scale, that is, on the scale of the log thresholds ($\psi_i = 0$) in the count measurement equation. Shaded areas indicate NBER recessions; the left and right panels split the sample as in Figure 1.

Figure 3 reports the shadow rates that arise when the interest rates hit the ELB. As discussed in Carriero *et al.* (2025), the shadow rates may be viewed as the monetary policy reactions to the variables in the VAR, and they capture the overall stance of policy when the central bank implements alternative policies, such as balance sheet policies and forward guidance, to stimulate the economy. The two episodes differ in depth and duration. In the first, which begins during the Great Recession and lasts until the liftoff in late 2015, the shadow federal funds rate declines gradually to about −0.9 percent in 2011–12. The shadow rates are ordered by maturity, with the one-year rate staying close to the threshold, and the five- and ten-year rates are never censored. In the second, the shadow federal funds rate falls

quickly to about −1.3 percent in 2020–21, with wider credible sets, and the shadow rates leave the ELB in early 2022 when policy tightens in response to inflation.

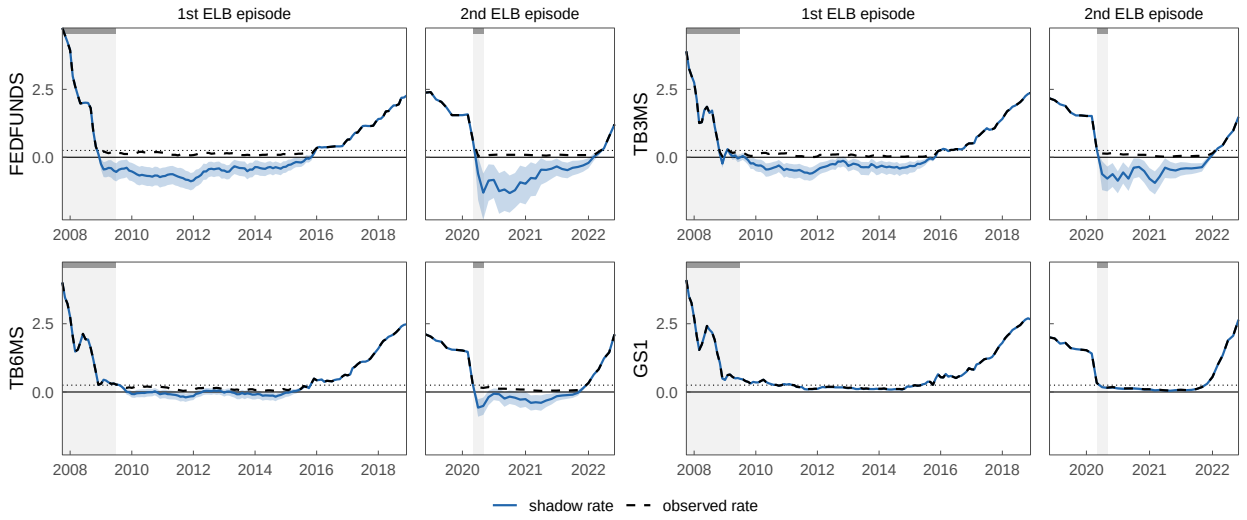

Figure 3: Estimates of the shadow rates underlying the interest rates subject to censoring at the ELB, around the two ELB episodes (2007:10 to 2018:12 and 2019:06 to 2022:06).

Notes: Solid lines show posterior medians of the shadow rates alongside 68/90 percent credible sets, dashed lines the observed interest rates. The panels show the four interest rates that have censored observations in the sample. The dotted horizontal line marks the 0.25 percent censoring threshold. Shaded areas indicate NBER recessions.

Figure 4 shows the excess bond premium and the financial conditions index, which are observed from 1973:01 and 1971:01 onwards, respectively; the model treats the earlier observations as missing and imputes them alongside the other latent states. The imputed paths rise into the 1969–70 recession, consistent with the tightening of credit and financial conditions at the end of the 1960s.

Figure 5 shows the reduced-form log innovation standard deviations of the recession indicator, the bank failure count and the federal funds rate; we report the log innovation standard deviations for the other variables in Appendix B. The innovation volatility of the federal funds rate is highest during the Volcker disinflation of 1979–82 and lowest in the 2010s, during and after the first ELB episode. The volatility of the recession indicator drifts down over the sample, and its time average is close to zero, which reflects the unit-scale normalization of Section 2.2. The circles mark the months in which the outlier component

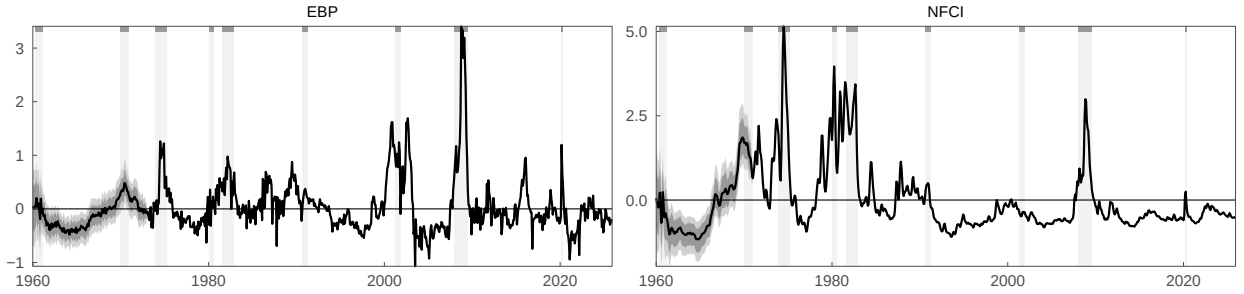

Figure 4: Imputed early observations of the excess bond premium (`EBP`) and the Chicago Fed national financial conditions index (`NFCI`), which are unavailable at the beginning of the sample.

Notes: Posterior median alongside 68/90 percent credible sets. `EBP` is observed from 1973:01, `NFCI` from 1971:01 onwards; earlier observations are imputed alongside the remaining latent states. Shaded areas indicate NBER recessions.

is active, among them several recession months and, for all three variables, the spring of 2020. There it raises the innovation standard deviations for a few months by up to a factor of about 20 without affecting the persistent volatility component.

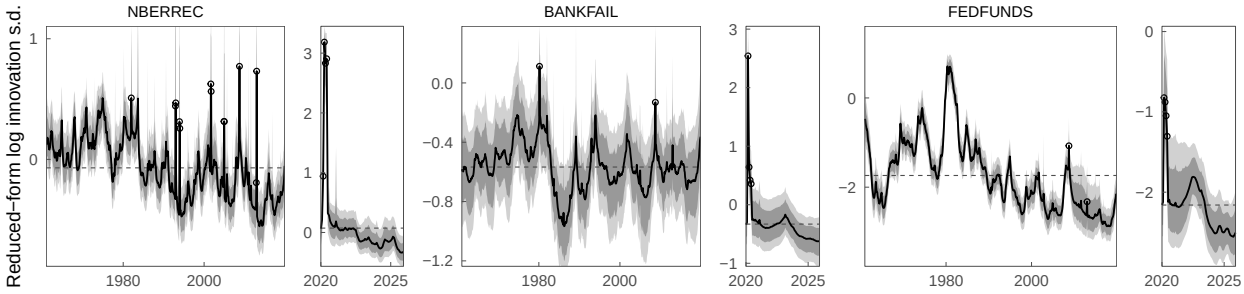

Figure 5: Estimates of the reduced-form log innovation standard deviations, $\log \sqrt{[\boldsymbol{\Sigma}_t]_{ii}}$ for `NBERREC`, `BANKFAIL` and `FEDFUNDS`.

Notes: Posterior median alongside the 68/90 percent credible sets. The reduced-form variances are computed from $\boldsymbol{\Sigma}_t = \boldsymbol{A}_0^{-1}\boldsymbol{O}_t\boldsymbol{H}_t\boldsymbol{O}_t\boldsymbol{A}_0^{-1\prime}$ and thus include the outlier component; circles mark the months in which the posterior median including the outlier component exceeds the one excluding it by a meaningful amount. Dashed horizontal lines indicate the time average of the posterior median within a panel.

We repeat Figures 1 to 4 for the homoskedastic version of the model in Appendix B. The turning points of the business cycle indicator, the two episodes of banking distress and the imputed series are essentially unchanged, although the homoskedastic indicator sepa-

rates recessions from expansions less sharply. The shadow rates differ. Without stochastic volatility, the innovation variances of the interest rates are averages over the whole sample, including the volatile years around 1980, and the shadow federal funds rate falls to about -2.2 percent in both ELB episodes, with wider credible sets. The low volatility that the stochastic volatility model estimates at the ELB keeps the shadow federal funds rate at about -0.9 and -1.3 percent in the two episodes, about one percentage point above its homoskedastic counterpart.

4.3. Conditional Forecasts and Scenario Analysis

Conditional forecasting is a popular tool for assessing the paths of a set of variables of interest under assumptions about the future paths of other variables (Waggoner and Zha, 1999; Bańbura *et al.*, 2015; Antolin-Diaz *et al.*, 2021). For example, McCracken *et al.* (2022) use it to examine the effects of monetary policy interventions and of increases in oil prices on the likelihood of a recession. We demonstrate the usefulness of the proposed framework with five conditional forecasting exercises, chosen to showcase the variable types that it accommodates: the scenarios concern the business cycle indicator, the count of bank failures and interest rates that are subject to the ELB, as well as continuous variables.

All scenarios start in 2026:01, the month after the final in-sample observation, and the forecast horizon is 36 months. The conditioning assumptions are imposed as restrictions on the joint predictive distribution derived in Appendix A, which allows us to condition on the observed variables and on their latent counterparts; the restrictions are imposed on the means while the variances of the unconditional predictive distribution are preserved (see Antolin-Diaz *et al.*, 2021). For the binary and count variables, the restrictions are placed on the latent variables: a target probability is imposed as the corresponding quantile of the business cycle indicator, and a count path as the logarithms of the interval midpoints of the counts. These scenarios fix the expected paths of the latent variables, and the realized binary and count outcomes remain random. The outlier factors are set to one over the forecast horizon, so

that the predictive distributions describe normal times and exclude the risk of further outlier episodes. For each scenario we report either the conditional forecasts of selected variables or generalized impulse response functions (GIRFs; see Koop *et al.*, 1996), computed draw by draw as the difference between the conditional and the unconditional forecasts, that is, the responses to the reduced-form innovations that implement the restrictions. Appendix B reports, for the most severe variant of each scenario, the conditional forecasts of all variables, including the latent states, alongside the unconditional forecasts.

