# Semiparametric Stochastic Volatility for Large Bayesian VARs

Frank C. Z. Wu*[1], Joshua C. C. Chan[2], and Yong Song[3]

[1]*Department of Statistics, Purdue University*
[2]*Department of Economics, Purdue University*
[3]*Department of Economics, University of Melbourne*

April 2026

**Abstract**

We propose a class of Bayesian VARs in which a Dirichlet process mixture governs the location-scale distribution of the innovations—allowing for asymmetry, heavy tails, and outliers—while the persistent volatility dynamics are specified through a separate parametric covariance process. We develop an efficient collapsed Gibbs sampler and show that the model can approximate a broad class of multivariate densities. Monte Carlo experiments confirm that the approach performs best when extreme shocks are clustered. Empirically, the semiparametric specification delivers the strongest joint density forecasts in a 42-variable macroeconomic application across pre-COVID, COVID, and post-COVID periods. In a financial application, it provides more robust bank connectedness estimates by separating scale-driven volatility spikes from genuine shifts in cross-bank correlation.

Keywords: Dirichlet process mixture, stochastic volatility, Bayesian VARs, heavy tails, density forecasting, financial connectedness

JEL classifications: C14, C53, C55, E17, E37

*We thank Yong Bao, Todd Clark, Mohitosh Kejriwal and Justin Tobias for valuable feedback on earlier versions of the paper. Any remaining errors are solely our own.

# 1 Introduction

Large Bayesian vector autoregressions (VARs) have become a central tool for empirical macroeconomic analysis and forecasting since the seminal contribution of Banbura, Giannone, and Reichlin (2010). Subsequent research has established the importance of time-varying volatility in both small and large systems.[1] As a result, considerable effort has been devoted to developing stochastic volatility specifications that remain computationally feasible in moderate- to high-dimensional VARs. Prominent approaches include common stochastic volatility models (Carriero, Clark, and Marcellino, 2016; Chan, 2020), Cholesky-based stochastic volatility models (Cogley and Sargent, 2005; Carriero, Clark, and Marcellino, 2019), and factor stochastic volatility models (Pitt and Shephard, 1999; Chib, Nardari, and Shephard, 2006; Kastner, 2019). A systematic comparison of these alternatives in large Bayesian VARs is provided by Chan (2023), who highlights the trade-offs between flexibility and computational tractability across different volatility structures. For broader overviews of stochastic volatility specifications in Bayesian VARs, see Clark and Mertens (2023) and Chan (2024).

Despite these advances, most large VARs continue to assume Gaussian innovations. Yet distributional assumptions matter for density forecasting, structural analysis, and the treatment of extreme observations, as large but infrequent shocks appear to play an important role in macroeconomic fluctuations (Cúrdia, Del Negro, and Greenwald, 2014). For small VARs, Cross and Poon (2016) and Chiu, Mumtaz, and Pinter (2017) show that heavy-tailed error distributions can improve forecasting performance even in the presence of stochastic volatility.[2] More generally, Bayesian semiparametric stochastic volatility models based on Dirichlet process mixtures have been shown to capture heavy tails and asymmetries effectively in univariate settings (Jensen and Maheu, 2010, 2014). Recently, Wu (2024) develop an efficient collapsed Gibbs sampler for stochastic volatility models with a Dirichlet process mixture prior, improving computational performance and extending applicability to more general settings.

This paper extends semiparametric stochastic volatility methods to the VAR framework.

[1] See, for example, Clark (2011), D'Agostino, Gambetti, and Giannone (2013), Koop and Korobilis (2013), Clark and Ravazzolo (2015), Cross and Poon (2016), and Chan and Eisenstat (2018).

[2] Acemoglu, Ozdaglar, and Tahbaz-Salehi (2017) provide theoretical motivation for non-Gaussian macroeconomic innovations.

We propose a class of Bayesian VARs in which the innovation distribution is modeled using a Dirichlet process mixture. The central idea is to decouple the modeling of tail behavior from the specification of volatility dynamics. The Dirichlet process mixture governs the location-scale distribution of the innovations, allowing for heavy tails, skewness, and outliers. At the same time, the covariance process $\{\boldsymbol{\Sigma}_t\}$ is specified parametrically and can be chosen according to the empirical context, including common, Cholesky-based, factor, or order-invariant stochastic volatility models. This modular structure allows researchers to tailor the volatility specification while retaining flexible, data-driven modeling of the innovation distribution.

The Dirichlet process mixture provides an infinite mixture-of-Gaussians representation that can approximate a broad class of multivariate densities. Unlike finite mixture models, it does not require the number of components to be fixed in advance; instead, the effective number of clusters is determined by the data. Extending the univariate semiparametric stochastic volatility models of Jensen and Maheu (2010) and Jensen and Maheu (2014) to the VAR setting poses nontrivial computational challenges. We address these by developing an efficient collapsed Gibbs sampler building on Neal (2000) and the computational improvements in Wu (2024), while leveraging existing algorithms for large Bayesian VARs with stochastic volatility.

In a series of Monte Carlo experiments, we clarify when the semiparametric specification is most valuable. The proposed approach delivers its largest gains when extreme shocks are clustered—precisely the environment encountered during episodes such as the COVID-19 pandemic—and remains competitive when innovations follow continuous heavy-tailed distributions.

We apply the framework to two empirical settings. First, in a macroeconomic forecasting application using 42 monthly variables, the semiparametric specification provides the strongest joint density forecast performance across pre-COVID, COVID, and post-COVID subsamples. Second, in a financial connectedness application using daily realized variances for 30 global banks, the transitory scale component absorbs common-scale crisis signals that would otherwise inflate connectedness estimates, giving measures that more accurately reflect genuine shifts in the cross-bank correlation structure.

Our approach relates to and complements several strands of the literature. On one end, Karlsson, Mazur, and Nguyen (2023) and Karlsson and Mazur (2020) develop parametric

heavy-tailed stochastic volatility models for VARs, which are computationally convenient but constrain the innovation distribution to a fixed parametric family. On the other end, fully nonparametric methods such as Clark, Huber, Koop, and Marcellino (2024) and Marcellino and Pfarrhofer (2024) offer greater flexibility but can be difficult to scale. Our semiparametric approach provides a middle ground: the Dirichlet process mixture flexibly learns the innovation distribution from the data, while the parametric volatility component ensures that the framework scales well to high-dimensional settings. We also connect to the growing literature on Dirichlet-process-based clustering in VARs, where Braun (2023) model structural shock distributions nonparametrically and Janssens and Lumsdaine (2025) cluster VAR coefficients using a matrix Dirichlet process prior. Our focus differs in that we target the time-varying innovation distribution while keeping the volatility dynamics parametric.

The framework is most closely related to Huber and Koop (2024), who introduce semiparametric shocks in large VARs via an additive error decomposition in which a Gaussian component captures the bulk of the distribution and a DPM-distributed component acts as a nonparametric correction. Our approach differs in two respects. First, rather than adding a nonparametric residual to Gaussian errors, we apply the Dirichlet process mixture directly to the full innovation distribution through a location-scale representation, so that the entire shape of the conditional distribution—including its location, scale, and tail behavior—is learned from the data. Second, we explicitly decouple the semiparametric innovation specification from the persistent volatility dynamics, treating the covariance process $\{\boldsymbol{\Sigma}_t\}$ as a modular parametric block. This separation allows the framework to be combined with any stochastic volatility specification and provides a natural decomposition of transitory and persistent sources of time-varying risk.

The remainder of the paper is organized as follows. Section 2 introduces the semiparametric VAR framework and discusses the range of compatible volatility specifications. Section 3 develops the collapsed Gibbs sampler. Section 4 uses Monte Carlo experiments to identify when the semiparametric specification delivers gains over parametric alternatives. Section 5 applies the framework to macroeconomic density forecasting with 42 variables, and Section 6 measures financial connectedness among 30 global banks. Section 7 concludes.

# 2 The Econometric Framework

Let $\mathbf{y}_t = (y_{1,t}, \ldots, y_{n,t})'$ be an $n \times 1$ vector of endogenous variables observed for $t = 1, \ldots, T$. We consider the standard VAR($p$):

$$\mathbf{y}_t = \mathbf{A}_1 \mathbf{y}_{t-1} + \cdots + \mathbf{A}_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t, \tag{1}$$

where $\mathbf{A}_1, \ldots, \mathbf{A}_p$ are all $n \times n$ coefficient matrices. In most applications, the innovations $\boldsymbol{\varepsilon}_t$ are assumed to follow a parametric distribution, typically Gaussian or Student-$t$. In contrast, we adopt a semiparametric approach that leaves the shape of the innovation distribution unspecified, while retaining a parametric structure for its second-order dynamics.

## 2.1 Semiparametric Stochastic Volatility

We model the innovations $\boldsymbol{\varepsilon}_t$ using a location-scale mixture of Gaussian distributions. Conditional on a mean vector $\boldsymbol{\mu}_t$, a positive scale $\lambda_t$, and a positive-definite covariance matrix $\boldsymbol{\Sigma}_t$, the innovations satisfy

$$(\boldsymbol{\varepsilon}_t \,|\, \boldsymbol{\mu}_t, \lambda_t, \boldsymbol{\Sigma}_t) \sim \mathcal{N}(\boldsymbol{\mu}_t, \lambda_t \boldsymbol{\Sigma}_t). \tag{2}$$

Because the innovation distribution includes a (potentially time-varying) mean component $\boldsymbol{\mu}_t$, we omit an explicit intercept term in (1) for identification; an intercept would be redundant and not separately identified from $\boldsymbol{\mu}_t$. This specification is used consistently throughout the paper for both exposition and estimation.

To model the distribution of $(\boldsymbol{\mu}_t, \lambda_t)$, we place a Dirichlet process prior on the joint location-scale parameters:

$$\begin{aligned} (\boldsymbol{\mu}_t, \lambda_t \,|\, G) &\sim G, \\ (G \,|\, G_0, \alpha) &\sim \mathcal{DP}(G_0, \alpha), \end{aligned} \tag{3}$$

where $G_0$ is a base measure and $\alpha$ is a concentration parameter. We take $G_0$ to be a multivariate normal-inverse-gamma distribution, $\mathcal{NIG}(\boldsymbol{\mu}_0, \mathbf{V}_{\boldsymbol{\mu}_0}, \nu_0, S_0)$.

The proposed approach in (2)–(3) combines both a parametric and a nonparametric components: the evolution of $\{\mathbf{\Sigma}_t\}_{t=1}^T$ is modeled parametrically, whereas the nonparametric component is constructed using the Dirichlet process prior that mixes potentially an infinite number of Gaussian distributions. To clarify the distributional flexibility implied by (3), consider the case in which $\mathbf{\Sigma}_t = \mathbf{\Sigma}$ is fixed over time. The Dirichlet process prior specified in (3) is capable of approximating virtually any continuous probability density function (pdf) in practice arbitrarily well. In particular, we consider a class of continuous pdfs that has the property of 'vanishing at infinity.'

**Assumption 1.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous probability density function that vanishes at infinity, i.e., for every $\varepsilon > 0$, there exists $M \in \mathbb{R}$ such that for all $\|\mathbf{x}\| > M$, $f(\mathbf{x}) < \varepsilon$. We denote this class of pdfs as $C_0(\mathbb{R}^n)$.*

Vanishing at infinity is a regularity condition that rules out some pathological cases.[3] A very wide class of pdfs satisfies Assumption 1. For example, any continuous pdf with bounded support is a member of $C_0(\mathbb{R}^n)$. In addition, if the tails of $f$ decay at a polynomial rate or faster, then $f$ is also a member of $C_0(\mathbb{R}^n)$. Hence, the pdfs of common distributions such as Gaussian, Laplace, Gamma, log-normal, Student-$t$, skew $t$, $\alpha$-stable all satisfy Assumption 1.

There are also a range of sufficient conditions that imply vanishing at infinity. For example, one can show that if a pdf $f$ is uniformly continuous, it vanishes at infinity. This fact implies that if $f$ is Lipschitz or has a bounded gradient (the former implies uniform continuity and the latter implies Lipschitz), then it also vanishes at infinity. The proofs of these claims are provided in Online Appendix A.

To formally state the claim that our construction using the Dirichlet process prior can approximate any $f \in C_0(\mathbb{R}^n)$ uniformly well, let $\phi(\mathbf{x}; \mathbf{m}, \mathbf{S})$ denote the $n$-dimensional Gaussian density with mean vector $\mathbf{m}$ and covariance matrix $\mathbf{S}$. Then, the setup in (3) admits the stick-breaking representation (Sethuraman, 1994):

$$G = \sum_{k=1}^{\infty} \pi_k \, \delta_{\boldsymbol{\theta}_k}, \quad \boldsymbol{\theta}_k = (\boldsymbol{\mu}_k, \lambda_k) \sim G_0, \quad \pi_k = v_k \prod_{j<k} (1 - v_j), \quad v_k \sim \text{Beta}(1, \alpha), \tag{4}$$

[3] If a pdf does not vanish at infinity, then it is possible to find infinitely many regions far away from the origin where the density jumps back up to a fixed level. To compensate, these regions must be infinitesimally narrow so that the pdf integrates to one. This is possible mathematically, but it is highly nonphysical.

where $\delta_{\boldsymbol{\theta}_k}$ is the Dirac measure centered at $\boldsymbol{\theta}_k$ and $\pi_k$ is the mixture weight for the $k$-th component. Substituting this representation into the mixture model in (2) yields an infinite Gaussian mixture representation:

$$p(\boldsymbol{\varepsilon} \mid \boldsymbol{\Sigma}) = \int \phi(\boldsymbol{\varepsilon}; \boldsymbol{\mu}, \lambda\boldsymbol{\Sigma})\, G(\mathrm{d}\boldsymbol{\mu}, \mathrm{d}\lambda) = \sum_{k=1}^{\infty} \pi_k\, \phi(\boldsymbol{\varepsilon}; \boldsymbol{\mu}_k, \lambda_k\boldsymbol{\Sigma}), \quad (\boldsymbol{\mu}_k, \lambda_k) \sim G_0.$$

This expresses the Dirichlet process mixture as an infinite mixture of Gaussian components, where the component parameters $(\boldsymbol{\mu}_k, \lambda_k)$ are drawn from the base measure $G_0$. This infinite mixture of Gaussian distributions is capable of uniformly approximating any pdf $f \in C_0(\mathbb{R}^n)$.

**Theorem 1.** *Let $f$ be a continuous probability density function that satisfies Assumption 1. Fix a positive definite matrix $\boldsymbol{\Sigma}$. Then, for every $\varepsilon > 0$, there exist a nonnegative integer $K$, $n$-vectors $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K$, positive scalars $\lambda_1, \ldots, \lambda_K$ and nonnegative weights $\pi_k$ with $\sum_{k=1}^{K} \pi_k = 1$ such that*

$$\sup_{\mathbf{x} \in \mathbb{R}^n} \left| f(\mathbf{x}) - \sum_{k=1}^{K} \pi_k\, \phi(\mathbf{x}; \boldsymbol{\mu}_k, \lambda_k\boldsymbol{\Sigma}) \right| < \varepsilon.$$

The proof involves verifying the premise of Theorem 5 in Nguyen, Nguyen, Chamroukhi, and McLachlan (2020); the details are given in Online Appendix A.