Recession probability scenarios. We condition the business cycle indicator underlying `NBERREC` on impact at the quantiles $p \in \{0.25, 0.9\}$ of the standard normal distribution. Given the scale normalization of the indicator, this targets a conditional recession probability of approximately p in the first month of the forecast horizon. Figure 6 shows the conditional forecasts of the recession probability and of the key indicators of the NBER dating committee. Under $p = 0.9$, the recession probability falls below 0.2 within about half a year and is close to zero by 2027: in the absence of further shocks, the model expects a recession to be short, as recessions have been in the sample. Employment, industrial production, real income and sales contract on impact and recover within a year. On impact, employment growth falls by about 0.2 percent per month and industrial production by more than one percent. Inflation is essentially unaffected, and the federal funds rate falls to about 30 basis points below its path under $p = 0.25$ and stays below it over the horizon. The $p = 0.25$ scenario is close to the unconditional forecast.

Bank failures. We impose the observed path of bank failures during the Global Financial Crisis, from 2008:09 to 2009:07, as the expected path of bank failures from 2026:01 onwards, which places the restriction on the latent intensity of banking distress. The top row of Figure 7 shows the imposed path, which rises from a handful of failures to 24 per month, alongside the unconditional forecast and the implied intensity of banking distress; the bottom row shows the GIRFs of selected variables. Housing starts fall by about 25 percent within two years and remain depressed, employment growth declines modestly in

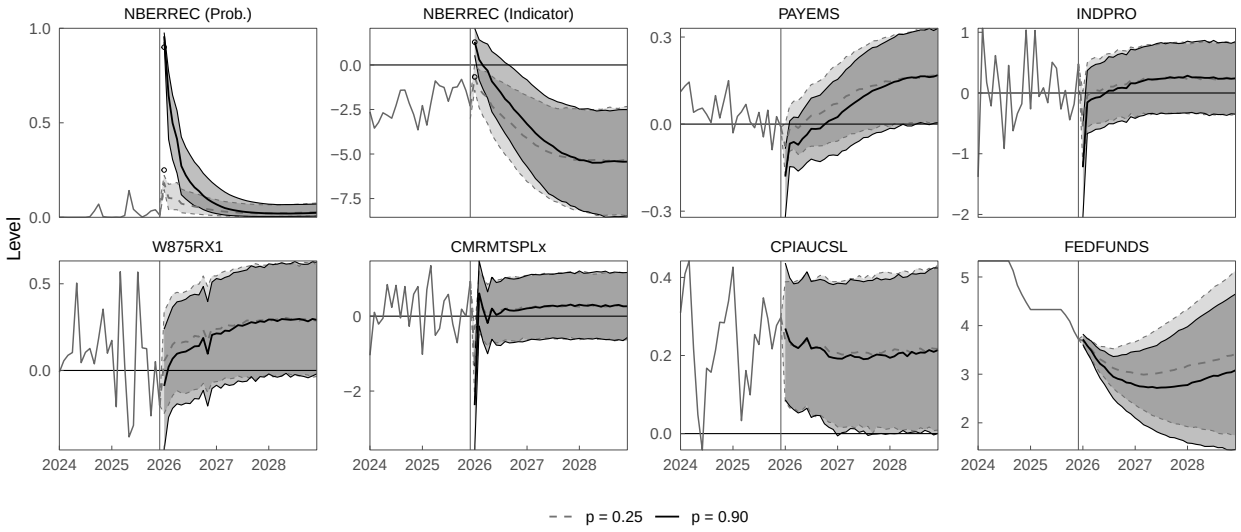

Figure 6: Recession probability scenarios: conditional forecasts when the business cycle indicator is fixed on impact at the $p = 0.25$ and $p = 0.9$ quantiles of the standard normal distribution.

Notes: Posterior medians alongside 68 percent credible sets of the conditional forecasts from 2026:01 to 2028:12, with the band boundaries drawn in the line type of the respective scenario; the grey line shows the in-sample posterior median for the preceding two years, up to the vertical line at the forecast origin. Circles mark the imposed restriction, shown on the probability scale for `NBERREC` (Prob.) and on the scale of the business cycle indicator for `NBERREC` (Indicator).

the second and third years, the excess bond premium rises after a year, and the business cycle indicator rises by about 0.4 units, which raises the recession probability only slightly from its low unconditional level. The federal funds rate rises by about 50 basis points over the first year and falls below its unconditional path at the end of the horizon. The GIRFs are responses to reduced-form innovations, and the response of the policy rate reflects the historical pattern in which waves of failures have followed periods of monetary tightening.

Growth-at-risk. Following the growth-at-risk literature (Adrian *et al.*, 2019), which relates the lower tail of the distribution of future growth to financial conditions, we shock the `NFCI` on impact by one and three standard deviations. Since the index is standardized, these correspond to moderate and severe tightenings of financial conditions. Figure 8 shows that financial conditions tighten further in the month after the shock and then ease gradually toward their unconditional path over the three-year horizon. Under the severe scenario, the

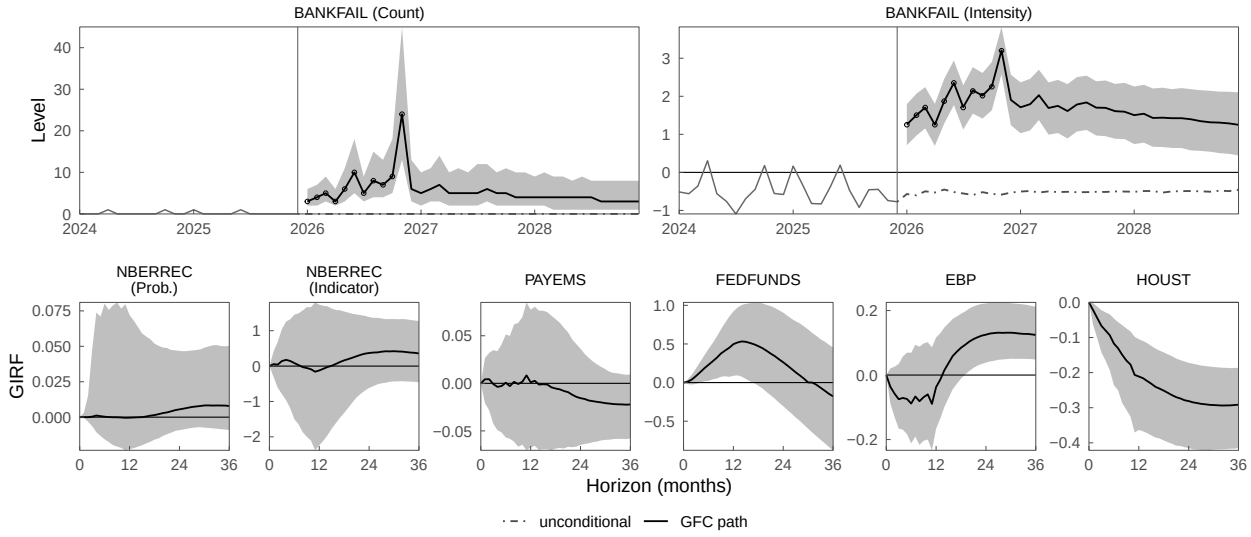

Figure 7: Bank failure scenario: the observed Global Financial Crisis path of `BANKFAIL` (2008:09 to 2009:07) is imposed from 2026:01 onwards.

Notes: The top row shows the imposed expected path of bank failures (circles mark the restrictions) alongside the unconditional forecast (dash-dotted), for the count and for the intensity of banking distress; posterior medians with 68 percent credible sets. The bottom row shows GIRFs, i.e., the draw-wise difference between the conditional and the unconditional predictive distribution, in units of the respective (transformed) variable against the forecast horizon.

growth rate of industrial production falls by about four percent on impact and remains below its unconditional path for about two years, so that the level remains permanently lower; employment growth declines for a year, to about -0.5 percent per month, and returns to its unconditional path after three years, while inflation hardly responds. The recession probability rises to nearly one for about a year and a half before it declines; under the moderate scenario it rises by a similar amount on impact but fades within about a year and a half. The responses of the continuous variables scale roughly with the size of the shock, whereas the recession probability does not, which reflects the nonlinearity of the measurement equation. Appendix B complements the GIRFs by reporting the tails of the conditional distribution, where asymmetries associated with downside risks to the real economy would materialize.

Stress tests. We impose the supervisory baseline and severely adverse scenarios of the 2026 Dodd-Frank Act stress test on the unemployment rate, CPI inflation, and the three-

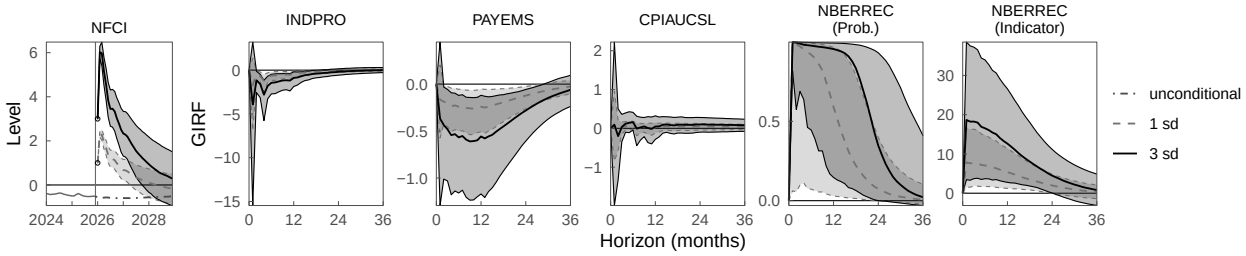

Figure 8: Growth-at-risk scenarios: `NFCI` is shocked on impact by 1 and 3 standard deviations.

Notes: The left panel shows the imposed impact restrictions (circles) alongside the unconditional forecast (dash-dotted); posterior medians with 68 percent credible sets. The remaining panels show GIRFs as in Figure 7; for `NBERREC` the responses of the recession probability (Prob.) and of the business cycle indicator (Indicator) are shown in separate panels. Figure B.7 shows the corresponding 5/95 percentile bands.

month, five-year and ten-year Treasury rates (`UNRATE`, `CPIAUCSL`, `TB3MS`, `GS5`, `GS10`). The scenario involves interest rates that are subject to censoring: where the supervisory path lies below the 0.25 percent threshold, we leave the rate unrestricted, and the model reports its shadow rate. Figure 9 shows the conditional forecasts. Under the severely adverse scenario, the unemployment rate rises to ten percent by mid-2027, the three-month rate reaches the ELB in the third quarter of 2026, and the federal funds rate, on which the scenario imposes no restriction, follows it, with a shadow rate of about -1.3 percent in 2027. The recession probability is close to one throughout 2026 and falls to near zero by the end of 2027, and employment growth falls to about -0.4 percent per month before it recovers. Bank failures hardly respond within the horizon, in line with the lag of two to three years with which failures followed the downturns of the early 1980s and of 2008. Under the baseline scenario, the recession probability stays near zero.

Oil price scenarios. We impose an adverse and a severe oil price path on `OILPRICEx`, taken from the scenarios that Kilian *et al.* (2026) calibrate to the 2026 Iran war: the adverse path raises the oil price sharply for one quarter, the severe path for three quarters, with an initial increase of about 40 percent. Figure 10 shows the imposed paths, the implied path of the federal funds rate, and the GIRFs of inflation, industrial production, the recession

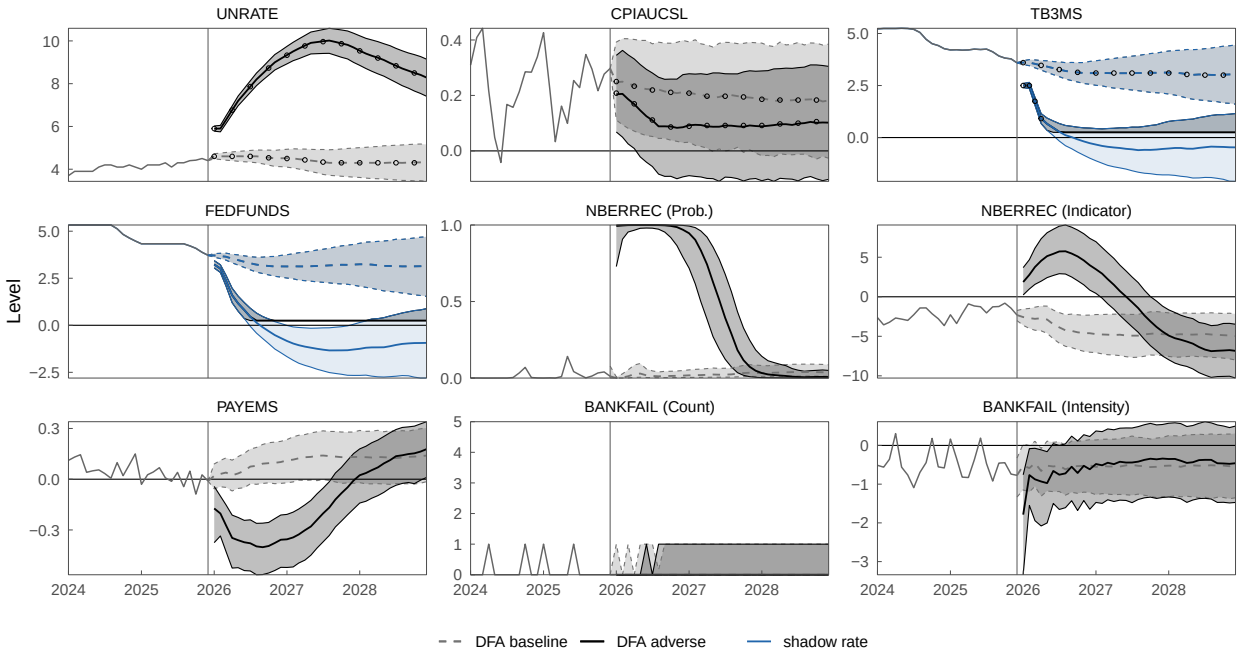

Figure 9: Stress test scenarios: conditional forecasts under the baseline and severely adverse supervisory scenarios of the 2026 Dodd-Frank Act stress test.

Notes: Posterior medians alongside 68 percent credible sets of the conditional forecasts; the grey line shows the in-sample posterior median for the preceding two years, and shadow rates are shown in blue. Circles mark the imposed restrictions on `UNRATE`, `CPIAUCSL` and `TB3MS` (the restrictions on `GS5` and `GS10` are not shown); the remaining panels show the implied paths of the other variables.

probability and the business cycle indicator. Under the severe scenario, monthly CPI inflation rises by up to 0.5 percentage points in the months of the shock and the effect dies out within two years; industrial production declines by about 0.1 percent after a year; and the recession probability rises by up to five percentage points. Under the adverse scenario, the initial response of inflation is the same but fades faster, and the responses of industrial production and of the recession probability are about half as large. The federal funds rate is essentially unchanged relative to its unconditional path under both scenarios.

The five exercises together illustrate the versatility of the framework. The questions they address would otherwise call for different models: a target for the recession probability is a question for a probit model, a wave of bank failures for a count model, a supervisory stress test with interest rates at the lower bound for a censored model, and financial and

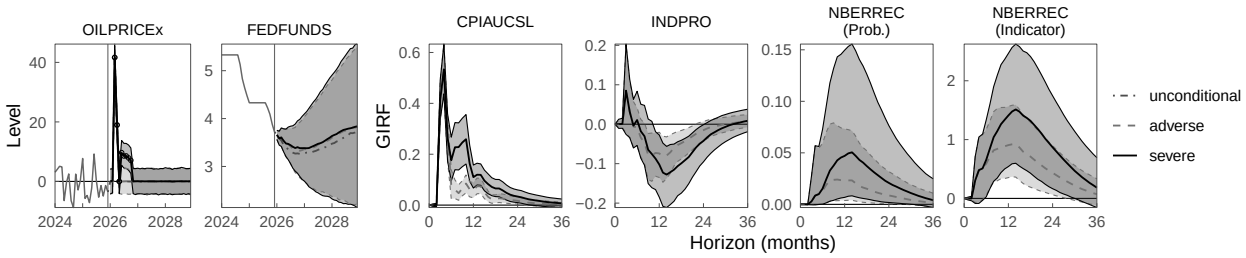

Figure 10: Oil price scenarios: adverse (one-quarter) and severe (three-quarter) oil price shock paths are imposed on `OILPRICEx`, following Kilian *et al.* (2026).

Notes: The two left panels show the imposed oil price paths (circles mark the restrictions) and the implied path of the federal funds rate alongside the unconditional forecast (dash-dotted); posterior medians with 68 percent credible sets. The remaining panels show GIRFs as in Figure 7.

oil price shocks for a standard VAR. In the proposed framework each is answered within one model, so that a scenario can combine restrictions on variables of different types, as the stress test does, and the responses of all variable types, including the recession probability, the expected number of bank failures and the shadow rates, are obtained from the same predictive distribution.

5. CONCLUDING REMARKS AND FUTURE RESEARCH

We have extended the VAR to a mix of binary, censored, count and unrestricted variables, with missing observations of any type handled in the same way. Each restricted observation is linked to a latent Gaussian variable, the counts through a threshold function from the Box–Cox family, and the latent variables are drawn, jointly or equation by equation, from a truncated normal distribution with a banded precision matrix. We have demonstrated the proposed framework in an application that delivers a business cycle indicator, the intensity of banking distress and shadow rates within one system, together with conditional forecasts in which the restrictions are placed on these variables themselves.

The framework can be extended in several directions. Ordinal variables, such as sovereign credit ratings, fit the same latent-variable representation with ordered thresholds (Zhang *et al.*, 2015), and the conditional forecasting machinery applies to them without

change. In addition, the latent representation makes structural analysis available for restricted variables. Impulse responses of the recession probability, of the expected number of bank failures and of the shadow rates to identified shocks can be computed from the same posterior draws. The identification schemes developed for standard VARs, such as recursive or sign restrictions, carry over to the latent system. We leave these extensions for future research.

DECLARATION OF GENERATIVE AI

We acknowledge the use of Claude Code and Codex for drafting, editing and brainstorming. A previous version of the paper was reviewed by Refine, and its suggestions are incorporated in the current draft. All AI-assisted suggestions were carefully reviewed, and we take full responsibility for the final content and accuracy of this paper.

REFERENCES

Adrian T, Boyarchenko N, and Giannone D (2019), "Vulnerable growth," *American Economic Review* **109**(4), 1263–1289.

Al-Osh MA, and Alzaid AA (1987), "First-order integer-valued autoregressive (INAR(1)) process," *Journal of Time Series Analysis* **8**(3), 261–275.

Albert J, and Chib S (1993), "Bayesian analysis of binary and polychotomous response data," *Journal of the American Statistical Association* **88**(422), 669–679.

Anceschi N, Fasano A, Durante D, and Zanella G (2023), "Bayesian conjugacy in probit, Tobit, multinomial probit and extensions: A review and new results," *Journal of the American Statistical Association* **118**(542), 1451–1469.

Antolin-Diaz J, Petrella I, and Rubio-Ramírez JF (2021), "Structural scenario analysis with SVARs," *Journal of Monetary Economics* **117**, 798–815.

Bańbura M, Giannone D, and Lenza M (2015), "Conditional forecasts and scenario analysis with vector autoregressions for large cross-sections," *International Journal of Forecasting* **31**(3), 739–756.

Bańbura M, Giannone D, and Reichlin L (2010), "Large Bayesian vector auto regressions," *Journal of Applied Econometrics* **25**(1), 71–92.

Berry LR, and West M (2020), "Bayesian forecasting of many count-valued time series," *Journal of Business & Economic Statistics* **38**(4), 872–887.

Bishop CM (2006), *Pattern Recognition and Machine Learning*, Springer.

Botev ZI (2017), "The normal law under linear restrictions: Simulation and estimation via minimax tilting," *Journal of the Royal Statistical Society Series B: Statistical Methodology* **79**(1), 125–148.

Carriero A, Clark TE, and Marcellino M (2019), "Large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors," *Journal of Econometrics* **212**(1), 137–154.

Carriero A, Clark TE, Marcellino M, and Mertens E (2024), "Addressing COVID-19 outliers in BVARs with stochastic volatility," *Review of Economics and Statistics* **106**(5), 1403–1417.

——— (2025), "Forecasting with shadow rate VARs," *Quantitative Economics* **16**(3), 795–822.

Carvalho CM, Polson NG, and Scott JG (2010), "The horseshoe estimator for sparse signals," *Biometrika* **97**(2), 465–480.

Chan JCC (2023), "Comparing stochastic volatility specifications for large Bayesian VARs," *Journal of Econometrics* **235**(2), 1419–1446.

Chan JCC, Koop G, and Yu X (2024), "Large order-invariant Bayesian VARs with stochastic volatility," *Journal of Business & Economic Statistics* **42**(2), 825–837.

Chan JCC, Pettenuzzo D, Poon A, and Zhu D (2025), "Conditional forecasts in large Bayesian VARs with multiple equality and inequality constraints," *Journal of Economic Dynamics and Control* **173**, 105061.

Chan JCC, Poon A, and Zhu D (2023), "High-dimensional conditionally Gaussian state space models with missing data," *Journal of Econometrics* **236**(1), 105468.

Chib S (1992), "Bayes inference in the Tobit censored regression model," *Journal of Econometrics* **51**(1–2), 79–99.

Clark TE (2011), "Real-time density forecasts from Bayesian vector autoregressions with stochastic volatility," *Journal of Business and Economic Statistics* **29**(3), 327–341.

Cogley T, and Sargent TJ (2005), "Drifts and volatilities: Monetary policies and outcomes in the post WWII US," *Review of Economic Dynamics* **8**(2), 262–302.

Cohn JB, Liu Z, and Wardlaw MI (2022), "Count (and count-like) data in finance," *Journal of Financial Economics* **146**(2), 529–551.

Davis RA, Fokianos K, Holan SH, Joe H, Livsey J, Lund R, Pipiras V, and Ravishanker N (2021), "Count time series: A methodological review," *Journal of the American Statistical Association* **116**(535), 1533–1547.