In particular, the proposed framework naturally allows for asymmetric innovation distributions. Because the Dirichlet process mixture combines Gaussian components with potentially nonzero mean vectors $\boldsymbol{\mu}_t$, the marginal distribution of $\boldsymbol{\varepsilon}_t$—after integrating out $(\boldsymbol{\mu}_t, \lambda_t)$—need not be symmetric around 0. This flexibility allows the model to capture both heavy left and heavy right tails in the marginal innovation distribution. As we document in the macroeconomic application below, the data indeed exhibit evidence of asymmetric tail behavior, with some variables displaying heavier right tails while others exhibit heavier left tails.

The parametric component involves the stochastic process governing $\{\boldsymbol{\Sigma}_t\}_{t=1}^{T}$; it can be specified in a variety of ways depending on the desired balance between parsimony and flexibility. The key feature of the framework is that the Dirichlet process mixture governing the innovation distribution is orthogonal to this choice. Any positive-definite

covariance process $\boldsymbol{\Sigma}_t$ can be embedded within the semiparametric structure in (2)–(3).

A natural starting point is the common stochastic volatility (e.g., Carriero, Clark, and Marcellino, 2016; Chan, 2020), denoted as CSV, where the entire error covariance matrix is scaled by a single latent volatility factor,

$$\boldsymbol{\Sigma}_t = \mathrm{e}^{h_t}\boldsymbol{\Sigma},$$

and the log-volatility follows a mean-zero AR(1) process,

$$h_t = \phi h_{t-1} + u_t, \quad |\phi| < 1.$$

This specification efficiently captures broad comovement in volatilities across macroeconomic series and permits fast estimation of order $O(n^3)$. The common stochastic volatility factor is typically interpreted as a measure of aggregate macroeconomic uncertainty, reflecting economy-wide fluctuations in risk. This interpretation is consistent with Carriero, Clark, and Marcellino (2018), who show that common volatility factors in large VARs provide informative summaries of macroeconomic uncertainty rather than variable-specific sources of risk.

A more flexible model is the Cholesky stochastic volatility introduced by Cogley and Sargent (2005) and further developed by Carriero, Clark, and Marcellino (2019). In this case, the covariance matrix is constructed via a modified Cholesky decomposition:

$$\boldsymbol{\Sigma}_t = \mathbf{B}_0^{-1}\mathbf{D}_t(\mathbf{B}_0^{-1})^\intercal, \quad \mathbf{D}_t = \mathrm{diag}(\mathrm{e}^{h_{1,t}}, \ldots, \mathrm{e}^{h_{n,t}}),$$

where $\mathbf{B}_0$ is a unit lower triangular matrix, and each $h_{i,t}$ following an independent AR(1) process. This structure allows for heterogeneous volatility dynamics across variables and time-varying correlations, albeit at the cost of higher computational complexity of order $O(n^4)$.

Cholesky-based specifications are inherently order-dependent, as the implied covariance dynamics depend on the chosen ordering of the variables. Several recent order-invariant stochastic volatility models address this limitation. For example, Arias, Rubio-Ramirez, and Shin (2023), Chan, Koop, and Yu (2024), and Wu and Koop (2025) construct $\boldsymbol{\Sigma}_t$ using dense impact matrices or eigendecompositions rather than triangular factorizations.

These order-invariant stochastic volatility (OISV) specifications retain tractable posterior simulation while avoiding dependence on variable ordering, and they can be accommodated naturally within the modular framework considered in this paper.

Finally, factor stochastic volatility (FSV) models (Pitt and Shephard, 1999; Aguilar and West, 2000; Chib, Nardari, and Shephard, 2006; Kastner, 2019) provide a dimension-reduction approach,

$$\boldsymbol{\Sigma}_t = \mathbf{L}\mathbf{G}_t\mathbf{L}^\intercal + \mathbf{D}_t,$$

where $\mathbf{L}$ is an $n \times r$ matrix of factor loadings, $\mathbf{G}_t = \text{diag}(\text{e}^{h_{n+1,t}}, \ldots, \text{e}^{h_{n+r,t}})$, and $\mathbf{D}_t = \text{diag}(\text{e}^{h_{1,t}}, \ldots, \text{e}^{h_{n,t}})$; each log-volatility $h_{i,t}$ again follows an AR(1) process. This formulation substantially reduces dimensionality and is particularly suited to high-dimensional VARs in which volatility comovement is driven by a small number of latent factors. This factor stochastic volatility has been incorporated into large VARs in a range of applications, such as Kastner and Huber (2020), Chan (2023) and Hauzenberger, Huber, Marcellino, and Petz (2025).

In what follows, we use the label DPM-X to denote a stochastic volatility model X in which the parametric innovation distribution is replaced by a Dirichlet process mixture. For example, DPM-CSV refers to the common stochastic volatility model with semiparametric innovations, whereas DPM-OISV refers to the order-invariant stochastic volatility model of Chan, Koop, and Yu (2024) combined with the same Dirichlet process mixture structure.

## 2.2 Relation to Existing Heavy-Tailed Stochastic Volatility

A growing literature extends Bayesian VARs by introducing heavy-tailed innovations and outliers in combination with stochastic volatility. From the perspective of the framework developed above, these models can be understood as imposing specific parametric or discrete restrictions on the innovation distribution, while retaining similar volatility dynamics.

An example is the Bayesian VAR with common stochastic volatility and Student-$t$ inno-

vations proposed by Chan (2020). In that model,

$$(\boldsymbol{\varepsilon}_t \mid \lambda_t, h_t, \boldsymbol{\Sigma}) \sim \mathcal{N}(\mathbf{0}, \lambda_t \mathrm{e}^{h_t} \boldsymbol{\Sigma}),$$

where $\lambda_t$ follows an independent inverse-gamma distribution. The Student-$t$ specification can be viewed as a *parametric scale mixture of Gaussians*, with a fixed mixing distribution that governs tail thickness. While this approach introduces heavy tails in a parsimonious manner, the shape of the innovation distribution—such as symmetry and tail behavior—is fully determined ex ante.

Related alternatives include the SVO and SVo models proposed in Carriero, Clark, Marcellino, and Mertens (2024), which extend the outlier framework of Stock and Watson (2016) to multivariate VARs with stochastic volatility. These models introduce latent multiplicative outlier states that inflate the conditional variance during extreme observations. In SVO, the outlier state is variable-specific, whereas in SVo it is common across variables. The latter feature makes SVo particularly attractive in large-scale applications, as it permits estimation in systems with a large cross-sectional dimension while retaining computational tractability. In both cases, the outlier states follow a discrete mixture distribution and are assumed to be independent across time.

Carriero, Clark, Marcellino, and Mertens (2024) further consider an extension denoted SVO-$t$, which augments the SVO outlier-state specification with $t$-distributed innovations. This model introduces additional parametric flexibility in the tails of the innovation distribution and is closely related to SV-$t$ specifications that replace Gaussian errors with a Student-$t$ distribution. While SVO, SVo, and SVO-$t$ allow for occasional large shocks through discrete or parametric tail adjustments, the functional form of the innovation distribution remains fixed *ex ante*.

The semiparametric model proposed in this paper departs from these approaches in two key respects. First, rather than imposing a parametric or discrete mixing distribution, we allow the location-scale parameters $(\boldsymbol{\mu}_t, \lambda_t)$ to be drawn from an unknown distribution governed by a Dirichlet process prior. This induces a continuous mixture of Gaussian components and allows the shape of the innovation distribution—including skewness, multimodality, and tail behavior—to be learned flexibly from the data. Second, although the innovation distribution is common across variables, the Dirichlet process induces *data-driven clustering over time*, enabling observations with similar distributional character-

istics to share information, in contrast to models in which outlier states are independent across periods.

We also note that unconditional skewness can arise in fully parametric settings, for example in stochastic volatility models with leverage effects or SV-in-mean specifications. In these cases, asymmetry is generated through dynamic dependence between shocks and volatility. By contrast, our approach captures asymmetric and heavy-tailed innovations directly through a flexible semiparametric error distribution, without imposing a specific parametric structure on this dependence.

It is important to emphasize, however, that the Dirichlet process in our model does not introduce cross-sectional heterogeneity: all innovations are drawn from the same mixture distribution, both *a priori* and *a posteriori*. Consequently, increasing the cross-sectional dimension $n$ does not expand the flexibility of the distributional class. Instead, larger systems provide more information about the common innovation distribution, leading to more precise inference, in a manner analogous to parametric models such as the Student-$t$, but without fixing the form of the mixing distribution in advance.

More broadly, SV-$t$, SVO, and SVo can all be embedded within the proposed framework, as they correspond to particular parametric specifications of the covariance process $\{\boldsymbol{\Sigma}_t\}$. SV-$t$ arises from a parametric scale mixture with a common volatility factor, while SVO and SVo correspond to Cholesky-based stochastic volatility models with variable-specific or common outlier states. In all cases, the Dirichlet process mixture governs a common location-scale innovation distribution, while cross-sectional dependence and heterogeneity are captured parametrically through $\{\boldsymbol{\Sigma}_t\}$.

By contrast, extending the Dirichlet process itself to introduce nonparametric cross-sectional heterogeneity—so that different variables are associated with different mixture distributions—poses a substantially more difficult problem. Such an extension would forgo the conjugate structure exploited in this paper and preclude the use of collapsed samplers, leading to a significant increase in computational complexity, particularly in large systems. For this reason, we focus on a scalable semiparametric framework in which flexibility in tail behavior and outliers is introduced through a common Dirichlet process mixture, while cross-sectional structure is modeled parametrically via $\{\boldsymbol{\Sigma}_t\}$.

# 3 Bayesian Estimation

The proposed Dirichlet process mixture VAR with stochastic volatility is estimated using Markov chain Monte Carlo methods. Below we outline the main components of the posterior sampler, with particular emphasis on the steps associated with the Dirichlet process mixture. Sampling the VAR coefficient $\mathbf{A} = (\mathbf{A}_1, \ldots, \mathbf{A}_p)$ and the covariance matrices $\boldsymbol{\Sigma}_{1:T} = (\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_T)$ follows standard procedures for Bayesian VARs with stochastic volatility and is therefore not described in detail here. In particular, the volatility states are sampled using the well-established algorithms for VAR models with stochastic volatility, implemented following the corrected sampling order discussed in Del Negro and Primiceri (2015). We refer readers to Carriero, Clark, and Marcellino (2016, 2019), Kastner and Huber (2020), and Carriero, Chan, Clark, and Marcellino (2022) for comprehensive treatments; in particular, Appendix A of Chan (2023) provides estimation details for large VARs under a range of commonly used stochastic volatility specifications. More broadly, our implementation builds on efficient simulation methods for state space models such as those developed in Chan and Jeliazkov (2009), which facilitate stable and computationally efficient posterior sampling in models with latent volatility states.

To facilitate posterior simulation of the nonparametric component, we augment the model with the latent cluster indicators $z_t, t = 1, \ldots, T$. Under the stick-breaking representation of the Dirichlet process in (4), each observation is associated with one mixture component through

$$\begin{aligned} z_t &\sim \text{Discrete}(\pi_1, \pi_2, \ldots, \pi_K), \\ (\mathbf{y}_t \mid z_t, \boldsymbol{\theta}_{1:K}, \boldsymbol{\Sigma}_t) &\sim \mathcal{N}(\boldsymbol{\mu}_{z_t}, \lambda_{z_t} \boldsymbol{\Sigma}_t), \end{aligned}$$

where $K$ denotes the (random) number of active clusters and $\boldsymbol{\theta}_{1:K} = (\boldsymbol{\theta}_1^\intercal, \ldots, \boldsymbol{\theta}_K^\intercal)^\intercal$ collects the cluster-specific parameters. Since the Dirichlet process defines an infinite mixture, $K$ is not fixed a priori and may vary across MCMC iterations.

## 3.1 Overview of the MCMC Algorithm

For clarity, we first provide a high-level outline of the MCMC algorithm. Each iteration proceeds as follows:

1. *Sample VAR coefficients* $\mathbf{A}$. Conditional on the stochastic volatility process $\boldsymbol{\Sigma}_{1:T}$ and the current state of the Dirichlet process mixture, draw $\mathbf{A}$ using standard methods for Bayesian VARs with stochastic volatility.

2. *Sample stochastic volatility* $\boldsymbol{\Sigma}_{1:T}$. Given $\mathbf{A}$ and the current mixture indicators and parameters, update the latent volatility states according to the chosen parametric stochastic volatility specification (e.g., common SV, Cholesky SV, or factor SV).

3. *Sample cluster assignments* $\mathbf{z} = (z_1, \ldots, z_T)$. Compute the VAR residuals $\boldsymbol{\varepsilon}_t = \mathbf{y}_t - \sum_{\ell=1}^{p} \mathbf{A}_\ell \mathbf{y}_{t-\ell}$ for $t = 1, \ldots, T$. Integrating out the mixture weights and cluster-specific parameters, sample each $z_t$ sequentially using a collapsed Gibbs sampler based on posterior and prior predictive densities.

4. *Update the number of clusters* $K$. After updating $\mathbf{z}$, determine the number of active clusters by counting the distinct labels.

5. *Sample the concentration parameter* $\alpha$. Given the current number of clusters $K$, update $\alpha$ using the auxiliary-variable method of Song (2014).

6. *Sample cluster-specific parameters* $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_K$. Conditional on $\mathbf{z}$ and $\boldsymbol{\Sigma}_{1:T}$, draw each $\boldsymbol{\theta}_k = (\boldsymbol{\mu}_k, \lambda_k)$ from its conjugate normal–inverse–gamma posterior distribution.

These steps are iterated for a fixed number of MCMC iterations, discarding an initial burn-in period. The use of conjugate priors and collapsed updates for the Dirichlet process mixture provides an efficient sampler that mixes well even in large VAR systems. We now describe in detail the collapsed Gibbs updates for the cluster assignments $\mathbf{z}$, as well as the sampling of $\alpha$ and $K$.

## 3.2 Sampling Cluster Assignments z, $\alpha$ and $K$

To efficiently sample the cluster indicators $\mathbf{z} = (z_1, \ldots, z_T)^\intercal$, we use the collapsed Gibbs sampling algorithm (Algorithm 3) proposed in Neal (2000). More specifically, we analytically integrate out the cluster-specific parameters $\boldsymbol{\theta}_{1:K}$ and the mixture weights $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_K)^\intercal$, and sample the cluster indicators unconditional on $\boldsymbol{\theta}_{1:K}$ and $\boldsymbol{\pi}$. Sudderth (2006) shows that due to the Rao-Blackwellized step of integrating out the cluster-specific parameters, this sampling scheme converges quickly, typically in a few

hundred of iterations. Below we provide details on sampling the cluster assignments $\mathbf{z}$, as well as drawing the concentration parameter $\alpha$, the number of clusters $K$ and the cluster parameters $\boldsymbol{\theta}$.