Doan T, Litterman RB, and Sims CA (1984), "Forecasting and conditional projection using realistic prior distributions," *Econometric Reviews* **3**(1), 1–100.

Dueker M (2005), "Dynamic forecasts of qualitative variables: A Qual VAR model of US recessions," *Journal of Business & Economic Statistics* **23**(1), 96–104.

Dueker M, and Assenmacher-Wesche K (2010), "Forecasting macro variables with a Qual VAR business cycle turning point index," *Applied Economics* **42**(23), 2909–2920.

Dueker M, and Nelson CR (2006), "Business-cycle filtering of macroeconomic data via a latent business-cycle index," *Macroeconomic Dynamics* **10**(5), 573–594.

Fokianos K (2024), "Multivariate count time series modelling," *Econometrics and Statistics* **31**, 100–116.

Giannone D, Lenza M, and Primiceri GE (2015), "Prior selection for vector autoregressions," *Review of Economics and Statistics* **97**(2), 436–451.

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

Gneiting T, and Raftery AE (2007), "Strictly proper scoring rules, prediction, and estimation," *Journal of the American statistical Association* **102**(477), 359–378.

Heinen A, and Rengifo E (2007), "Multivariate autoregressive modeling of time series count data using copulas," *Journal of Empirical Finance* **14**(4), 564–583.

Jia Y, Kechagias S, Livsey J, Lund R, and Pipiras V (2023), "Latent Gaussian count time series," *Journal of the American Statistical Association* **118**(541), 596–606.

Johannsen BK, and Mertens E (2021), "A time-series model of interest rates with the effective lower bound," *Journal of Money, Credit and Banking* **53**(5), 1005–1046.

Kilian L, Plante MD, Richter AW, and Zhou X (2026), "The impact of the 2026 Iran war on U.S. inflation: A scenario analysis," Working Paper 2609, Federal Reserve Bank of Dallas.

Koop G (2013), "Forecasting with medium and large Bayesian VARs," *Journal of Applied Econometrics* **28**(2), 177–203.

Koop G, Pesaran MH, and Potter SM (1996), "Impulse response analysis in nonlinear multivariate models," *Journal of Econometrics* **74**(1), 119–147.

Kowal DR, and Canale A (2020), "Simultaneous transformation and rounding (STAR) models for integer-valued data," *Electronic Journal of Statistics* **14**(1), 1744–1772.

McCracken MW, McGillicuddy JT, and Owyang MT (2022), "Binary conditional forecasts," *Journal of Business & Economic Statistics* **40**(3), 1246–1258.

McCracken MW, and Ng S (2016), "FRED-MD: A monthly database for macroeconomic research," *Journal of Business & Economic Statistics* **34**(4), 574–589.

Neal RM (2011), "MCMC Using Hamiltonian Dynamics," in S Brooks, A Gelman, GL Jones, and XL Meng (eds.) "Handbook of Markov Chain Monte Carlo," 113–162, Boca Raton: Chapman & Hall/CRC.

Nishimura A, Dunson DB, and Lu J (2020), "Discontinuous Hamiltonian Monte Carlo for discrete parameters and discontinuous likelihoods," *Biometrika* **107**(2), 365–380.

Nishimura A, Zhang Z, and Suchard MA (2025), "Zigzag path connects two Monte Carlo samplers: Hamiltonian counterpart to a piecewise deterministic Markov process," *Journal of the American Statistical Association* **120**(550), 1077–1089.

Pakman A, and Paninski L (2014), "Exact Hamiltonian Monte Carlo for truncated multivariate Gaussians," *Journal of Computational and Graphical Statistics* **23**(2), 518–542.

Poon A, and Zhu D (2024), "Do recessions and bear markets occur concurrently across countries? A multinomial logistic approach," *Journal of Financial Econometrics* **22**(5), 1482–1502.

Sims CA (1980), "Macroeconomics and reality," *Econometrica* **48**(1), 1–48.

Waggoner DF, and Zha T (1999), "Conditional forecasts in dynamic multivariate models," *Review of Economics and Statistics* **81**(4), 639–651.

Zhang X, Boscardin WJ, Belin TR, Wan X, He Y, and Zhang K (2015), "A Bayesian method for analyzing combinations of continuous, ordinal, and nominal categorical data with missing values," *Journal of Multivariate Analysis* **135**, 43–58.

Zhang Z, Chin A, Nishimura A, and Suchard MA (2026), "`hdtg`: An R package for high-dimensional truncated normal simulation," *The R Journal* **18**(3).

Zito J, and Kowal DR (2025), "A dynamic copula model for probabilistic forecasting of non-Gaussian multivariate time series," arXiv preprint arXiv:2502.16874.

Online Appendix

A. JOINT PREDICTIVE DISTRIBUTION

This appendix derives the joint predictive distribution of the latent and observed variables over the forecast horizon, conditional on the parameters and the volatilities. It then shows how the conditioning assumptions of Section 4.3 are imposed on this distribution, which shifts the means of the restricted elements and preserves the covariance of the unconditional forecast.

The model in (4) can be written in structural form as

$$\boldsymbol{B}_0\boldsymbol{y}_t = \boldsymbol{b} + \boldsymbol{B}_1\boldsymbol{y}_{t-1} + \cdots + \boldsymbol{B}_P\boldsymbol{y}_{t-P} + \boldsymbol{O}_t\boldsymbol{H}_t^{1/2}\boldsymbol{\epsilon}_t, \tag{A.1}$$

where $\boldsymbol{B}_0 = \boldsymbol{A}_0$, $\boldsymbol{B}_p = \boldsymbol{A}_0\boldsymbol{A}_p$ for $p = 1, \ldots, P$ and $\boldsymbol{b} = \boldsymbol{A}_0\boldsymbol{a}$. Following Chan *et al.* (2025), we stack (A.1) over the forecast horizon $t = T+1, \ldots, T+F$. Let $\boldsymbol{y}_f = (\boldsymbol{y}_{T+1}', \ldots, \boldsymbol{y}_{T+F}')'$ and define

$$\widetilde{\boldsymbol{M}}_f = \begin{bmatrix} \boldsymbol{B}_0 & \boldsymbol{0}_{n\times n} & \cdots & \cdots & \cdots & \cdots & \boldsymbol{0}_{n\times n} \\ -\boldsymbol{B}_1 & \boldsymbol{B}_0 & \boldsymbol{0}_{n\times n} & \cdots & \cdots & \cdots & \boldsymbol{0}_{n\times n} \\ -\boldsymbol{B}_2 & -\boldsymbol{B}_1 & \boldsymbol{B}_0 & \boldsymbol{0}_{n\times n} & \cdots & \cdots & \boldsymbol{0}_{n\times n} \\ \vdots & \ddots & \ddots & \ddots & \ddots & & \vdots \\ -\boldsymbol{B}_P & \cdots & & -\boldsymbol{B}_1 & \boldsymbol{B}_0 & \boldsymbol{0}_{n\times n} & \vdots \\ \boldsymbol{0}_{n\times n} & \ddots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \vdots \\ \boldsymbol{0}_{n\times n} & \cdots & \boldsymbol{0}_{n\times n} & -\boldsymbol{B}_P & \cdots & -\boldsymbol{B}_1 & \boldsymbol{B}_0 \end{bmatrix}, \quad \widetilde{\boldsymbol{m}}_f = \begin{bmatrix} \boldsymbol{b} + \sum_{p=1}^{P} \boldsymbol{B}_p\boldsymbol{y}_{T+1-p} \\ \boldsymbol{b} + \sum_{p=2}^{P} \boldsymbol{B}_p\boldsymbol{y}_{T+2-p} \\ \boldsymbol{b} + \sum_{p=3}^{P} \boldsymbol{B}_p\boldsymbol{y}_{T+3-p} \\ \vdots \\ \boldsymbol{b} + \boldsymbol{B}_P\boldsymbol{y}_T \\ \boldsymbol{b} \\ \vdots \\ \boldsymbol{b} \end{bmatrix},$$

and $\boldsymbol{\Omega}_f^{1/2} = \text{bdiag}(\boldsymbol{O}_{T+1}\boldsymbol{H}_{T+1}^{1/2}, \ldots, \boldsymbol{O}_{T+F}\boldsymbol{H}_{T+F}^{1/2})$, where the log-volatilities over the forecast horizon are simulated forward from their state equations and the outlier factors are set to one. The stacked system is $\widetilde{\boldsymbol{M}}_f\boldsymbol{y}_f = \widetilde{\boldsymbol{m}}_f + \boldsymbol{\Omega}_f^{1/2}\boldsymbol{\epsilon}_f$ with $\boldsymbol{\epsilon}_f \sim \mathcal{N}(\boldsymbol{0}_{nF}, \boldsymbol{I}_{nF})$. With $\boldsymbol{M}_f = \boldsymbol{\Omega}_f^{-1/2}\widetilde{\boldsymbol{M}}_f$

and $\boldsymbol{m}_f = \boldsymbol{\Omega}_f^{-1/2}\widetilde{\boldsymbol{m}}_f$, it becomes $\boldsymbol{M}_f\boldsymbol{y}_f = \boldsymbol{m}_f + \boldsymbol{\epsilon}_f$, so that $\boldsymbol{y}_f = \boldsymbol{M}_f^{-1}\boldsymbol{m}_f + \boldsymbol{M}_f^{-1}\boldsymbol{\epsilon}_f$ and, conditional on the parameters and on the volatilities over the forecast horizon,

$$\boldsymbol{y}_f \sim \mathcal{N}\left(\boldsymbol{\mu}_f, \boldsymbol{\Sigma}_f\right), \quad \boldsymbol{\mu}_f = \boldsymbol{M}_f^{-1}\boldsymbol{m}_f, \quad \boldsymbol{\Sigma}_f = (\boldsymbol{M}_f'\boldsymbol{M}_f)^{-1}. \tag{A.2}$$

Draws from (A.2) are obtained by drawing $\boldsymbol{\epsilon}_f$ and solving the block lower triangular system $\boldsymbol{M}_f\boldsymbol{y}_f = \boldsymbol{m}_f + \boldsymbol{\epsilon}_f$ by forward substitution, and the unconditional forecasts collect these draws across the posterior draws of the parameters and the simulated volatilities. The draws of the latent binary, count and censored variables are mapped to the observed outcomes through the measurement equations (1) to (3), so that the predictive distribution of an observed variable is the distribution of these mapped draws. For the homoskedastic version of the model, in which $\boldsymbol{\Sigma}$ is drawn directly, we set $\boldsymbol{B}_0$ to the inverse of the lower Cholesky factor of $\boldsymbol{\Sigma}$, which satisfies $\boldsymbol{B}_0\boldsymbol{\Sigma}\boldsymbol{B}_0' = \boldsymbol{I}_n$, and $\boldsymbol{\Omega}_f^{1/2} = \boldsymbol{I}_{nF}$; then $\boldsymbol{M}_f = \widetilde{\boldsymbol{M}}_f$ and $\boldsymbol{m}_f = \widetilde{\boldsymbol{m}}_f$.

A scenario restricts q elements or linear combinations of $\boldsymbol{y}_f$, observed or latent, through a $q \times nF$ matrix $\boldsymbol{R}$. We impose the restriction $\boldsymbol{R}\boldsymbol{y}_f \sim \mathcal{N}(\boldsymbol{r}, \boldsymbol{W})$, where $\boldsymbol{r}$ collects the imposed values and $\boldsymbol{W}$ is the restriction covariance; hard restrictions correspond to $\boldsymbol{W} = \boldsymbol{0}$. With $\boldsymbol{Q} = \boldsymbol{R}\boldsymbol{\Sigma}_f\boldsymbol{R}'$, the conditional forecast distribution is Gaussian with mean $\boldsymbol{\mu}_f + \boldsymbol{\Sigma}_f\boldsymbol{R}'\boldsymbol{Q}^{-1}(\boldsymbol{r} - \boldsymbol{R}\boldsymbol{\mu}_f)$ and covariance $\boldsymbol{\Sigma}_f - \boldsymbol{\Sigma}_f\boldsymbol{R}'\boldsymbol{Q}^{-1}\boldsymbol{R}\boldsymbol{\Sigma}_f + \boldsymbol{\Sigma}_f\boldsymbol{R}'\boldsymbol{Q}^{-1}\boldsymbol{W}\boldsymbol{Q}^{-1}\boldsymbol{R}\boldsymbol{\Sigma}_f$. In Section 4.3 we set $\boldsymbol{W} = \boldsymbol{Q}$, so that the restricted elements have mean $\boldsymbol{r}$ and the covariance of the conditional forecast equals $\boldsymbol{\Sigma}_f$: the restrictions shift the means and preserve the uncertainty of the unconditional forecast. The conditional distribution depends on the reduced-form objects $\boldsymbol{\mu}_f$ and $\boldsymbol{\Sigma}_f$ only, so no structural identification is needed. For computation, we write $\boldsymbol{y}_f = \boldsymbol{\mu}_f + \boldsymbol{M}_f^{-1}\boldsymbol{\epsilon}_f$ and obtain the conditional draw by adjusting the draw of $\boldsymbol{\epsilon}_f$ used for the unconditional path (see Antolin-Diaz *et al.*, 2021), so that the difference between the conditional and the unconditional path of each draw is the response to the restrictions; the GIRFs in Section 4.3 are the posterior distributions of these differences. A restriction on a count is imposed at the logarithm of the midpoint of the interval of the restricted count, and

a restriction on the recession probability at the corresponding quantile of the business cycle indicator, which targets the probability exactly when the innovation variance of the indicator equals one and approximately under stochastic volatility. In both cases the restriction fixes the mean of the latent variable, and the realized count or binary outcome remains random under the scenario.

B. DATA AND ADDITIONAL EMPIRICAL RESULTS

This appendix lists the variables entering the model (Table B.1), plots them (Figure B.1), and reports the reduced-form log innovation standard deviations for the full set of variables (Figure B.2). It then repeats the estimates of the latent states of Section 4.2 for the homoskedastic version of the model. This version replaces stochastic volatility and the outlier component by a constant covariance matrix $\boldsymbol{\Sigma}_t = \boldsymbol{\Sigma}$ with the inverse Wishart prior $\boldsymbol{\Sigma} \sim \mathcal{W}^{-1}(n+2, \boldsymbol{I}_n)$; all other specification choices are unchanged. Each draw of $\boldsymbol{\Sigma}$ is normalized as $\boldsymbol{\Sigma}^* = \boldsymbol{D}^{-1}\boldsymbol{\Sigma}\boldsymbol{D}^{-1}$ with $d_i = \sqrt{[\boldsymbol{\Sigma}]_{ii}}$ for the binary variables and $d_i = 1$ otherwise, which is the normalization of Dueker (2005). Finally, it reports additional results from the conditional forecasting exercises of Section 4.3: Figure B.7 shows the 5/95 percentile bands of the growth-at-risk GIRFs, and Figures B.8 to B.12 show, for the most severe variant of each scenario family, the conditional forecasts of all variables and their latent states.

Table B.1: The 25 monthly US macro-financial variables used in the application together with their transformations.

Mnemonic	Description	Type	Tfm.
NBERREC	NBER recession indicator	b	–
BANKFAIL	Number of FDIC-insured bank failures	z	–[a]
FEDFUNDS	Effective federal funds rate	c	–
TB3MS	3-month Treasury bill rate	c	–
TB6MS	6-month Treasury bill rate	c	–
GS1	1-year Treasury rate	c	–
GS5	5-year Treasury rate	c	–
GS10	10-year Treasury rate	c	–
PAYEMS	All employees: total nonfarm	u	$100\,\Delta\log$
INDPRO	Industrial production index	u	$100\,\Delta\log$
CMRMTSPLx	Real manufacturing and trade industries sales	u	$100\,\Delta\log$
W875RX1	Real personal income ex transfer receipts	u	$100\,\Delta\log$
UNRATE	Civilian unemployment rate	u	–
CES0600000007	Average weekly hours: goods-producing	u	log
CES0600000008	Average hourly earnings: goods-producing	u	$100\,\Delta\log$
PCEPI	Personal consumption expenditures price index	u	$100\,\Delta\log$
CPIAUCSL	Consumer price index: all items	u	$100\,\Delta\log$
HOUST	Housing starts: total new privately owned	u	log
SP500	S&P 500 composite stock price index	u	$100\,\Delta\log$
EXUSUKx	U.S.–U.K. foreign exchange rate	u	$100\,\Delta\log$
OILPRICEx	Crude oil price (spliced WTI and Cushing)	u	$100\,\Delta\log$
BAA	Moody's seasoned Baa corporate bond yield	u	–
TS10G1	Term spread: 10-year minus 1-year Treasury rate	u	–
EBP[b]	Excess bond premium	u	–
NFCI[c]	Chicago Fed national financial conditions index	u	–

Notes: Monthly data; the estimation sample runs from 1960-01 to 2025-12. Type indicates the variable type: binary (b), count (z), censored (c), or unrestricted (u). Tfm. denotes the transformation applied to the raw series: none (–), natural logarithm (log), or log-difference in percent ($100\,\Delta\log$). Sources: FRED-MD (2026-06 vintage), NBER business cycle chronology, Gilchrist and Zakrajšek (2012) for the excess bond premium (Federal Reserve Board update), Federal Reserve Bank of Chicago (NFCI), and the FDIC failures and assistance database (BANKFAIL). [a] Enters as an integer count; the count measurement equation uses the log thresholds ($\psi_i = 0$). [b] Available from 1973-01 onwards. [c] Available from 1971-01 onwards.

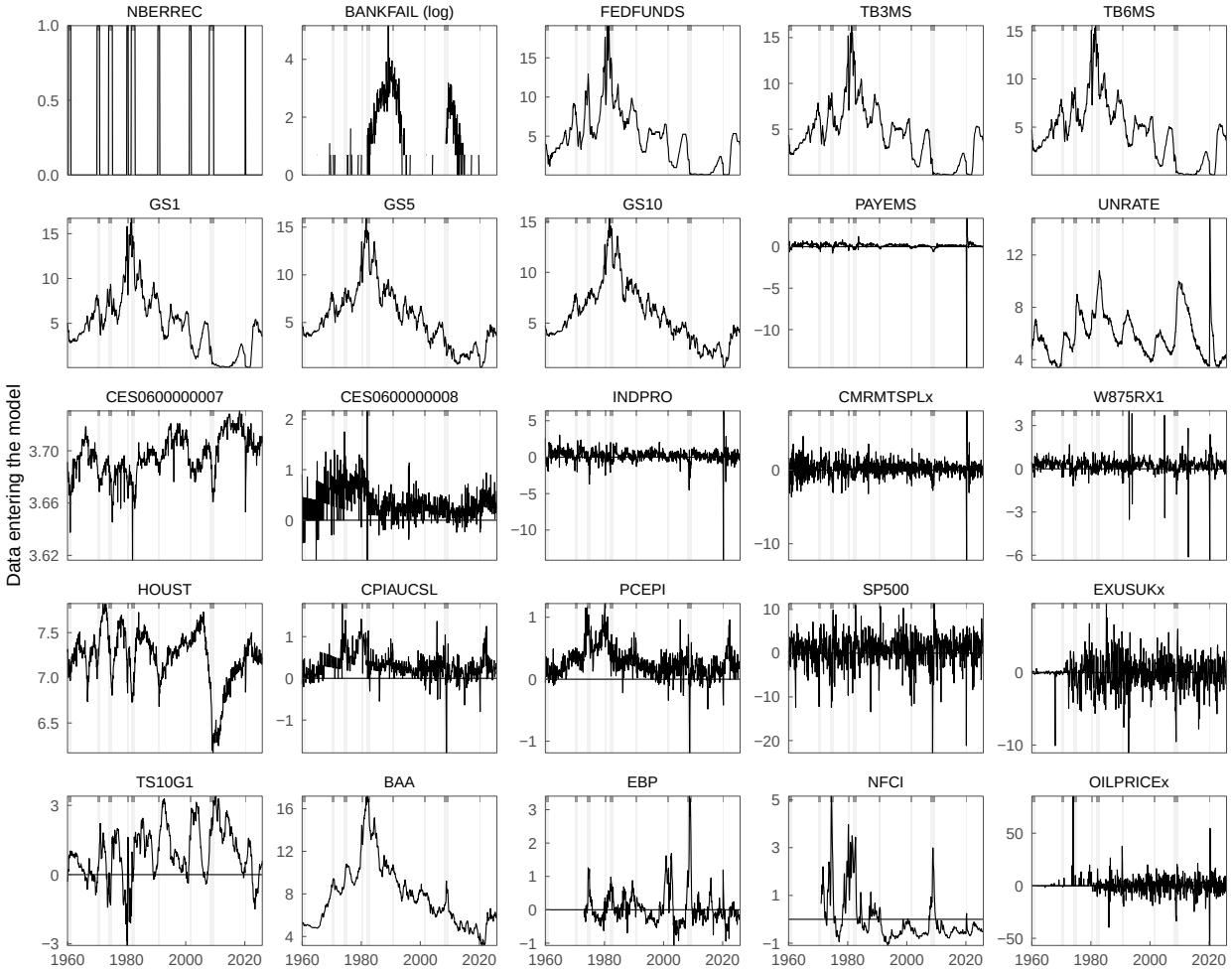

Figure B.1: Time series of the 25 variables entering the model, 1960:01 to 2025:12.

Notes: Monthly observations from 1960:01 to 2025:12, in the units in which they enter the VAR, ordered by type: the binary, the count, the censored, and the unrestricted variables. Table B.1 lists definitions and transformations. The count variable is plotted as the logarithm of the count, which is the threshold scale of its measurement equation; months without bank failures have no logarithm and are left blank, and they are not missing observations. The remaining gaps mark the observations that the model treats as missing, `EBP` before 1973:01 and `NFCI` before 1971:01.