Given the data $\mathbf{y}_t$ and the VAR coefficients $\mathbf{A}$, the residual $\boldsymbol{\varepsilon}_t = \mathbf{y}_t - \mathbf{A}_1\mathbf{y}_{t-1} - \cdots - \mathbf{A}_p\mathbf{y}_{t-p}$ can be readily computed. Stack $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1^\intercal, \ldots, \boldsymbol{\varepsilon}_T^\intercal)^\intercal$. The conditional posterior density of $z_t$ given other cluster indicators $\mathbf{z}_{\setminus t} = (z_1, \ldots, z_{t-1}, z_{t+1}, \ldots, z_T)^\intercal$ and $\boldsymbol{\Sigma}_{1:T}$, but unconditional on $\boldsymbol{\theta}_{1:K}$ and $\boldsymbol{\pi}$, can be written as

$$p(z_t = k \,|\, \mathbf{z}_{\setminus t}, \boldsymbol{\varepsilon}, \boldsymbol{\Sigma}_{1:T}) \propto p(z_t = k \,|\, \mathbf{z}_{\setminus t}) \times p(\boldsymbol{\varepsilon}_t \,|\, \boldsymbol{\Sigma}_{1:T}, \{\boldsymbol{\varepsilon}_s : z_s = k, s \neq t\}),$$

where the second density on the right-hand side of the equation depends only on $\{\boldsymbol{\varepsilon}_s : z_s = k, s \neq t\}$, the set of residuals (excluding $\boldsymbol{\varepsilon}_t$) that belong to the same cluster $k$. More specifically, it is the posterior predictive density evaluated at $\boldsymbol{\varepsilon}_t$, obtained by integrating out $\boldsymbol{\theta}_k = (\boldsymbol{\mu}_k, \lambda_k)$ with respect to its posterior density:

$$p(\boldsymbol{\varepsilon}_t \,|\, \boldsymbol{\Sigma}_{1:T}, \{\boldsymbol{\varepsilon}_s : z_s = k, s \neq t\}) = \int \phi(\boldsymbol{\varepsilon}_t; \boldsymbol{\mu}_k, \lambda_k\boldsymbol{\Sigma}_t) p(\boldsymbol{\mu}_k, \lambda_k \,|\, \{\boldsymbol{\varepsilon}_s : z_s = k, s \neq t\}, \boldsymbol{\Sigma}_{1:T}) \mathrm{d}(\boldsymbol{\mu}_k, \lambda_k).$$

Note that the posterior density of $\boldsymbol{\theta}_k$ is computed using only the set of residuals $\{\boldsymbol{\varepsilon}_s : z_s = k, s \neq t\}$. Owing to the conjugate base measure $G_0$, the associated posterior predictive density can be obtained analytically. We refer readers to Online Appendix B for technical details.[4]

For notational convenience, we denote this posterior predictive density for the $k$-th cluster by

$$f_k(\cdot) \equiv p(\cdot \,|\, \boldsymbol{\Sigma}_{1:T}, \{\boldsymbol{\varepsilon}_s : z_s = k, s \neq t\}).$$

In addition, $f_{\bar{K}}(\cdot)$ denotes the *prior predictive density*, obtained by integrating the likelihood with respect to the base measure $G_0$. Here $\bar{K}$ represents a new cluster containing no observations.

Putting it all together, $z_t$ follows a $(K+1)$-dimensional multinomial distribution with

[4]For the linear regression with the normal-inverse-gamma prior, Zellner (1971) shows that the posterior predictive density is a Student-$t$. For the homoskedastic VARs with the normal-inverse-Wishart prior, the posterior predictive density is a multivariate Student-$t$ (see, e.g., Karlsson, 2013). In our setting with time-varying covariance matrices $\boldsymbol{\Sigma}_{1:T}$, we show that the posterior predictive density also takes the form of a multivariate Student-$t$ density.

$k = 1, \ldots, K$, and $\bar{K}$. Specifically, $z_t$ is distributed as:

$$p(z_t = k \mid \mathbf{z}_{\setminus t}, \boldsymbol{\varepsilon}, \boldsymbol{\Sigma}_{1:T}) = \begin{cases} \frac{T_k^{\setminus t}}{c} f_k(\boldsymbol{\varepsilon}_t), & k = 1, \ldots, K, \\ \frac{\alpha}{c} f_{\bar{K}}(\boldsymbol{\varepsilon}_t), & k = \bar{K}, \end{cases}$$

where $c = \alpha f_{\bar{K}}(\boldsymbol{\varepsilon}_t) + \sum_{k=1}^{K} T_k^{\setminus t} f_k(\boldsymbol{\varepsilon}_t)$ and $T_k^{\setminus t}$ is the number of observations in cluster $k$, excluding $\boldsymbol{\varepsilon}_t$.

Next, given the draw $z_t$, we update the value of $K$. If $z_t = \bar{K}$, assign it a new, unique cluster label. Update $\mathbf{z}$ with the new value of $z_t$. Then, determine the total number of active clusters $K$ by counting the unique labels in $\mathbf{z}$. The total number of clusters $K$ changes in two scenarios: $K$ increases by 1 if a new cluster is drawn, while the previous cluster of the residual $\boldsymbol{\varepsilon}_t$ remains non-empty. In the second scenario, the cluster label of $\boldsymbol{\varepsilon}_t$ changes and its previous cluster becomes empty (i.e., $\boldsymbol{\varepsilon}_t$ was the only member of the cluster), then $K$ decreases by 1.[5]

To sample the concentration parameter $\alpha$, we follow the method proposed in Song (2014), which only uses one auxiliary variable and is more efficient than the two-step procedure developed in Escobar and West (1995). We assume $\alpha$ has a gamma prior $\alpha \sim \mathcal{G}(a_0, b_0)$.[6] Given this gamma prior, we first sample an auxiliary variable $\xi$ from the beta distribution $\xi \sim \mathcal{B}(\alpha, T)$, followed by sampling $\alpha$ from $\mathcal{G}(a_0 + K, b_0 - \log \xi)$.

## 3.3 Sampling Cluster Parameters $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_K$

For each $\boldsymbol{\theta}_k, k = 1, \ldots, K$, its conditional posterior distribution depends only on the observations from its own cluster. Given the cluster assignments $\mathbf{z}$, we can therefore partition the data into $K$ clusters and sample $\boldsymbol{\theta}_k$ sequentially.

The prior on $\boldsymbol{\theta}_k = (\boldsymbol{\mu}_k, \lambda_k)$ is the normal-inverse-gamma distribution $\mathcal{NIG}(\boldsymbol{\mu}_0, \mathbf{V}_{\boldsymbol{\mu}_0}, \nu_0, S_0)$, which is conjugate to the likelihood. Hence, the conditional posterior is also a normal-

[5] Here, we sample $z_t$ sequentially to mitigate the issue of label switching. What matters for our purpose is whether two observations belong to the same cluster, rather than the values of their labels.

[6] We set the hyperparameters $a_0$ and $b_0$ so that they reflect the expected number of clusters. For example, for monthly datasets, one might expect a new cluster (e.g., outliers) occurs roughly once every 4 years. In that case, the expected number of clusters is approximately $T/48$. Accordingly, one might set $a_0 = 20$ and $b_0 = 8$.

inverse-gamma distribution:

$$(\boldsymbol{\theta}_k \,|\, \boldsymbol{\Sigma}_{1:T}, \mathbf{z}, \{\boldsymbol{\varepsilon}_t : z_t = k\}) \sim \mathcal{NIG}(\boldsymbol{\mu}_k, \mathbf{V}_{\boldsymbol{\mu}_k}, \nu_k, S_k),$$

where $\nu_k = \nu_0 + nT_k/2$ with $T_k$ representing the size of the $k$-th cluster,

$$\boldsymbol{\mu}_k = \mathbf{V}_{\boldsymbol{\mu}_k}\left(\mathbf{V}_{\boldsymbol{\mu}_0}^{-1}\boldsymbol{\mu}_0 + \sum_{t:z_t=k}\boldsymbol{\Sigma}_t^{-1}\boldsymbol{\varepsilon}_t\right), \quad \mathbf{V}_{\boldsymbol{\mu}_k} = \left(\mathbf{V}_{\boldsymbol{\mu}_0}^{-1} + \sum_{t:z_t=k}\boldsymbol{\Sigma}_t^{-1}\right)^{-1},$$
$$S_k = S_0 + \frac{1}{2}\boldsymbol{\mu}_0^{\intercal}\mathbf{V}_{\boldsymbol{\mu}_0}^{-1}\boldsymbol{\mu}_0 + \frac{1}{2}\sum_{t:z_t=k}\boldsymbol{\varepsilon}_t^{\intercal}\boldsymbol{\Sigma}_t^{-1}\boldsymbol{\varepsilon}_t - \frac{1}{2}\boldsymbol{\mu}_k^{\intercal}\mathbf{V}_{\boldsymbol{\mu}_k}^{-1}\boldsymbol{\mu}_k.$$

The notation $\sum_{t:z_t=k}$ indicates that the summation is over all the time periods $t = 1, \ldots, T$ that are assigned to cluster $k$. The details of the derivations of the conditional posterior distribution are provided in Online Appendix B.

# 4 Monte Carlo Experiments

This section presents Monte Carlo experiments designed to clarify when flexible semi-parametric modeling of the innovation distribution delivers gains relative to parametric alternatives. The simulations are deliberately constructed to isolate three empirically relevant features of macroeconomic innovations: heavy tails, discrete outliers, and clustered extreme episodes (pandemic-style shocks).

## 4.1 Simulation Design

All simulations are based on a fixed $n$-variable VAR($p$), where the coefficient matrices are set equal to the posterior means from the estimated 16-variable SVo benchmark model in Section 5, with $p = 12$. Fixing the VAR dynamics in this way serves two purposes. First, it calibrates persistence and cross-variable propagation to match the empirical application. Second, it ensures that differences across models in the Monte Carlo study are driven primarily by the innovation specification rather than by differences in the conditional mean.

Innovations are generated according to a common stochastic volatility structure augmented by a transitory scale component:

$$(\boldsymbol{\varepsilon}_t \mid h_t, \lambda_t) \sim \mathcal{N}(\mathbf{0}, \lambda_t \mathrm{e}^{h_t} \boldsymbol{\Sigma}),$$

where $\boldsymbol{\Sigma}$ is fixed at the estimated time-invariant covariance matrix from the same 16-variable SVo benchmark. The latent log-volatility factor evolves as

$$h_t = \phi h_{t-1} + u_t, \quad u_t \sim \mathcal{N}(0, \sigma_h^2).$$

This design mirrors the decomposition emphasized in the paper: $\mathrm{e}^{h_t}$ captures persistent movements in volatility, while $\lambda_t$ captures transitory departures associated with outliers and tail events.

We consider two sample sizes, $T \in \{300, 800\}$, chosen to span settings that are representative of moderately sized macroeconomic datasets and longer panels (e.g., quarterly versus monthly data after the post-war period, depending on the application and transformations). For each design we run $R = 20$ independent Monte Carlo replications.

For each simulated dataset, we estimate three models: (i) DPM-CSV (a proposed semi-parametric innovation model), (ii) SVo (the discrete outlier-state specification), and (iii) SV-$t$ (common stochastic volatility with Student-$t$ innovations).

We consider three DGPs that differ only in the transitory scale process $\{\lambda_t\}$, holding the VAR dynamics and the persistent volatility process fixed.

*DGP1: Fat-tailed mixture.* To generate heavy tails without discrete outlier regimes, $\lambda_t$ is drawn iid from a two-component continuous mixture:

$$\lambda_t \sim \begin{cases} \mathcal{IG}(\nu_1/2, \nu_1/2), & \text{with probability } 1-q, \\ \kappa \cdot \mathcal{IG}(\nu_2/2, \nu_2/2), & \text{with probability } q, \end{cases}$$

with $q = 0.2$, $\nu_1 = 10$, $\nu_2 = 3$ and $\kappa = 5$. This DGP provides a setting in which a parametric SV-$t$ model is a natural competitor, while allowing the DPM-CSV model to adapt flexibly to tail behavior.

*DGP2: Discrete outliers.* To mimic the outlier-state mechanism used in SVo, the scale

process follows

$$\lambda_t = \begin{cases} 1, & \text{with probability } 1-q, \\ o_t, & \text{with probability } q, \quad o_t \sim \mathcal{U}\{2,\dots,20\}, \end{cases}$$

with $q = 0.05$. This DGP favors discrete outlier-state models and serves as a benchmark for whether DPM-CSV can recover outlier magnitudes when the truth is inherently discrete.

*DGP3: Clustered outlier block.* To mimic pandemic-style episodes, we introduce a short contiguous block of extreme observations:

$$\lambda_t = \begin{cases} 1, & t \notin \mathcal{T}_{\text{out}}, \\ \lambda^\star, & t \in \mathcal{T}_{\text{out}}, \end{cases}$$

where $\mathcal{T}_{\text{out}} = \{t_0, \dots, t_0 + L - 1\}$, $t_0 = \lfloor 0.7T \rfloor$, $L = 6$, and $\lambda^\star = 25$. This DGP is designed to isolate the mechanism emphasized in the empirical application: outliers are not merely rare, but clustered in time, so that the model must accommodate a short-lived regime of extreme volatility.

We evaluate performance by the recovery of latent volatility states $h_t$ and $\lambda_t$. In particular, for each replication, we compute

$$\text{MSE}_h = \frac{1}{RT} \sum_{r=1}^{R} \sum_{t=1}^{T} \big(\widehat{h}_t^{(r)} - h_t^{(r)}\big)^2, \qquad \text{MSE}_\lambda = \frac{1}{RT} \sum_{r=1}^{R} \sum_{t=1}^{T} \big(\widehat{\lambda}_t^{(r)} - \lambda_t^{(r)}\big)^2.$$

## 4.2 Simulation Results

Tables 1 and 2 report relative mean squared errors (MSEs) for the transitory scale $\lambda_t$ and the persistent volatility component $h_t$, normalized so that the DPM-CSV model equals one in each design.

Under DGP1, $\lambda_t$ follows a two-component continuous mixture. As expected, the SV-$t$ model performs best in large samples, since the true DGP closely aligns with a parametric $t$-mixture representation. For $T = 800$, SV-$t$ achieves the lowest MSE for both $\lambda_t$ and $h_t$, while DPM-CSV performs competitively. In smaller samples ($T = 300$), performance

differences are more muted, reflecting the greater flexibility of the semiparametric approach. Notably, DPM-CSV performs uniformly better than SVo under this DGP, as the discrete outlier structure of SVo is not well aligned with the continuous heavy-tailed DGP. These results confirm that when the DGP is well captured by a parametric heavy-tailed model, SV-$t$ can dominate asymptotically, while DPM-CSV remains robust and avoids the misspecification costs incurred by discrete outlier models.