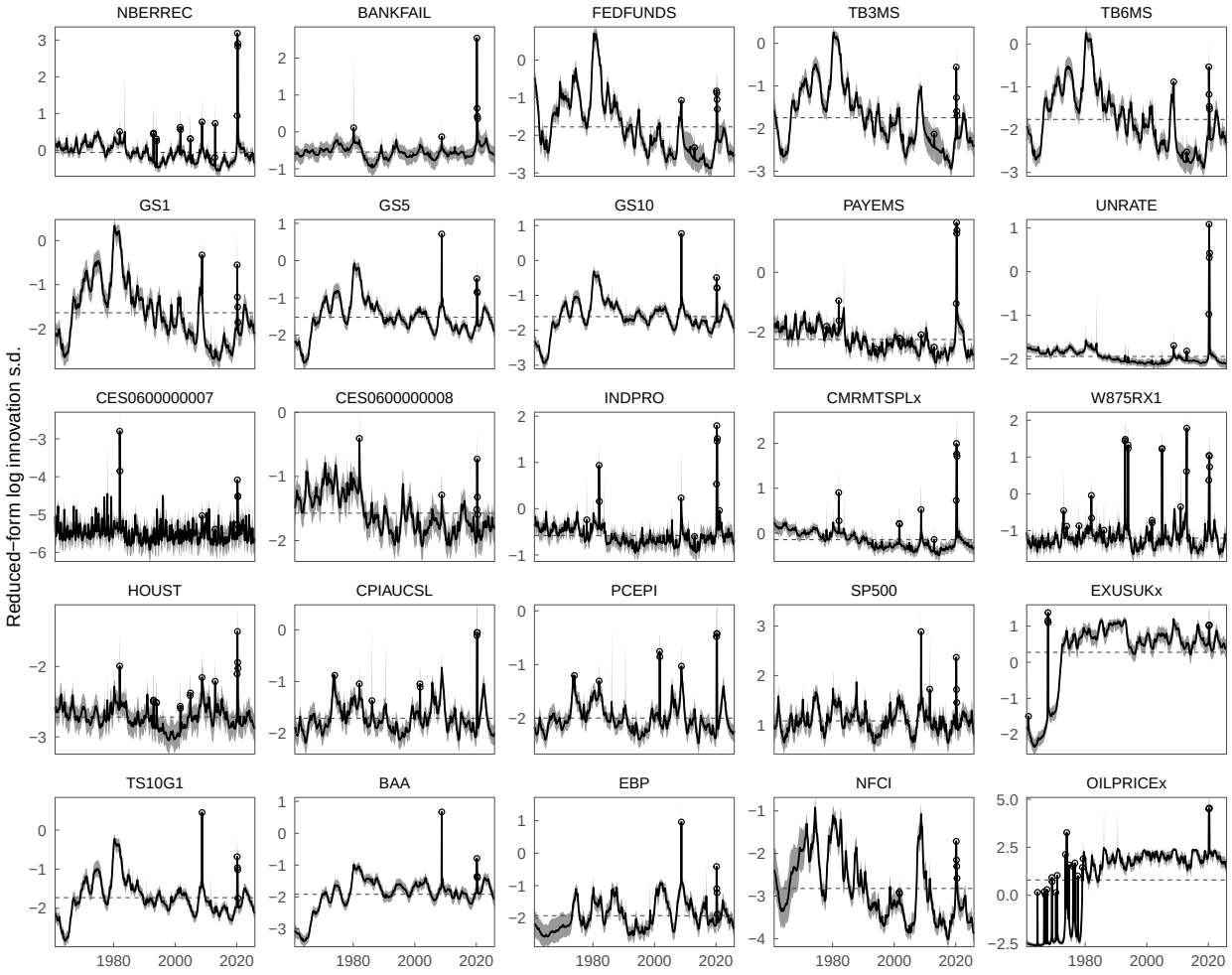

Figure B.2: Estimates of the reduced-form log innovation standard deviations, $\log \sqrt{[\boldsymbol{\Sigma}_t]_{ii}}$, for all variables, on the scale of the latent variable for the binary, count and censored variables and on the scale of the observed data otherwise.

Notes: See Figure 5, which shows the first three variables of this figure for the pre-COVID-19 and the COVID-19 sample separately. Here, each panel covers the full sample and the vertical scale is not restricted to the range of the posterior median.

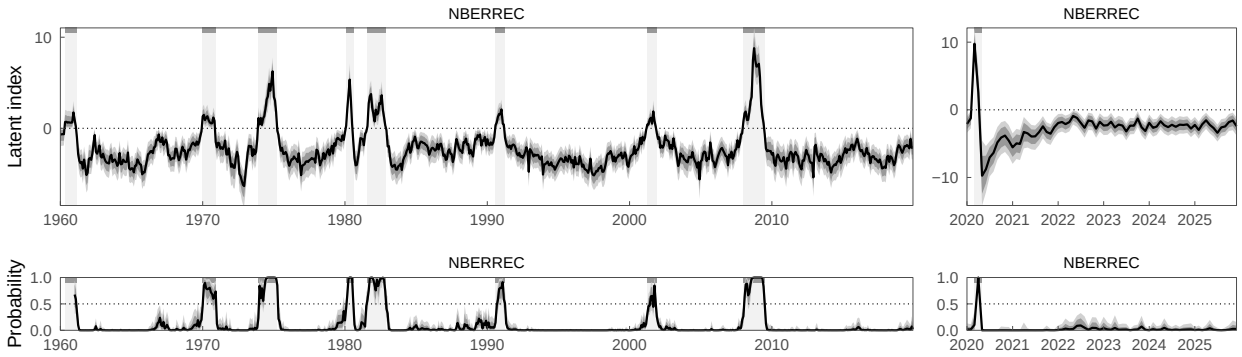

Figure B.3: Homoskedastic model: estimates of the business cycle indicator underlying `NBERREC` (upper panels) and the implied recession probabilities (lower panels).

Notes: See Figure 1.

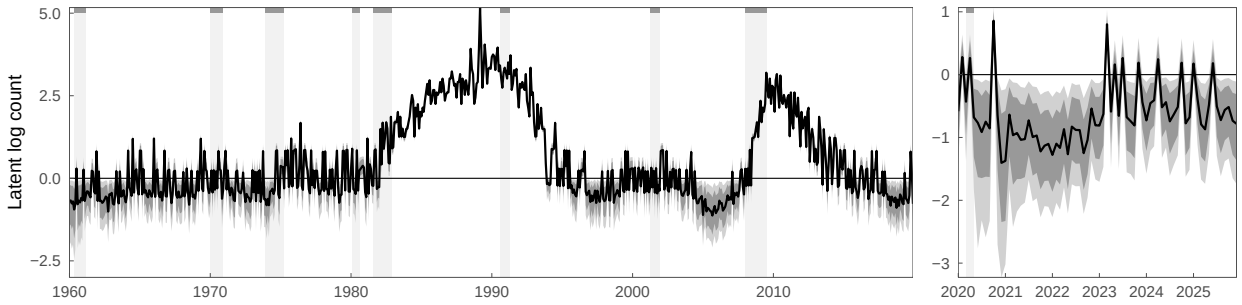

Figure B.4: Homoskedastic model: estimates of the intensity of banking distress underlying the number of FDIC-insured bank failures (`BANKFAIL`).

Notes: See Figure 2.

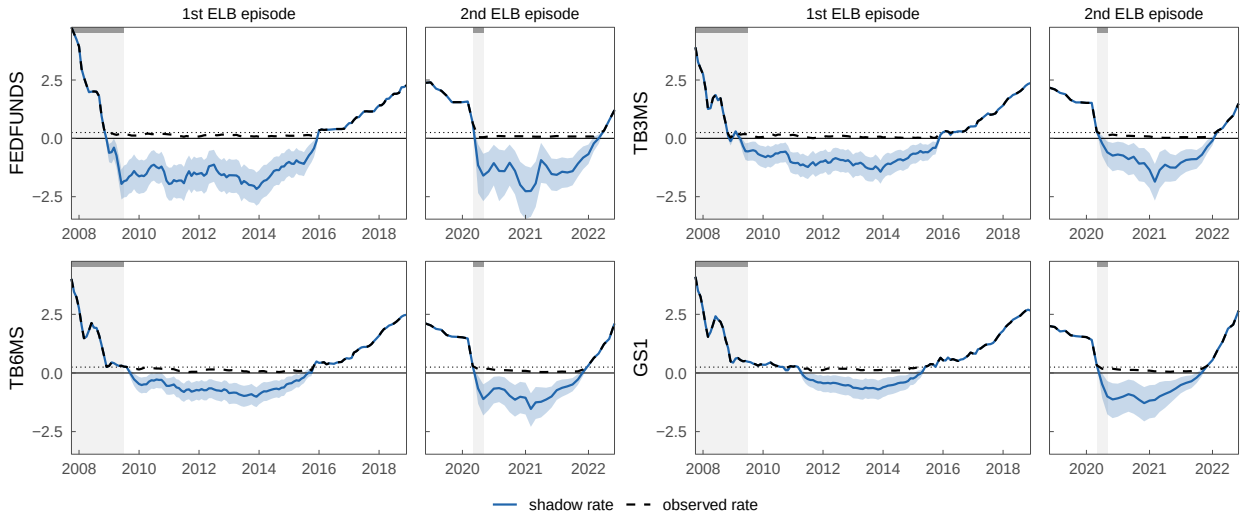

Figure B.5: Homoskedastic model: estimates of the shadow rates underlying the interest rates subject to censoring at the ELB.

Notes: See Figure 3.

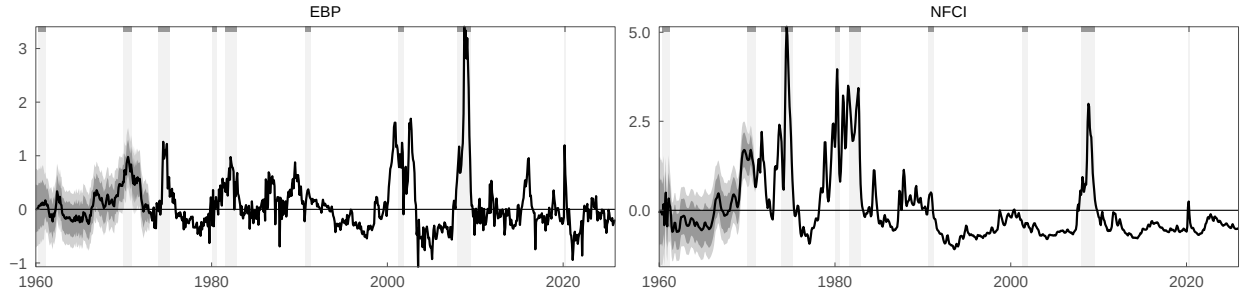

Figure B.6: Homoskedastic model: imputed early observations of the excess bond premium (`EBP`) and the Chicago Fed national financial conditions index (`NFCI`), which are unavailable at the beginning of the sample.