Table 1: Relative MSE for $\lambda_t$ (DPM-CSV normalized to 1)

| | $T = 300$ | | | $T = 800$ | | |
|---|---|---|---|---|---|---|
| DGP | DPM | SVo | SV-$t$ | DPM | SVo | SV-$t$ |
| 1 (fat tails) | 1.000 | 2.267 | 1.851 | 1.000 | 1.065 | 0.327 |
| 2 (discrete) | 1.000 | 0.250 | 1.258 | 1.000 | 0.351 | 0.933 |
| 3 (clustered) | 1.000 | 1.878 | 2.582 | 1.000 | 2.125 | 2.314 |

Table 2: Relative MSE for $h_t$ (DPM-CSV normalized to 1)

| | $T = 300$ | | | $T = 800$ | | |
|---|---|---|---|---|---|---|
| DGP | DPM | SVo | SV-$t$ | DPM | SVo | SV-$t$ |
| 1 (fat tails) | 1.000 | 1.306 | 0.640 | 1.000 | 1.368 | 1.002 |
| 2 (discrete) | 1.000 | 1.414 | 1.170 | 1.000 | 1.429 | 1.020 |
| 3 (clustered) | 1.000 | 1.043 | 1.117 | 1.000 | 1.024 | 1.050 |

Under DGP2, $\lambda_t$ follows a discrete outlier process with fixed, finite support. Consistent with its design, the SVo model performs best in recovering $\lambda_t$, particularly in smaller samples. Although the Dirichlet process mixture is almost surely discrete, it learns the support of the scale distribution nonparametrically rather than imposing a fixed grid. When the DGP coincides exactly with a known discrete outlier structure, the corresponding parametric model has an advantage in finite samples. Importantly, however, DPM-CSV performs similarly to, and in some cases better than, SV-$t$ under this DGP, indicating that the semiparametric approach remains competitive even when the true scale process is discrete.

DGP3 introduces a short contiguous block of extreme observations, mimicking pandemic-style episodes. In this setting, the DPM-CSV model clearly dominates in recovering $\lambda_t$ across both sample sizes. Table 3, which decomposes MSEs into outlier and non-outlier regimes for DGP3, shows that these gains are concentrated within the outlier block, while performance differences during tranquil periods are modest. Taken together, these results indicate that under DGP3 the advantages of the semiparametric specification stem primarily from its ability to adapt to clustered extreme shocks, where flexible modeling of the innovation distribution is most valuable. In stable regimes, by contrast, the competing parametric models perform similarly.

Table 3: MSE for $\lambda_t$ by regime under DGP3

| | $T = 300$ | | | $T = 800$ | | |
|---|---|---|---|---|---|---|
| Regime | DPM | SVo | SV-$t$ | DPM | SVo | SV-$t$ |
| Outlier block | 217 | 410 | 565 | 158 | 316 | 459 |
| Non-outlier | 0.034 | 0.013 | 0.004 | 0.032 | 0.005 | 0.004 |

Across all designs, no model uniformly dominates. Each specification performs best when its structural assumptions align closely with the DGP. The DPM-CSV model delivers the largest gains in the presence of clustered extreme episodes, while remaining competitive under continuous heavy-tailed innovations. These results clarify the conditions under which flexible semiparametric modeling of the innovation distribution is most valuable and highlight the trade-offs between parametric structure and distributional flexibility.

# 5 Macroeconomic Application: Forecasting

We illustrate the proposed semiparametric framework through a macroeconomic forecasting application. This section has two objectives. First, we examine how the proposed models fit a large macroeconomic dataset and assess their ability to capture extreme movements in the innovation distribution through the transitory scale component $\lambda_t$ and the persistent volatility component. Second, we evaluate the forecasting performance of the proposed models relative to commonly used benchmark specifications.

We begin by describing the data, model setup, and priors. We then present full-sample estimation results that illustrate the behavior of the latent volatility and outlier components. Next, we outline the forecasting design and evaluation metrics. Finally, we report results from a pseudo out-of-sample forecasting exercise, comparing the proposed models with several benchmark VAR specifications using both point and density forecast measures.

## 5.1 Data and Model Setup

We evaluate the proposed models using the FRED-MD dataset maintained by the Federal Reserve Bank of St. Louis (McCracken and Ng, 2016). For the full-sample estimation results, we follow Carriero, Clark, Marcellino, and Mertens (2024) and use the same 16 monthly macroeconomic and financial variables, transformed according to their data transformation codes to ensure stationarity. The sample begins in March 1959 and extends to September 2025 using the January 2026 vintage of FRED-MD, yielding a total of $T = 788$ observations.

For the forecasting exercise, we enlarge the information set by adding 26 macroeconomic and financial variables, resulting in a 42-variable VAR. This broader system is intended to assess whether the proposed semiparametric framework remains effective in a moderately large forecasting environment.

We compare the proposed semiparametric specifications with several benchmark models considered in Carriero, Clark, Marcellino, and Mertens (2024): SVo, SVO, and SV-$t$. For the semiparametric models, we estimate two specifications: DPM-CSV (common stochastic volatility) and DPM-OISV (order-invariant stochastic volatility). Both models are estimated using 10,000 MCMC draws after a burn-in period of 5,000 iterations.[7]

We assign Minnesota-type priors to the VAR coefficients (Doan, Litterman, and Sims, 1984; Karlsson, 2013). For DPM-CSV, we use a standard Minnesota prior with overall shrinkage hyperparameter $c_1 = 0.04$ and no intercept term. For DPM-OISV, we adopt the Minnesota-type horseshoe prior of Chan, Koop, and Yu (2024), in which the global and local shrinkage parameters satisfy $\kappa_1, \kappa_2, \psi_{ij} \sim \mathcal{C}^+(0, 1)$. The Dirichlet process con-

[7]Replication code, datasets, and estimation and forecasting results are available on the GitHub repository at `https://github.com/sisurain/DPMSV2024`.

centration parameter is assigned a Gamma prior $\alpha \sim \mathcal{G}(20, 8)$, implying a prior mean of 2.5 clusters.

For the stochastic volatility component, the DPM-CSV model uses the prior $\sigma_h^2 \sim \mathcal{IG}(11, 0.1)$, implying a prior mean of 0.01. For DPM-OISV, each volatility innovation variance $\sigma_{h_i}^2$ follows $\mathcal{IG}(11, 0.1)$, providing a comparable prior scale across the two specifications.

## 5.2 Full-Sample Estimates

Before turning to the forecasting exercises, we first examine the full-sample estimation results to understand how the semiparametric framework captures extreme movements in the innovation distribution. Throughout this subsection, we focus on the DPM-OISV specification, which combines the Dirichlet process mixture with the order-invariant stochastic volatility model of Chan, Koop, and Yu (2024). This specification provides a flexible multivariate volatility structure while maintaining stable posterior sampling.

To assess the reliability of the posterior sampler, we compute convergence diagnostics based on Geweke statistics and inefficiency factors. The results indicate satisfactory convergence for the Dirichlet process parameters and most stochastic volatility parameters, with inefficiency factors typically below 80. The Geweke diagnostics generally fail to reject equality of means between the early and late portions of the chain, suggesting that the Markov chain has reached its stationary distribution.

The estimation reveals that the transitory scale component $\lambda_t$ varies substantially over time. The posterior mean ranges from approximately 1.8 to 9.5, with both the minimum and median close to 1.8. This pattern indicates that most observations belong to a baseline regime with relatively stable innovation scale, while a small number of periods exhibit substantially larger shocks. Figure 1 plots the posterior mean of $\lambda_t$. Several well-known episodes of macroeconomic turbulence are clearly visible, including the high-inflation period of the late 1970s and early 1980s and the COVID-19 pandemic. The largest values occur during the early months of the pandemic, reflecting the unprecedented volatility observed across many macroeconomic series.

Importantly, these transitory outlier states occur on top of the persistent stochastic

volatility component governed by $\{\boldsymbol{\Sigma}_t\}$. In other words, the stochastic volatility process captures gradual and persistent movements in macroeconomic volatility, while $\lambda_t$ accounts for short-lived bursts of unusually large shocks. Quantitatively, the maximum posterior mean reaches about 9.5, indicating that the scale of innovations during the most turbulent episodes is more than five times larger than in normal periods. The sharp contrast between the baseline level and these spikes suggests that the model attributes most observations to a regular regime, while reserving substantially larger scale adjustments for a small number of extreme events.

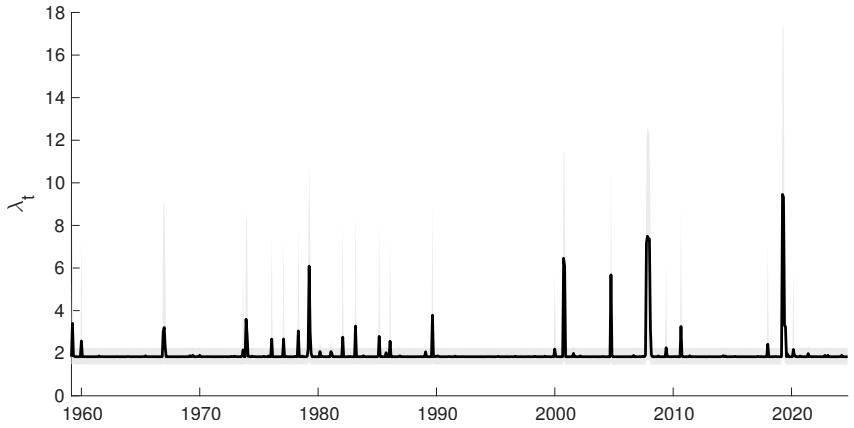


Figure 1: Posterior mean of $\lambda_t$ for the full sample. The shaded region represents the pointwise 90% credible interval.

These results illustrate how the semiparametric framework automatically detects periods of unusually large shocks without imposing a fixed outlier structure. In this sense, $\lambda_t$ acts as a flexible outlier-state variable whose magnitude is learned from the data rather than constrained to a predetermined parametric form. Consistent with this interpretation, the posterior mean of the concentration parameter $\alpha$ is about 1.5, with a 90% credible interval of $(1.01, 2.1)$, implying an average of roughly two active clusters over the sample. Moreover, across posterior draws the number of active clusters $K$ ranges between one and three, with a mode of two. Broadly speaking, these clusters correspond to a baseline regime and periods of elevated volatility.

Next, to quantity the amount of asymmetry in the implied innovation distribution, we compute a tail-balance diagnostic based on posterior predictive draws of the innovations. Let $\varepsilon_{i,t}^{(s)}$ denote the $s$-th posterior predictive draw of the innovation for variable $i$ at time $t$, $s = 1, \ldots, S$. Define a fixed scale for each $(i, t)$ as $\sigma_{i,t} = \sqrt{\mathbb{E}[\lambda_t \Sigma_{ii,t}]}$, where the

expectation is taken with respect to the posterior draws. We then construct standardized predictive draws $\widetilde{\varepsilon}_{i,t}^{(s)} = \varepsilon_{i,t}^{(s)}/\sigma_{i,t}$.

For a given threshold $q > 0$ (we use $q = 2$), we estimate the upper-and lower-tail exceedance probabilities by Monte Carlo frequencies:

$$p_{i,t}^{+} = \frac{1}{S}\sum_{s=1}^{S} \mathbf{1}\left(\widetilde{\varepsilon}_{i,t}^{(s)} > q\right), \quad p_{i,t}^{-} = \frac{1}{S}\sum_{s=1}^{S} \mathbf{1}\left(\widetilde{\varepsilon}_{i,t}^{(s)} < -q\right).$$

These probabilities are then averaged across time:

$$p_i^{+} = \frac{1}{T}\sum_{t=1}^{T} p_{i,t}^{+}, \qquad p_i^{-} = \frac{1}{T}\sum_{t=1}^{T} p_{i,t}^{-}.$$

Finally, we summarize tail asymmetry using the index $A_i = \log(p_i^{+}/p_i^{-})$.

Under a symmetric innovation distribution, the probabilities of large positive and negative realizations should coincide and $A_i \approx 0$. Positive values of $A_i$ indicate relatively heavier right tails, whereas negative values indicate heavier left tails. Because $A_i = \log(p_i^{+}/p_i^{-})$, its magnitude measures the degree of tail imbalance. For example, $A_i = 0.5$ implies that large positive shocks are about $\mathrm{e}^{0.5} \approx 1.65$ times as likely as large negative shocks.

Figure 2 plots the asymmetry index $A_i$ for the standardized innovations. Several variables exhibit clear evidence of tail asymmetry. Real activity indicators such as industrial production and capacity utilization display heavier right tails, reflecting occasional large positive rebounds. In contrast, financial variables such as equity prices and exchange rates exhibit heavier left tails, consistent with the presence of crash risk. Overall, these results provide direct evidence that the innovation distribution is asymmetric across a number of variables.

An important advantage of the semiparametric specification is that it can accommodate both heavy left and right tails simultaneously across variables. By contrast, commonly used parametric skewed distributions typically impose a single direction of skewness. The Dirichlet process mixture therefore provides greater flexibility in capturing heterogeneous and asymmetric tail behavior in multivariate macroeconomic data.

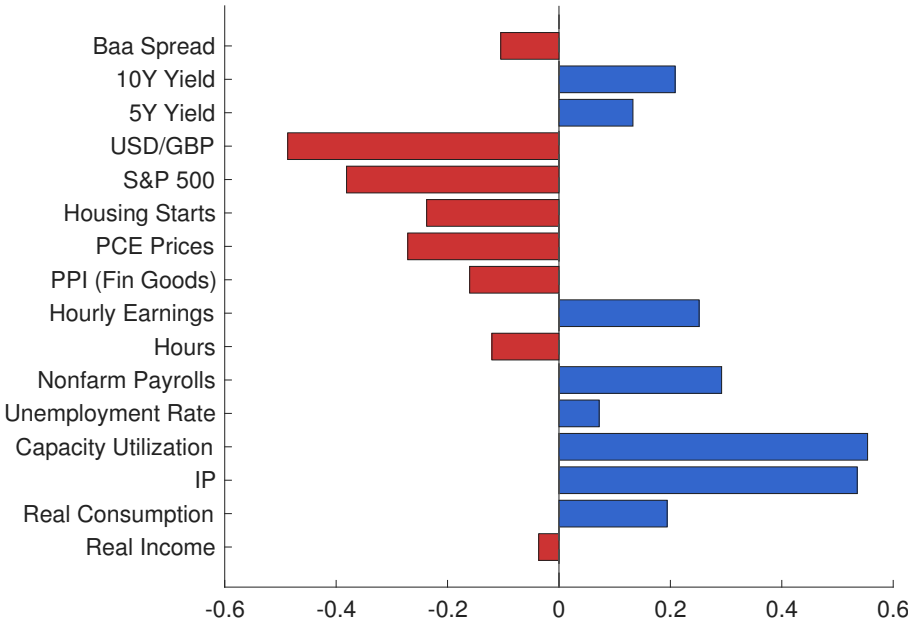


Figure 2: Tail asymmetry diagnostic for standardized innovations from DPM-OISV. Positive values of $A_i$ indicate heavier right tails, while negative values indicate heavier left tails.

## 5.3 Forecasting Results

We now turn to the pseudo out-of-sample forecasting exercise. We begin by outlining the forecasting design and evaluation metrics, then report the results for joint density forecasts and selected individual series.

### 5.3.1 Forecasting Design

Forecasts are generated using an expanding-window scheme with a fixed initial sample beginning in March 1959. At each forecast origin, all models are re-estimated using the available data, and forecasts are produced for horizons ranging from 1 to 12 months ahead. The forecasting system includes 42 monthly macroeconomic and financial variables from FRED-MD, and the evaluation period spans 250 forecast origins from December 2004 to September 2025.

We compare the proposed semiparametric specifications with four benchmark models:

SV, SVo, SVO, and SV-$t$. Forecast accuracy is evaluated along three dimensions. Point forecasts are assessed using root mean squared errors (RMSE), marginal density forecasts are assessed using continuous ranked probability scores (CRPS), and multivariate density forecasts are assessed using joint log predictive scores. Specifically, the joint log predictive score is computed as the sum of the one-step-ahead joint log predictive likelihoods, obtained by evaluating the predictive density at the realized vector of outcomes at each forecast origin. To better understand performance across different environments, we report results for the full sample as well as for three subsamples: pre-COVID (December 2004 to February 2020), COVID (March 2020 to December 2021), and post-COVID (January 2022 to September 2025).

### 5.3.2 Joint Density Forecast Performance

We begin with joint density forecast performance, as this provides the most direct comparison of the competing multivariate models. Table 4 reports the joint log predictive scores over the full evaluation period, together with the corresponding results for the pre-COVID, COVID, and post-COVID subsamples. This decomposition helps distinguish relatively normal forecasting environments from periods characterized by unusually large and clustered shocks, where the semiparametric specification is expected to be especially valuable. The reported values are sums of the one-step-ahead joint log predictive likelihoods evaluated at the realized outcomes, so higher values indicate better multivariate density forecast performance.

Table 4: Joint log predictive scores for the competing models over the full evaluation sample and the three subsamples (pre-COVID, COVID, and post-COVID). Higher values indicate better multivariate density forecast performance.

| Period | SVo | SVO | SV-$t$ | DPM-CSV | DPM-OISV |
|---|---|---|---|---|---|
| Pre-COVID | −91.1 | −102.4 | −71.7 | −74.7 | −54.9 |
| COVID | −177.0 | −543.9 | −204.6 | −195.5 | −332.0 |
| Post-COVID | −84.3 | −83.6 | −68.7 | −69.8 | −56.1 |
| Full sample | −97.5 | −138.1 | −82.9 | −84.5 | −79.6 |

Several findings emerge. First, DPM-OISV delivers the strongest joint density forecast

performance over the full sample, with a joint log predictive score of $-79.6$, followed by $-82.9$ for SV-$t$, $-84.5$ for DPM-CSV, and $-97.5$ for SVo. Second, its gains are especially strong outside the pandemic period: DPM-OISV attains the highest score in both the pre-COVID and post-COVID subsamples. These results are important because they show that the proposed semiparametric specification is not useful only during the COVID episode, but also improves multivariate density forecasts in more stable environments.

During the COVID subsample, however, DPM-OISV underperforms SVo in terms of the joint log score, with a value of $-332.0$ compared with $-177.0$ for SVo. One possible explanation is that the richer 42-dimensional volatility structure in DPM-OISV, while beneficial in normal and post-pandemic periods, becomes harder to estimate precisely during the most turbulent part of the sample. This interpretation is consistent with the poor performance of SVO, which allows for highly flexible variable-specific outlier dynamics and performs worst among all the models in this subsample. More generally, additional flexibility may come at a cost in joint density evaluation when many series are simultaneously subject to extreme and heterogeneous shocks. Even so, DPM-OISV still dominates SVO during COVID and remains the best-performing model overall on the full sample.

### 5.3.3 Selected Series and Density Calibration

To better understand where the gains in density forecasting come from, we focus on three representative series: industrial production, PCE inflation, and the unemployment rate. These variables capture three distinct dimensions of macroeconomic fluctuations—real activity, inflation, and labor market conditions—and are also among the most informative series in the forecasting exercise. Online Appendix C reports the corresponding results for all 42 series, together with RMSE-based comparisons for point forecasts across the full sample and the pre-COVID, COVID, and post-COVID subsamples.

Table 5 reports the CRPS of the benchmark models relative to DPM-OISV for these three series at horizons $h = 1, 6, 12$. Values below 1 indicate that DPM-OISV produces more accurate density forecasts, whereas values above 1 favor the benchmark model. Several patterns emerge. First, the gains from DPM-OISV are especially pronounced for PCE inflation and industrial production, particularly at the longer horizons. For example,

for PCE inflation at $h = 6$ the relative CRPS is about 0.76-0.78 against SVo and SVO, implying a substantial improvement in density forecast accuracy. Industrial production shows a similar, though somewhat smaller, pattern.

Second, for PCE inflation and industrial production DPM-OISV also tends to outperform SV-$t$ and DPM-CSV, suggesting that both the semiparametric innovation specification and the more flexible stochastic volatility model contribute to the forecasting gains. By contrast, the unemployment rate shows much weaker evidence of improvement, and at longer horizons some benchmark models perform better. More broadly, the full results reported in the appendix suggest that the semiparametric specification is especially competitive for financial and price-related variables, whereas the gains are less uniform for labor-market variables. This pattern is consistent with the view that the conditional distributions of financial and price variables are more likely to exhibit heavy tails, skewness, or clustered extreme observations, features that the DPM-based approach is better suited to capture than fully parametric alternatives.

Table 5: CRPS ratios for the competing models relative to DPM-OISV over the full sample period for selected series. Values below 1 favor DPM-OISV.

| Variable | Horizon | SVo | SVO | SV-$t$ | DPM-CSV |
|---|---|---|---|---|---|
| PCE prices | $h = 1$ | $0.84^{*}$ | $0.83^{**}$ | 0.90 | 0.95 |
| | $h = 6$ | $0.78^{***}$ | $0.76^{***}$ | 0.96 | 0.99 |
| | $h = 12$ | $0.49^{***}$ | $0.50^{***}$ | 0.94 | 0.94 |
| Industrial production | $h = 1$ | 0.88 | $0.90^{*}$ | 1.00 | 1.01 |
| | $h = 6$ | $0.86^{*}$ | $0.83^{*}$ | 0.99 | 0.98 |
| | $h = 12$ | $0.58^{***}$ | $0.57^{**}$ | 0.97 | 0.95 |
| Unemployment rate | $h = 1$ | 0.93 | 0.96 | 1.03 | 1.13 |
| | $h = 6$ | 1.19 | 1.24 | 1.28 | 1.24 |
| | $h = 12$ | 1.07 | 1.08 | 1.25 | 1.15 |

Each entry is the ratio of the DPM-OISV forecast score to that of the indicated benchmark model. Values below 1 favor DPM-OISV, while values above 1 favor the benchmark. Asterisks denote Diebold-Mariano-West tests based on Newey-West HAC standard errors: $^{***}$ $p < 0.001$, $^{**}$ $0.001 \leqslant p < 0.01$, and $^{*}$ $0.01 \leqslant p < 0.05$.

Overall, these results reinforce the broader pattern in the forecasting exercise. The pro-

posed semiparametric models deliver sizable gains for some series, especially those whose forecast distributions appear to exhibit stronger non-Gaussian features, but the improvements are not uniform across all variables and horizons.

To shed further light on the sources of the forecasting gains, Figures 3 and 4 present spaghetti plots for PCE inflation and the unemployment rate. Each panel overlays the sequence of $h$-step-ahead predictive mean paths ($h = 1, \ldots, 12$) against the realized series, with the shaded region marking the COVID subsample.

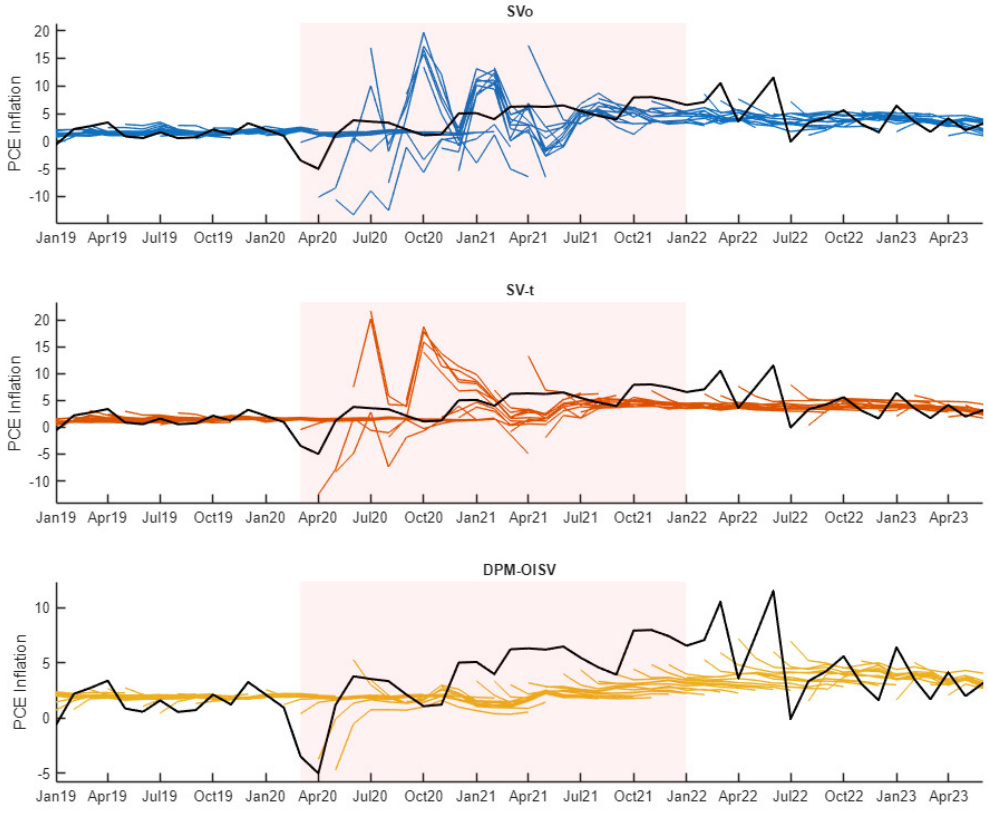


Figure 3: Forecast paths ($h = 1, \ldots, 12$ steps ahead) for PCE inflation from SVo (top), SV-$t$ (middle), and DPM-OISV (bottom). The black line is the realized series; colored lines are predictive mean paths issued at each forecast origin. The shaded region marks the COVID subsample (March 2020-December 2021).

The two figures illustrate the contrasting behavior of the models during the pandemic

and help explain the differential performance documented in Table 5. For PCE inflation, the forecast paths from SVo and SV-$t$ are highly erratic during the COVID period, with individual paths swinging sharply in both directions and failing to track the persistent inflation surge that materialized from 2021 onward. The DPM-OISV paths are considerably more stable and better centered around the realized series throughout the pandemic and recovery period, though all models consistently underpredict the magnitude of the post-pandemic inflation surge. This pattern reflects the ability of the Dirichlet process mixture to assign pandemic observations to a distinct cluster with an appropriately elevated but controlled scale, whereas the parametric outlier states in SVo and the heavy-tailed innovations in SV-$t$ tend to generate excessive dispersion in the predictive paths without improving their central tendency.

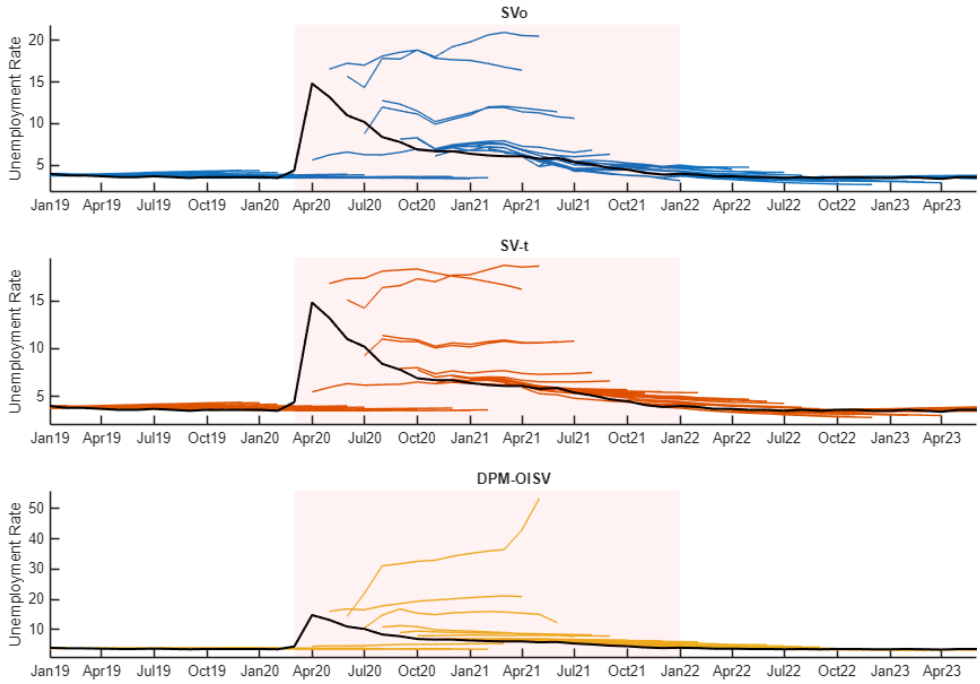


Figure 4: Forecast paths ($h = 1, \ldots, 12$ steps ahead) for the unemployment rate from SVo (top), SV-$t$ (middle), and DPM-OISV (bottom). The black line is the realized series; colored lines are predictive mean paths issued at each forecast origin. The shaded region marks the COVID subsample (March 2020-December 2021).

The unemployment rate panel tells a different story. Here, the DPM-OISV predictive

mean paths during the pandemic extend far above the realized series, with some trajectories reaching implausibly large values. Because the scale component $\lambda_t$ is common across all variables, the large shocks concentrated in financial and price series during 2020 inflate the shared innovation scale and shift the predictive means upward for unemployment, whose pandemic response—while large—was more structured and mean-reverting than suggested by the DPM-OISV paths. This cross-variable contamination of the predictive mean is an inherent limitation of the common innovation distribution and provides a direct explanation for the weaker performance of DPM-OISV on labor-market variables documented in Table 5.

Finally, to assess the calibration of the predictive densities, Figure 5 presents the probability integral transform (PIT) histograms at the one-step-ahead horizon for industrial production, PCE inflation, and the unemployment rate across the four models. A well-calibrated predictive density implies that the PIT values follow a standard uniform distribution, corresponding to a flat histogram aligned with the dashed reference line. Deviations from uniformity signal misspecification of the predictive densities: a hump shape indicates overdispersion, a U shape suggests underdispersion, and skewness points to a systematic bias in the location of the predictive density.