Notes: See Figure 4.

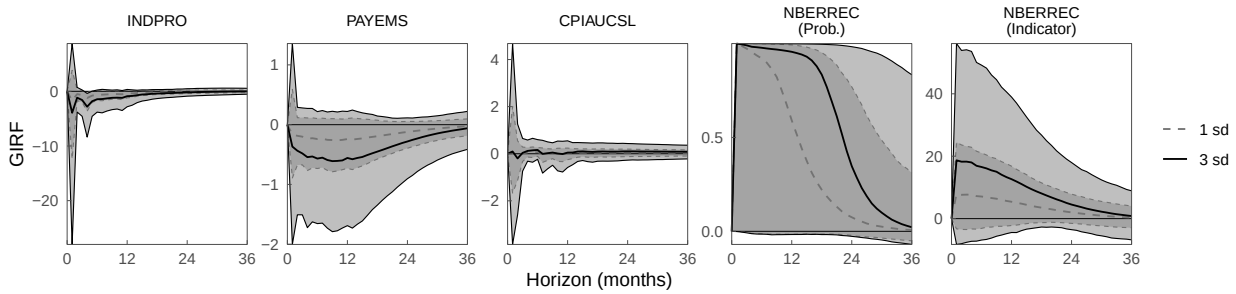

Figure B.7: Growth-at-risk scenarios: 5/95 percentile bands of the GIRFs shown in Figure 8.

Notes: Posterior medians alongside 90 percent credible sets of the draw-wise difference between the conditional and the unconditional predictive distribution.

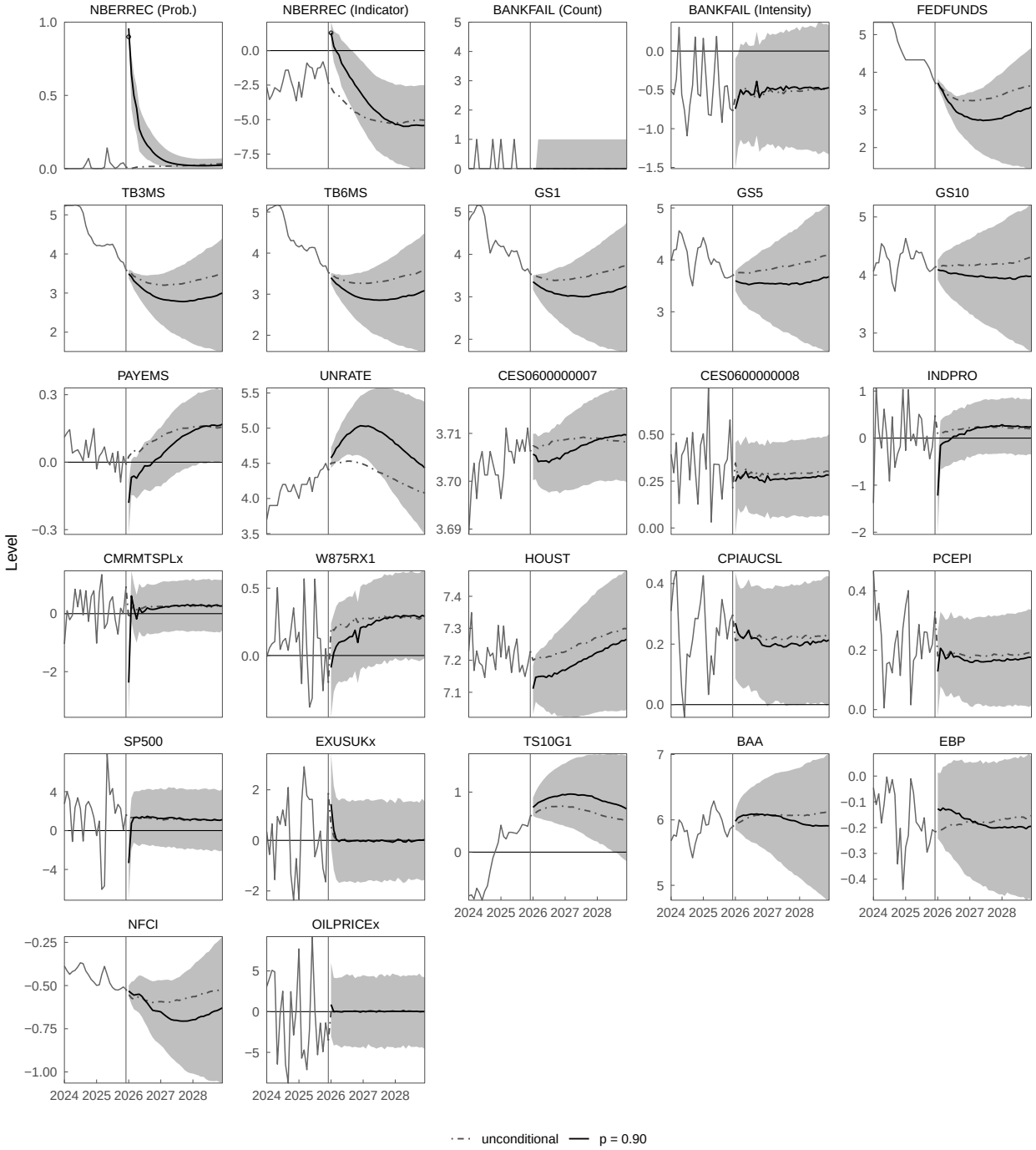

Figure B.8: Recession probability scenario ($p = 0.9$): conditional forecasts for all variables.

Notes: Posterior medians of the conditional (black solid) and unconditional (grey dash-dotted) forecasts, with 68 percent credible sets for the conditional forecast; the preceding two years show the in-sample posterior median. Circles mark the imposed restrictions. Panels labeled "(Indicator)", "(Intensity)" and "(Shadow rate)" show the associated latent states on their model scale.

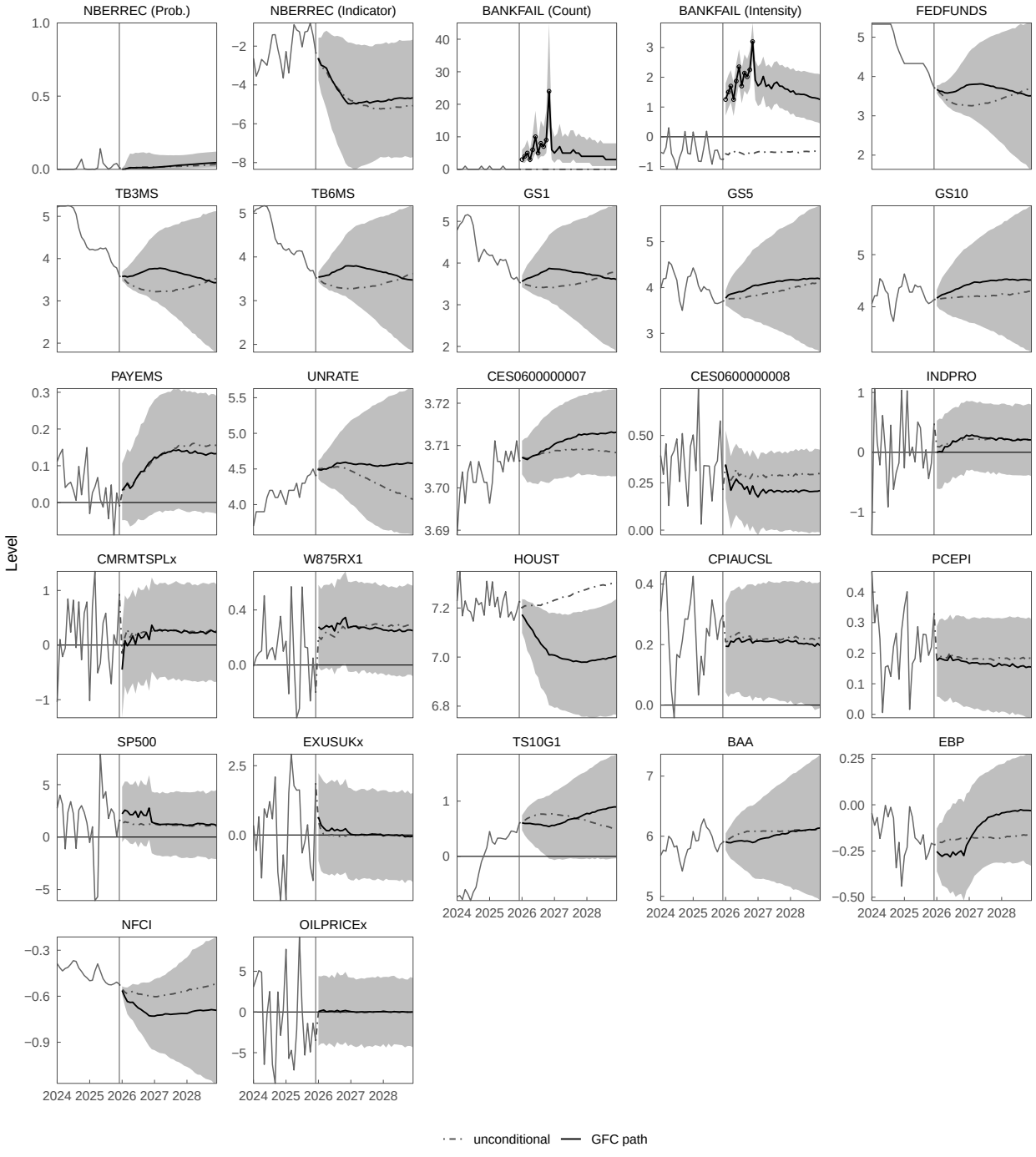

Figure B.9: Bank failure scenario (Global Financial Crisis path): conditional forecasts for all variables.

Notes: See Figure B.8.