The PIT histograms highlight several differences across models. The SVo model exhibits noticeable departures from uniformity, particularly for industrial production, where the histogram displays excess mass in the left tail, suggesting that the predictive densities do not fully account for large negative realizations. The SV-$t$ and DPM-CSV models offer some improvement, producing somewhat more uniform histograms, though mild irregularities remain for PCE inflation and the unemployment rate. Among all four specifications, the DPM-OISV model yields the most uniform PIT histograms across all three variables, indicating that its predictive densities are the best calibrated. This finding is consistent with the log predictive score results reported above and confirms that the flexible nonparametric modeling of both outliers and stochastic volatility translates into tangible gains in density forecast calibration.

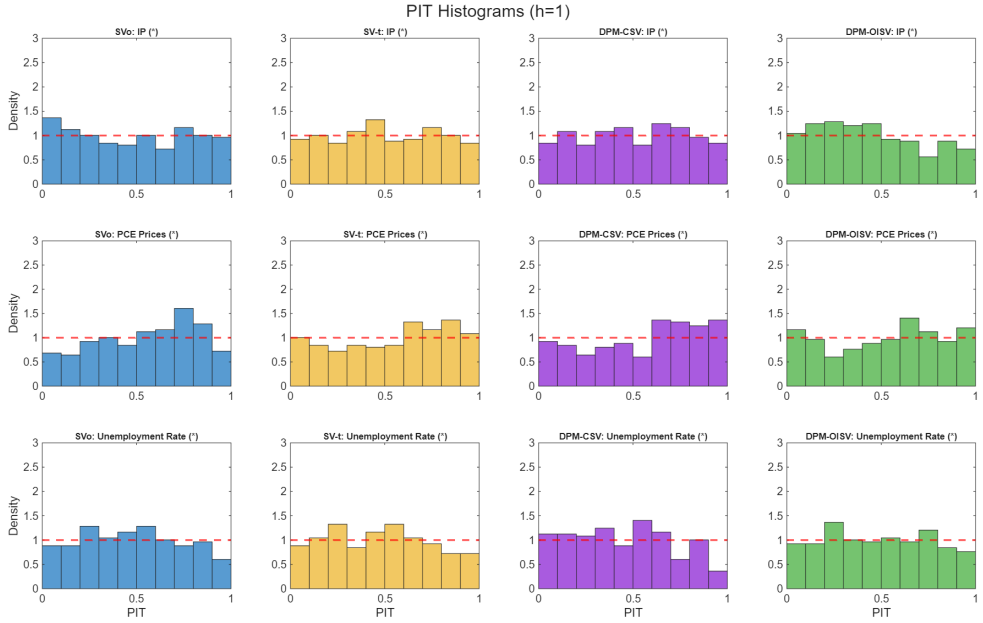


Figure 5: PIT histograms at the one-step-ahead horizon ($h = 1$) for industrial production (top), PCE inflation (middle), and the unemployment rate (bottom). Columns correspond to SVo, SV-$t$, DPM-CSV, and DPM-OISV, respectively. The dashed red line marks the uniform density. A well-calibrated predictive density produces a flat histogram.

# 6 Financial Application: Bank Connectedness

In this section we apply the proposed semiparametric framework to measure financial connectedness among global banks, using the generalized forecast error variance decomposition (GFEVD) of Diebold and Yılmaz (2014). Because the GFEVD depends on the correlation structure of the innovation covariance rather than its scale, the transitory scale component $\lambda_t$ in DPM-OISV provides a natural channel for separating scale-driven volatility spikes from genuine shifts in financial transmission. We compare the connectedness estimates from DPM-OISV with those from a standard stochastic volatility VAR and examine how the two specifications differ in their characterization of crisis dynamics.

## 6.1 Data and Setup

We use a subset of the bank-level dataset of Demirer, Diebold, Liu, and Yilmaz (2018), consisting of the 30 largest global banks by total assets. The data contain daily log realized variances, and the sample covers September 17, 2003 to February 7, 2014, yielding $T = 2,673$ trading-day observations.

Both models are estimated as VAR(3) specifications with the Minnesota-type horseshoe prior on the VAR coefficients. As a benchmark, we estimate a standard Cholesky stochastic volatility VAR (SV) with a lower-triangular impact matrix and independent AR(1) log-volatility processes. The posterior sampler for each model is run for 10,000 iterations after a burn-in period of 5,000 iterations, retaining every 10th draw.

Financial connectedness is measured using the GFEVD of Diebold and Yılmaz (2014) at horizon $H = 10$. The total connectedness index is defined as

$$\mathcal{C}_t^H = \frac{\sum_{i \neq j} \theta_{ij,t}^H}{\sum_{i,j} \theta_{ij,t}^H} \times 100,$$

where $\theta_{ij,t}^H$ is the $H$-step-ahead share of the forecast error variance of bank $i$ attributable to shocks from bank $j$ at time $t$. Connectedness is evaluated at $T_{\text{sub}} = 134$ subsampled dates (every 20 trading days) and averaged over the posterior draws at each date, yielding full posterior distributions for all connectedness measures.

## 6.2 Connectedness Estimates

The two models agree closely on both the level and the cross-sectional pattern of connectedness. The full-sample means are 73.45% (DPM-OISV) and 73.48% (SV), and the Spearman rank correlation of per-bank net transmission across the two models is 0.99. Under both specifications, the top net transmitters are BNP Paribas, Deutsche Bank, and Société Générale, while the largest net receivers are concentrated among among Asian, Canadian, and peripheral European banks. The semiparametric innovation specification therefore preserves the economic characterization of financial transmission in normal times.

Where the two models differ is in their temporal dynamics during crises. Figure 6 plots

the posterior median of $\mathcal{C}_t^H$ under DPM-OISV and SV over 2007–2014, together with 16th-84th percentile credible bands. DPM-OISV produces a visibly smoother path and tighter credible bands, with a dynamic range of 16.9 percentage points compared to 20.3 for SV. The reason is that $\lambda_t$ absorbs common-scale movements in the innovation covariance without affecting its correlation structure. The GFEVD depends only on correlations, so the connectedness estimates from DPM-OISV respond to genuine shifts in the cross-bank correlation structure, whereas SV assumes Gaussian innovations, so large shocks are absorbed by the log-volatility processes or changes in the impact matrix, which in turn inflates the estimated connectedness.

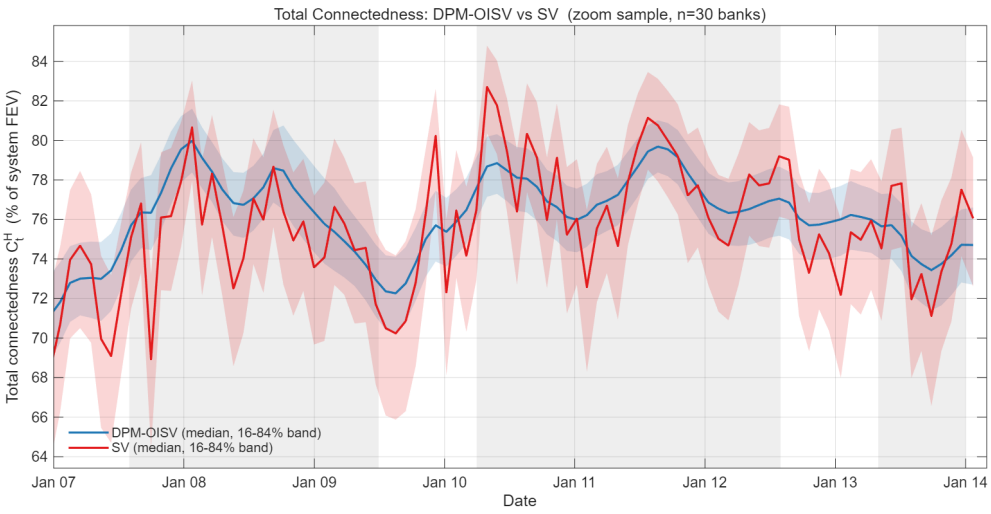


Figure 6: Total connectedness $\mathcal{C}_t^H$ under DPM-OISV (blue) and SV (red), 2007-2014. Solid lines are posterior medians; shaded regions are 16th-84th percentile credible bands.

The gap between the two models is non-monotonic across crisis episodes. Table 6 reports total connectedness at five crisis dates. SV reads substantially higher than DPM-OISV during the Flash Crash/Greek crisis (+5.15 pp) and the U.S. downgrade (+3.34 pp)—episodes driven by large but proportionally correlated moves across banks, where the Gaussian assumption in SV forces these large shocks into the log-volatility processes, inflating the estimated correlation structure. By contrast, the two models are closest at the collapse of Lehman Brothers (+0.53 pp), when the shift in cross-bank correlation was genuine and both specifications respond similarly. This pattern is consistent with the view that DPM-OISV separates scale from correlation through $\lambda_t$, producing connectedness estimates that are less sensitive to the magnitude of shocks and more reflective of actual

changes in financial transmission.

Table 6: Total connectedness at five crisis dates.

| Date | Event | DPM-OISV | SV | Difference |
|---|---|---|---|---|
| 2007-08-09 | BNP liquidity freeze | 75.90 | 77.17 | +1.27 |
| 2008-09-15 | Lehman Brothers | 78.90 | 79.43 | +0.53 |
| 2010-05-06 | Flash Crash / Greek crisis | 78.74 | 83.89 | +5.15 |
| 2011-08-08 | U.S. downgrade | 79.80 | 83.14 | +3.34 |
| 2013-06-19 | Taper Tantrum | 75.72 | 79.34 | +3.62 |

Overall, these results demonstrate that the choice of innovation distribution has substantive implications for the measurement of financial connectedness, particularly during periods of acute stress. By providing a dedicated scale channel through $\lambda_t$, the semi-parametric framework yields connectedness estimates that are robust to common-scale volatility spikes while remaining responsive to genuine shifts in the cross-bank correlation structure.

# 7 Concluding Remarks

We have introduced a novel approach to modeling the transitory component of stochastic volatility in Bayesian VARs using a Dirichlet process mixture prior. We have also developed an efficient collapsed Gibbs sampling algorithm for the proposed models in the presence of stochastic volatility. The framework offers several advantages: it accommodates a wide range of transitory volatility dynamics, allows for asymmetric and heavy-tailed innovation distributions, and provides a flexible way to handle unknown outliers. In addition, the proposed approach is scalable to large systems. Our empirical results show that it delivers meaningful gains in forecasting performance, especially in settings characterized by large and clustered shocks.

There are several promising directions for future research. One is to explore alternative Bayesian nonparametric priors for VARs, such as the two-parameter Poisson-Dirichlet process, which may allow for richer clustering behavior and better information shar-

ing across mixture components. Another is to combine the Dirichlet process mixture framework with nonparametric VAR specifications, potentially yielding a more flexible approach for large datasets and settings with mixed-frequency observations. For example, Kalli and Griffin (2018) develop a nonparametric VAR framework, although its scalability to high-dimensional systems remains unclear. Extending such approaches to large VARs would likely require more advanced computational tools, including efficient missing-data imputation and variational inference methods.

The proposed framework may also prove useful for measuring uncertainty and for structural analysis. In settings where stochastic volatility is interpreted as a measure of macroeconomic uncertainty (see, e.g., Jurado, Ludvigson, and Ng, 2015; Carriero, Clark, and Marcellino, 2018), more flexible modeling of the innovation distribution and covariance process could affect both economic interpretation and statistical inference. In structural VARs, improved modeling of heteroskedasticity and shock distributions may enhance identification and impulse response inference. Exploring these connections, including links to recent work such as Braun (2023), is a natural direction for future research.

# References


Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi (2017): "Microeconomic Origins of Macroeconomic Tail Risks," *The American Economic Review*, 107(1), 54–108.

Aguilar, O., and M. West (2000): "Bayesian dynamic factor models and portfolio allocation," *Journal of Business and Economic Statistics*, 18(3), 338–357.

Arias, J. E., J. F. Rubio-Ramirez, and M. Shin (2023): "Macroeconomic forecasting and variable ordering in multivariate stochastic volatility models," *Journal of Econometrics*, 235(2), 1054–1086.

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

Braun, R. (2023): "The importance of supply and demand for oil prices: Evidence from non-Gaussianity," *Quantitative Economics*, 14(4), 1163–1198.

Carriero, A., J. C. C. Chan, T. E. Clark, and M. G. Marcellino (2022): "Corrigendum to "Large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors"," *Journal of Econometrics*, 227(2), 506–512.

Carriero, A., T. E. Clark, and M. Marcellino (2018): "Measuring Uncertainty and Its Impact on the Economy," *Review of Economics and Statistics*, 100(5), 799–815.

Carriero, A., T. E. Clark, M. Marcellino, and E. Mertens (2024): "Addressing COVID-19 Outliers in BVARs with Stochastic Volatility," *The Review of Economics and Statistics*, 106(5), 1403–1417.

Carriero, A., T. E. Clark, and M. G. Marcellino (2016): "Common drifting volatility in large Bayesian VARs," *Journal of Business and Economic Statistics*, 34(3), 375–390.

——— (2019): "Large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors," *Journal of Econometrics*, 212(1), 137–154.

Chan, J. C. C. (2020): "Large Bayesian VARs: A Flexible Kronecker Error Covariance Structure," *Journal of Business and Economic Statistics*, 38(1), 68–79.

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

——— (2024): "BVARs and Stochastic Volatility," in *Handbook of Research Methods and Applications in Macroeconomic Forecasting*, ed. by M. Clements, and A. Galvão, pp. 43–67. Edward Elgar Publishing.

Chan, J. C. C., and E. Eisenstat (2018): "Bayesian Model Comparison for Time-Varying Parameter VARs with Stochastic Volatility," *Journal of Applied Econometrics*, 33(4), 509–532.

Chan, J. C. C., and I. Jeliazkov (2009): "Efficient Simulation and Integrated Likelihood Estimation in State Space Models," *International Journal of Mathematical Modelling and Numerical Optimisation*, 1(1), 101–120.

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

Chib, S., F. Nardari, and N. Shephard (2006): "Analysis of high dimensional multivariate stochastic volatility models," *Journal of Econometrics*, 134(2), 341–371.

Chiu, C. J., H. Mumtaz, and G. Pinter (2017): "Forecasting with VAR models: Fat tails and stochastic volatility," *International Journal of Forecasting*, 33(4), 1124–1143.

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

Clark, T. E., F. Huber, G. Koop, and M. Marcellino (2024): "Forecasting US inflation using Bayesian nonparametric models," *The Annals of Applied Statistics*, 18(2), 1421–1444.

Clark, T. E., and E. Mertens (2023): "Stochastic Volatility in Bayesian Vector Autoregressions," in *Oxford Research Encyclopedia of Economics and Finance*. Oxford University Press.

Clark, T. E., and F. Ravazzolo (2015): "Macroeconomic Forecasting Performance under Alternative Specifications of Time-Varying Volatility," *Journal of Applied Econometrics*, 30(4), 551–575.

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

Cross, J., and A. Poon (2016): "Forecasting structural change and fat-tailed events in Australian macroeconomic variables," *Economic Modelling*, 58, 34–51.

Cúrdia, V., M. Del Negro, and D. L. Greenwald (2014): "Rare shocks, great recessions," *Journal of Applied Econometrics*, 29(7), 1031–1052.

D'Agostino, A., L. Gambetti, and D. Giannone (2013): "Macroeconomic forecasting and structural change," *Journal of Applied Econometrics*, 28, 82–101.

Del Negro, M., and G. E. Primiceri (2015): "Time-varying structural vector autoregressions and monetary policy: a corrigendum," *Review of Economic Studies*, 82(4), 1342–1345.

Demirer, M., F. X. Diebold, L. Liu, and K. Yilmaz (2018): "Estimating Global Bank Network Connectedness," *Journal of Applied Econometrics*, 33(1), 1–15.

Diebold, F. X., and K. Yilmaz (2014): "On the Network Topology of Variance Decompositions: Measuring the Connectedness of Financial Firms," *Journal of Econometrics*, 182(1), 119–134.

Doan, T., R. Litterman, and C. Sims (1984): "Forecasting and conditional projection using realistic prior distributions," *Econometric reviews*, 3(1), 1–100.

Escobar, M. D., and M. West (1995): "Bayesian Density Estimation and Inference Using Mixtures," *Journal of the American Statistical Association*, 90(430), 577–588.

Hauzenberger, N., F. Huber, M. Marcellino, and N. Petz (2025): "Gaussian process vector autoregressions and macroeconomic uncertainty," *Journal of Business & Economic Statistics*, 43(1), 27–43.

Huber, F., and G. Koop (2024): "Fast and order-invariant inference in Bayesian VARs with nonparametric shocks," *Journal of Applied Econometrics*, 39(7), 1301–1320.

Janssens, E. F., and R. L. Lumsdaine (2025): "Restricted Large Bayesian Vector Autoregressions," *Available at SSRN 5342958*.

Jensen, M. J., and J. M. Maheu (2010): "Bayesian semiparametric stochastic volatility modeling," *Journal of Econometrics*, 157(2), 306–316.

Jensen, M. J., and J. M. Maheu (2014): "Estimating a semiparametric asymmetric stochastic volatility model with a Dirichlet process mixture," *Journal of Econometrics*, 178, 523–538.

Jurado, K., S. C. Ludvigson, and S. Ng (2015): "Measuring Uncertainty," *American Economic Review*, 105(3), 1177–1216.

Kalli, M., and J. E. Griffin (2018): "Bayesian nonparametric vector autoregressive models," *Journal of econometrics*, 203(2), 267–282.

Karlsson, S. (2013): "Forecasting with Bayesian vector autoregressions," in *Handbook of Economic Forecasting*, ed. by G. Elliott, and A. Timmermann, vol. 2, pp. 791–897. Elsevier.

Karlsson, S., and S. Mazur (2020): "Flexible fat-tailed vector autoregression," Working Paper No. 5/2020, Örebro University School of Business, Örebro.

Karlsson, S., S. Mazur, and H. Nguyen (2023): "Vector autoregression models with skewness and heavy tails," *Journal of Economic Dynamics and Control*, 146, 104580.

Kastner, G. (2019): "Sparse Bayesian time-varying covariance estimation in many dimensions," *Journal of Econometrics*, 210(1), 98–115.

Kastner, G., and F. Huber (2020): "Sparse Bayesian vector autoregressions in huge dimensions," *Journal of Forecasting*, 39(7), 1142–1165.

Koop, G., and D. Korobilis (2013): "Large time-varying parameter VARs," *Journal of Econometrics*, 177(2), 185–198.

Marcellino, M., and M. Pfarrhofer (2024): "Bayesian Nonparametric Methods for Macroeconomic Forecasting," in *Handbook of Research Methods and Applications in Macroeconomic Forecasting*, ed. by M. P. Clements, and A. B. Galvão, pp. 90–125. Edward Elgar.

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

Neal, R. M. (2000): "Markov Chain Sampling Methods for Dirichlet Process Mixture Models," *Journal of Computational and Graphical Statistics*, 9(2), 249–265.

Nguyen, T. T., H. D. Nguyen, F. Chamroukhi, and G. J. McLachlan (2020): "Approximation by finite mixtures of continuous density functions that vanish at infinity," *Cogent Mathematics & Statistics*, 7(1), 1750861.

Pitt, M., and N. Shephard (1999): "Time varying covariances: a factor stochastic volatility approach," *Bayesian Statistics*, 6, 547–570.

Sethuraman, J. (1994): "A constructive definition of Dirichlet priors," *Statistica Sinica*, pp. 639–650.

Song, Y. (2014): "Modelling regime switching and structural breaks with an infinite hidden Markov model," *Journal of Applied Econometrics*, 29(5), 825–842.

Stock, J. H., and M. W. Watson (2016): "Core Inflation and Trend Inflation," *The Review of Economics and Statistics*, 98(4), 770–784.

Sudderth, E. B. (2006): "Graphical models for visual object recognition and tracking," Ph.D. thesis, Massachusetts Institute of Technology.

Wu, F. C. Z. (2024): "Bayesian collapsed Gibbs sampling for a stochastic volatility model with a Dirichlet process mixture," *Journal of Applied Econometrics*, 39(4), 697–704.

Wu, P., and G. Koop (2025): "Fast, order-invariant Bayesian inference in VARs using the eigendecomposition of the error covariance matrix," *Journal of Business & Economic Statistics*, just-accepted, 1–21.

Zellner, A. (1971): *An Introduction to Bayesian Inference in Econometrics*. Wiley, New York.

# Online Appendices:

# A Proofs of Propositions

This appendix provides proofs of Theorem 1 and a few propositions. Let $C_0(\mathbb{R}^n)$ denote the class of continuous probability density functions on $\mathbb{R}^n$ that vanishes at infinity, i.e., $f \in C_0(\mathbb{R}^n)$ if $f$ is a continuous pdf and for every $\varepsilon > 0$, there exists $M$ such that for all $\|\mathbf{x}\| > M$, $f(\mathbf{x}) < \varepsilon$.

We first prove the following useful lemma.

**Lemma 1.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuous. Suppose there exists a function $\psi : [0, \infty) \to [0, \infty)$ with $\psi(r) \to 0$ as $r \to \infty$ and $R \geqslant 0$ such that*

$$|f(\mathbf{x})| \leqslant \psi(\|\mathbf{x}\|) \quad \textit{for all} \quad \|\mathbf{x}\| \geqslant R.$$

*Then $f$ vanishes at infinity.*

*Proof.* Fix $\varepsilon > 0$. Since $\psi(r) \to 0$ as $r \to \infty$, there exists $M$ such that $\phi(r) < \varepsilon$ for all $r > M$. Without loss of generality we can choose $M \geqslant R$. Then, for $\|\mathbf{x}\| > M$ we have $|f(\mathbf{x})| \leqslant \psi(\|\mathbf{x}\|) < \varepsilon$. Hence, $f$ vanishes at infinity. □

**Proposition 1** (Polynomial decay of tails)**.** *Let $f$ be a continuous pdf on $\mathbb{R}^n$. Suppose there exist constants $C > 0$, $\alpha > 0$, and $R \geqslant 0$ such that*

$$f(\mathbf{x}) \leqslant \frac{C}{(1 + \|\mathbf{x}\|)^{\alpha}} \quad \textit{for all} \quad \|\mathbf{x}\| \geq R.$$

*Then $f \in C_0(\mathbb{R}^n)$.*

*Proof.* Applying Lemma 1 with $\psi(r) = C(1 + r)^{-\alpha}$, $f$ vanishes at infinity. Since $f$ is a continuous pdf, $f \in C_0(\mathbb{R}^n)$. □

**Proposition 2** (Exponential decay of tails)**.** *Let $f$ be a continuous pdf on $\mathbb{R}^n$. Suppose there exist constants $C > 0$, $c > 0$, $\beta > 0$, and $R \geqslant 0$ such that*

$$f(\mathbf{x}) \leqslant Ce^{-c\|\mathbf{x}\|^{\beta}} \quad \textit{for all} \quad \|\mathbf{x}\| \geq R. \tag{5}$$

*Then $f \in C_0(\mathbb{R}^n)$. In particular, the density*

$$\phi(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = (2\pi)^{-n/2}|\boldsymbol{\Sigma}|^{-1/2}e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}$$

*associated with the $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ distribution is a member of $C_0(\mathbb{R}^n)$.*

*Proof.* We apply Lemma 1 with $\psi(r) = C\mathrm{e}^{-c\,r^\beta}$ to conclude that $f$ vanishes at infinity. To show that $\phi(\mathbf{x};\boldsymbol{\mu},\boldsymbol{\Sigma}) \in C_0(\mathbb{R}^n)$, let $\lambda_{\max}$ be the largest eigenvalue of $\boldsymbol{\Sigma}$. Then, for all $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x}'\boldsymbol{\Sigma}^{-1}\mathbf{x} \geqslant \|\mathbf{x}\|^2/\lambda_{\max}$. It follows that

$$\phi(\mathbf{x};\boldsymbol{\mu},\boldsymbol{\Sigma}) \leqslant (2\pi)^{-n/2}|\boldsymbol{\Sigma}|^{-1/2}\mathrm{e}^{-\frac{\|\mathbf{x}-\boldsymbol{\mu}\|^2}{2\lambda_{\max}}}.$$

For $\|\mathbf{x}\| \geqslant 2\|\boldsymbol{\mu}\|$, $\|\mathbf{x}-\boldsymbol{\mu}\| \geqslant \|\mathbf{x}\| - \|\boldsymbol{\mu}\| \geqslant \frac{1}{2}\|\mathbf{x}\|$, giving

$$\phi(\mathbf{x};\boldsymbol{\mu},\boldsymbol{\Sigma}) \leq (2\pi)^{-n/2}|\boldsymbol{\Sigma}|^{-1/2}\mathrm{e}^{-\frac{\|\mathbf{x}\|^2}{8\lambda_{\max}}} = C\mathrm{e}^{-c\|\mathbf{x}\|^2},$$

which is in the form of (5). □

**Proposition 3** (Uniform continuity, Lipschitz or bounded gradient)**.** *Suppose $f$ is a probability density function. If $f$ is uniformly continuous, then $f \in C_0(\mathbb{R}^n)$. Therefore, if $f$ is Lipschitz or has a bounded gradient, then $f \in C_0(\mathbb{R}^n)$.*

*Proof.* For a probability density function $f$, we first prove by contradiction that if it is uniformly continuous, then it vanishes at infinity. The idea is to construct infinitely many $\delta$-balls of the same volume on which $f$ is bounded below by a positive value. This implies the integral of $f$ over $\mathbb{R}^n$ diverges to infinity, contradicting the assumption that it is a pdf.

Suppose on the contrary that $f$ does not vanish at infinity. Then, there exist $\varepsilon > 0$ and a sequence $(\mathbf{x}_k)$ in $\mathbb{R}^n$ such that

$$\|\mathbf{x}_k\| \to \infty \quad \text{and} \quad f(\mathbf{x}_k) > \varepsilon \quad \text{for all } k.$$

Since $f$ is uniformly continuous, there exists $\delta > 0$ such that

$$\|\mathbf{y}-\mathbf{z}\| < \delta \implies |f(\mathbf{y}) - f(\mathbf{z})| < \frac{\varepsilon}{2}.$$

Hence, for all $\mathbf{y} \in B(\mathbf{x}_k,\delta) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}-\mathbf{x}_k\| < \delta\}$,

$$f(\mathbf{y}) > f(\mathbf{x}_k) - \frac{\varepsilon}{2} > \frac{\varepsilon}{2}.$$

Because $\|\mathbf{x}_k\| \to \infty$, we can choose a subsequence (with a slight abuse of notations, it is still denoted as $\mathbf{x}_k$) such that the balls $B(\mathbf{x}_k,\delta)$ are pairwise disjoint. Therefore,

$$\int_{\mathbb{R}^n} f(\mathbf{y})\,\mathrm{d}\mathbf{y} \geqslant \sum_k \int_{B(\mathbf{x}_k,\delta)} f(\mathbf{y})\,\mathrm{d}\mathbf{y} > \sum_k \frac{\varepsilon}{2}\,\mathrm{vol}(B(\mathbf{0},\delta)),$$

where $\mathrm{vol}(B(\mathbf{0},\delta))$ denotes the volume of the $\delta$-ball centered at the origin. The right-hand side diverges to $\infty$, contradicting the fact that $f$ is a pdf with $\int_{\mathbb{R}^n} f(\mathbf{y})\,\mathrm{d}\mathbf{y} = 1$. Hence,

the assumption is false and $f$ vanishes at infinity.

Next, suppose $f$ is Lipschitz on $\mathbb{R}^n$. That is, there exists $L \geq 0$ such that

$$\|f(\mathbf{x}) - f(\mathbf{y})\| \leqslant L\|\mathbf{x} - \mathbf{y}\| \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

If $L = 0$, then $f$ is constant and thus uniformly continuous. Assume $L > 0$.

Given $\varepsilon > 0$, set $\delta = \varepsilon/L$. Then for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ satisfying $\|\mathbf{x} - \mathbf{y}\| < \delta$, we have

$$\|f(\mathbf{x}) - f(\mathbf{y})\| \leqslant L\|\mathbf{x} - \mathbf{y}\| < L \cdot \frac{\varepsilon}{L} = \varepsilon.$$

The choice of $\delta$ depends only on $\varepsilon$ and not on the specific points $\mathbf{x}, \mathbf{y}$, so $f$ is uniformly continuous on $\mathbb{R}^n$. Therefore, by the earlier result, $f \in C_0(\mathbb{R}^n)$.

Finally, suppose the gradient of $f(\mathbf{x})$ is bounded, particularly, there exists $L \geqslant 0$ such that

$$\|\nabla f(\mathbf{x})\| \leqslant L \quad \text{for all } \mathbf{x} \in \mathbb{R}^n.$$

We show that $f$ is Lipschitz, and therefore $f \in C_0(\mathbb{R}^n)$.