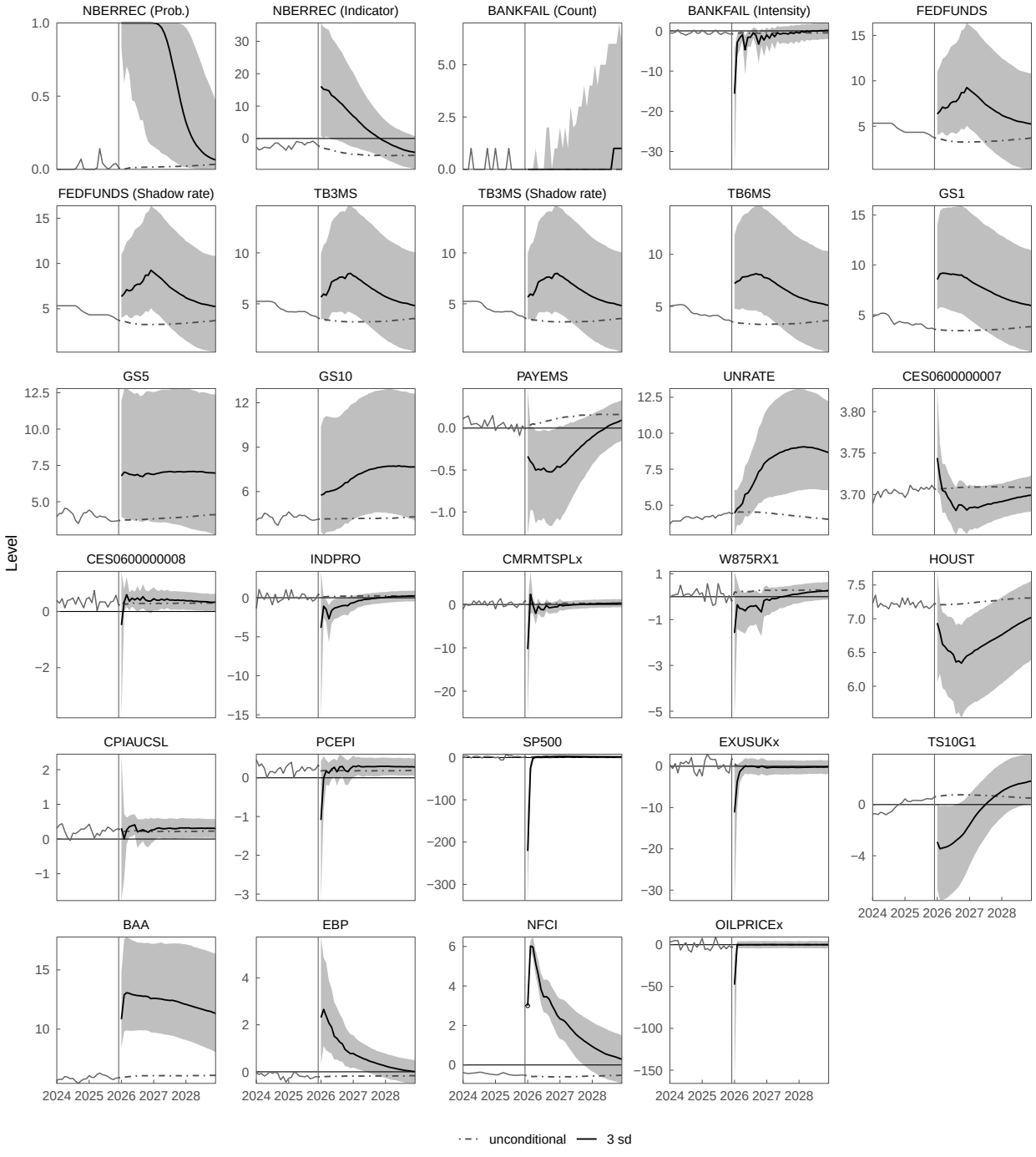

Figure B.10: Growth-at-risk scenario (3 standard deviations): conditional forecasts for all variables.

Notes: See Figure B.8.

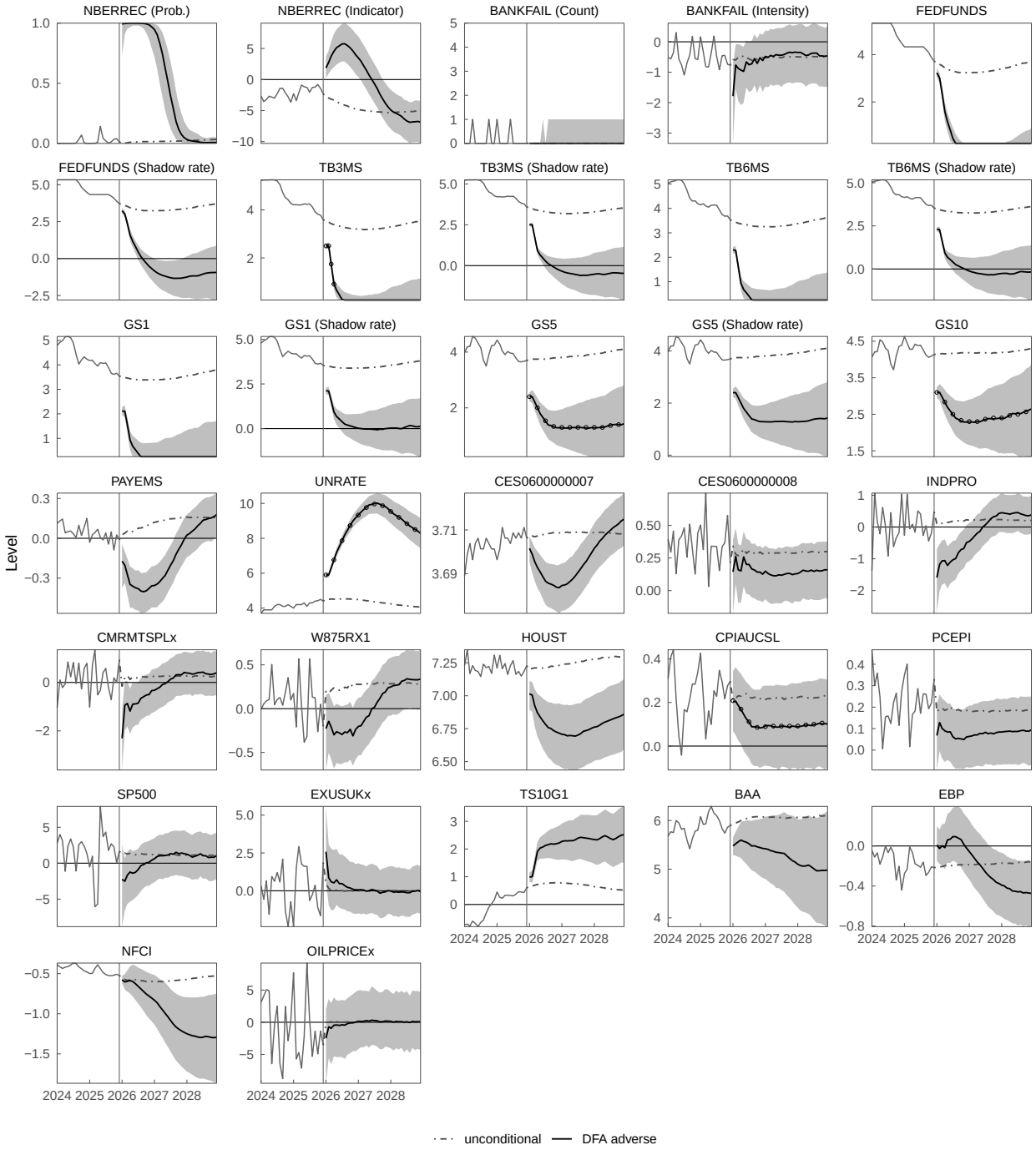

Figure B.11: Stress test scenario (severely adverse): conditional forecasts for all variables.

Notes: See Figure B.8.

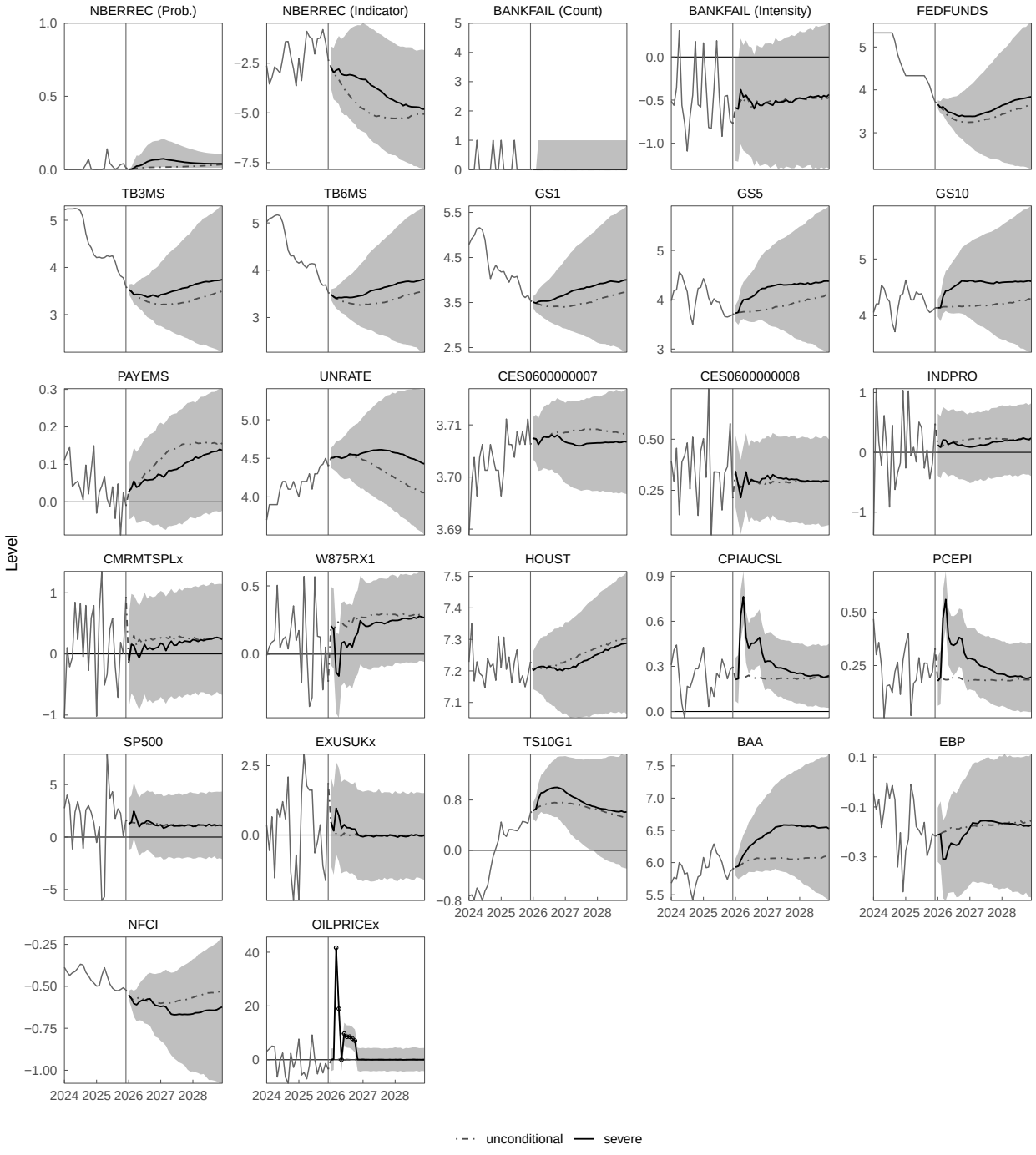

Figure B.12: Oil price scenario (severe): conditional forecasts for all variables.

Notes: See Figure B.8.