To that end, fix $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and consider the segment $\mathbf{x}_t = \mathbf{x} + t(\mathbf{y} - \mathbf{x})$, $t \in [0, 1]$. Define $g(t) = f(\mathbf{x}_t)$. By the chain rule,

$$g'(t) = \nabla f(\mathbf{x}_t) \cdot (\mathbf{y} - \mathbf{x}).$$

By the Fundamental Theorem of Calculus,

$$f(\mathbf{y}) - f(\mathbf{x}) = g(1) - g(0) = \int_0^1 g'(t)\,\mathrm{d}t = \int_0^1 \nabla f(\mathbf{x}_t) \cdot (\mathbf{y} - \mathbf{x})\,\mathrm{d}t.$$

Taking absolute values and applying the Cauchy-Schwarz inequality,

$$|f(\mathbf{y}) - f(\mathbf{x})| \leqslant \int_0^1 \|\nabla f(\mathbf{x}_t)\|\,\|\mathbf{y} - \mathbf{x}\|\,\mathrm{d}t \leqslant \int_0^1 L\,\|\mathbf{y} - \mathbf{x}\|\,\mathrm{d}t = L\,\|\mathbf{y} - \mathbf{x}\|.$$

Hence, $f$ is Lipschitz. □

*Proof of Theorem 1.* The proof involves verifying the premises of Theorem 5 (a) in Nguyen, Nguyen, Chamroukhi, and McLachlan (2020). First, the base pdf $g$ that is used to generate the functional classes $\{\mathcal{G}_k^g\}_{k=1}^{\infty}$ to approximate $f$ is the Gaussian density $g(\mathbf{x}) = \phi(\mathbf{x}; \mathbf{0}, \boldsymbol{\Sigma})$, and the functional class with $K$ components is given by

$$\mathcal{G}_K^g = \left\{ g : g(\mathbf{x}) = \sum_{k=1}^{K} \pi_k\,\phi(\mathbf{x}; \boldsymbol{\mu}_k, \lambda_k\boldsymbol{\Sigma}), \pi_k \geqslant 0, \sum_{k=1}^{K} \pi_k = 1, \boldsymbol{\mu}_k \in \mathbb{R}^n, \lambda_k > 0, k = 1, \ldots, K \right\}.$$

The Gaussian density $\phi(\mathbf{x}; \mathbf{0}, \boldsymbol{\Sigma})$ is a continuous pdf, and by Proposition 2, it vanishes at infinity and therefore belongs to $C_0(\mathbb{R}^n)$. Hence, the claim follows from Theorem 5 (a) in Nguyen, Nguyen, Chamroukhi, and McLachlan (2020). □

# B Estimation Details

This appendix provides the derivations of the conditional posterior distribution for $\boldsymbol{\theta}_k = (\boldsymbol{\mu}_k, \lambda)$ and the posterior predictive density under the proposed semiparametric VAR.

## B.1 Conditional Posterior Distribution for $\boldsymbol{\theta}_k$

Since the cluster-specific parameter $\boldsymbol{\theta}_k$ depends only on the residuals assigned to the $k$-th cluster, we relabel these residuals for convenience as $\boldsymbol{\varepsilon}_s \equiv \mathbf{y}_s - \mathbf{A}_1\mathbf{y}_{s-1} - \cdots - \mathbf{A}_p\mathbf{y}_{s-p}$ so that $s = 1, \ldots, T_k$, where $T_k$ denotes the size of the cluster. Define $\mathbf{E}_k = (\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_{T_k})$ and $\boldsymbol{\Sigma}_{1:T} = (\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_T)$. We also drop the subscript $k$ on the component-specific parameter and write $\boldsymbol{\theta}_k = (\boldsymbol{\mu}_k, \lambda_k)$ simply as $(\boldsymbol{\mu}, \lambda)$. Then, the conditional posterior for $(\boldsymbol{\mu}, \lambda)$ is given by:

$$
\begin{aligned}
p(\boldsymbol{\mu}, \lambda \,|\, \mathbf{E}_k, \boldsymbol{\Sigma}_{1:T}) &\propto p(\mathbf{E}_k \,|\, \boldsymbol{\mu}, \lambda, \boldsymbol{\Sigma}_{1:T}) p(\boldsymbol{\mu}, \lambda) \\
&= \prod_{s=1}^{T_k} (2\pi)^{-n/2} |\lambda \boldsymbol{\Sigma}_s|^{-1/2} \exp\left( -\frac{1}{2\lambda} (\boldsymbol{\varepsilon}_s - \boldsymbol{\mu})^\intercal \boldsymbol{\Sigma}_s^{-1} (\boldsymbol{\varepsilon}_s - \boldsymbol{\mu}) \right) \\
&\quad \times (2\pi)^{-n/2} |\mathbf{V}_{\boldsymbol{\mu}_0}|^{-1/2} \frac{S_0^{\nu_0}}{\Gamma(\nu_0)} \lambda^{-(\nu_0 + 1 + \frac{n}{2})} \exp\left( -\frac{1}{2\lambda} [2S_0 + (\boldsymbol{\mu} - \boldsymbol{\mu}_0)^\intercal \mathbf{V}_{\boldsymbol{\mu}_0}^{-1} (\boldsymbol{\mu} - \boldsymbol{\mu}_0)] \right) \\
&= (2\pi)^{-nT_k/2} \left( \prod_{s=1}^{T_k} |\boldsymbol{\Sigma}_s|^{-1/2} \right) \lambda^{-nT_k/2} \exp\left( -\frac{1}{2\lambda} \sum_{s=1}^{T_k} (\boldsymbol{\varepsilon}_s - \boldsymbol{\mu})^\intercal \boldsymbol{\Sigma}_s^{-1} (\boldsymbol{\varepsilon}_s - \boldsymbol{\mu}) \right) \\
&\quad \times (2\pi)^{-n/2} |\mathbf{V}_{\boldsymbol{\mu}_0}|^{-1/2} \frac{S_0^{\nu_0}}{\Gamma(\nu_0)} \lambda^{-(\nu_0 + 1 + \frac{n}{2})} \exp\left( -\frac{1}{2\lambda} [2S_0 + (\boldsymbol{\mu} - \boldsymbol{\mu}_0)^\intercal \mathbf{V}_{\boldsymbol{\mu}_0}^{-1} (\boldsymbol{\mu} - \boldsymbol{\mu}_0)] \right)
\end{aligned}
$$

To obtain a known kernel for the posterior, we focus on the exponential terms and complete the square in $\boldsymbol{\mu}$. Consider

$$
\begin{aligned}
&\sum_{s=1}^{T_k} (\boldsymbol{\varepsilon}_s - \boldsymbol{\mu})^\intercal \boldsymbol{\Sigma}_s^{-1} (\boldsymbol{\varepsilon}_s - \boldsymbol{\mu}) + (\boldsymbol{\mu} - \boldsymbol{\mu}_0)^\intercal \mathbf{V}_{\boldsymbol{\mu}_0}^{-1} (\boldsymbol{\mu} - \boldsymbol{\mu}_0) \\
&\quad = \sum_{s=1}^{T_k} \boldsymbol{\varepsilon}_s^\intercal \boldsymbol{\Sigma}_s^{-1} \boldsymbol{\varepsilon}_s - 2 \sum_{s=1}^{T_k} \boldsymbol{\mu}^\intercal \boldsymbol{\Sigma}_s^{-1} \boldsymbol{\varepsilon}_s + \sum_{s=1}^{T_k} \boldsymbol{\mu}^\intercal \boldsymbol{\Sigma}_s^{-1} \boldsymbol{\mu} + \boldsymbol{\mu}^\intercal \mathbf{V}_{\boldsymbol{\mu}_0}^{-1} \boldsymbol{\mu} - 2\boldsymbol{\mu}^\intercal \mathbf{V}_{\boldsymbol{\mu}_0}^{-1} \boldsymbol{\mu}_\mathbf{0} + \boldsymbol{\mu}_\mathbf{0}^\intercal \mathbf{V}_{\boldsymbol{\mu}_0}^{-1} \boldsymbol{\mu}_\mathbf{0} \\
&\quad = \boldsymbol{\mu}^\intercal (\mathbf{V}_{\boldsymbol{\mu}_0}^{-1} + \sum_{s=1}^{T_k} \boldsymbol{\Sigma}_s^{-1}) \boldsymbol{\mu} - 2\boldsymbol{\mu}^\intercal (\mathbf{V}_{\boldsymbol{\mu}_0}^{-1} \boldsymbol{\mu}_0 + \sum_{s=1}^{T_k} \boldsymbol{\Sigma}_s^{-1} \boldsymbol{\varepsilon}_s) + \sum_{s=1}^{T_k} \boldsymbol{\varepsilon}_s^\intercal \boldsymbol{\Sigma}_s^{-1} \boldsymbol{\varepsilon}_s + \boldsymbol{\mu}_\mathbf{0}^\intercal \mathbf{V}_{\boldsymbol{\mu}_0}^{-1} \boldsymbol{\mu}_\mathbf{0}.
\end{aligned}
$$

Define the updated parameters

$$\widehat{\mathbf{V}}_{\boldsymbol{\mu}_k} = \left( \mathbf{V}_{\boldsymbol{\mu}_0}^{-1} + \sum_{s=1}^{T_k} \boldsymbol{\Sigma}_s^{-1} \right)^{-1}, \quad \widehat{\boldsymbol{\mu}}_k = \widehat{\mathbf{V}}_{\boldsymbol{\mu}_k} \left( \mathbf{V}_{\boldsymbol{\mu}_0}^{-1} \boldsymbol{\mu}_0 + \sum_{s=1}^{T_k} \boldsymbol{\Sigma}_s^{-1} \boldsymbol{\varepsilon}_s \right).$$

Then the quadratic term becomes

$$(\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}}_k)^\intercal \widehat{\mathbf{V}}_{\boldsymbol{\mu}_k}^{-1} (\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}}_k) - \widehat{\boldsymbol{\mu}}_k^\intercal \widehat{\mathbf{V}}_{\boldsymbol{\mu}_k}^{-1} \widehat{\boldsymbol{\mu}}_k + \sum_{s=1}^{T_k} \boldsymbol{\varepsilon}_s^\intercal \boldsymbol{\Sigma}_s^{-1} \boldsymbol{\varepsilon}_s + \boldsymbol{\mu}_\mathbf{0}^\intercal \mathbf{V}_{\boldsymbol{\mu}_0}^{-1} \boldsymbol{\mu}_\mathbf{0}.$$

Next define the updated scale parameters

$$\widehat{\nu}_k = \nu_0 + \frac{nT_k}{2}, \quad \widehat{S}_k = S_0 + \frac{1}{2} \boldsymbol{\mu}_0^\intercal \mathbf{V}_{\boldsymbol{\mu}_0}^{-1} \boldsymbol{\mu}_0 + \frac{1}{2} \sum_{s=1}^{T_k} \boldsymbol{\varepsilon}_s^\intercal \boldsymbol{\Sigma}_s^{-1} \boldsymbol{\varepsilon}_s - \frac{1}{2} \boldsymbol{\mu}_k^\intercal \mathbf{V}_{\boldsymbol{\mu}_k}^{-1} \boldsymbol{\mu}_k.$$

Collecting the remaining terms, the conditional posterior distribution is

$$(\boldsymbol{\mu}, \lambda \,|\, \mathbf{E}_k, \boldsymbol{\Sigma}_{1:T}) \sim \mathcal{NIG}(\widehat{\boldsymbol{\mu}}_k, \widehat{\mathbf{V}}_{\boldsymbol{\mu}_k}, \widehat{\nu}_k, \widehat{S}_k).$$

## B.2 Posterior Predictive Density

We derive the posterior predictive density of a new residual $\boldsymbol{\varepsilon}^*$ for the $k$-th cluster, conditional on a new covariance matrix $\boldsymbol{\Sigma}^*$, which is one of the covariance matrices in $\boldsymbol{\Sigma}_{1:T}$:

$$(\boldsymbol{\varepsilon}^* \,|\, \mathbf{E}_k, \boldsymbol{\Sigma}_{1:T}, \boldsymbol{\Sigma}^*) = (\boldsymbol{\varepsilon}^* \,|\, \mathbf{E}_k, \boldsymbol{\Sigma}_{1:T}).$$

Let the conditional posterior distribution of $(\boldsymbol{\mu}, \lambda)$ from Online Appendix B.1 be

$$(\boldsymbol{\mu}, \lambda \,|\, \mathbf{E}_k, \boldsymbol{\Sigma}_{1:T}) \sim \mathcal{NIG}\left(\widehat{\boldsymbol{\mu}}_k, \widehat{\mathbf{V}}_{\boldsymbol{\mu}_k}, \widehat{\nu}_k, \widehat{S}_k\right).$$

That is,

$$(\lambda \,|\, \mathbf{E}_k, \boldsymbol{\Sigma}_{1:T}) \sim \mathcal{IG}(\widehat{\nu}_k, \widehat{S}_k), \quad (\boldsymbol{\mu} \,|\, \lambda, \mathbf{E}_k, \boldsymbol{\Sigma}_{1:T}) \sim \mathcal{N}(\widehat{\boldsymbol{\mu}}_k, \lambda \widehat{\mathbf{V}}_{\boldsymbol{\mu}_k}).$$

The posterior predictive density is obtained by integrating out the cluster-specific parameters:

$$p(\boldsymbol{\varepsilon}^* \,|\, \mathbf{E}_k, \boldsymbol{\Sigma}_{1:T}) = \int p(\boldsymbol{\varepsilon}^* \,|\, \boldsymbol{\mu}, \lambda, \mathbf{E}_k, \boldsymbol{\Sigma}^*) p(\boldsymbol{\mu}, \lambda \,|\, \mathbf{E}_k, \boldsymbol{\Sigma}_{1:T}) \mathrm{d}(\boldsymbol{\mu}, \lambda).$$

This can be computed by first integrating out $\boldsymbol{\mu}$ conditional on $\lambda$, followed by marginalizing over $\lambda$.

For the first step, observe that $(\boldsymbol{\varepsilon}^* \,|\, \boldsymbol{\mu}, \lambda, \mathbf{E}_k, \boldsymbol{\Sigma}^*) \sim \mathcal{N}(\boldsymbol{\mu}, \lambda\boldsymbol{\Sigma}^*)$, and since $(\boldsymbol{\mu} \,|\, \lambda, \mathbf{E}_k, \boldsymbol{\Sigma}_{1:T}) \sim \mathcal{N}(\widehat{\boldsymbol{\mu}}_k, \lambda\widehat{\mathbf{V}}_{\boldsymbol{\mu}_k})$, it follows by convolution of normals that

$$(\boldsymbol{\varepsilon}^* \,|\, \lambda, \mathbf{E}_k, \boldsymbol{\Sigma}_{1:T}) \sim \mathcal{N}(\widehat{\boldsymbol{\mu}}_k, \lambda(\boldsymbol{\Sigma}^* + \widehat{\mathbf{V}}_{\boldsymbol{\mu}_k})).$$

Since $(\lambda \,|\, \mathbf{E}_k, \boldsymbol{\Sigma}_{1:T}) \sim \mathcal{IG}(\widehat{\nu}_k, \widehat{S}_k)$, the joint distribution of $(\boldsymbol{\varepsilon}^*, \lambda)$ is also normal-inverse-gamma:

$$(\boldsymbol{\varepsilon}^*, \lambda \,|\, \mathbf{E}_k, \boldsymbol{\Sigma}_{1:T}) \sim \mathcal{NIG}\left(\widehat{\boldsymbol{\mu}}_k, \boldsymbol{\Sigma}^* + \widehat{\mathbf{V}}_{\boldsymbol{\mu}_k}, \widehat{\nu}_k, \widehat{S}_k\right).$$

To obtain the marginal distribution of $\boldsymbol{\varepsilon}^*$, we invoke the following lemma.

**Lemma 2.** *(Normal–inverse-gamma implies Student-t marginal) Suppose*

$$(\mathbf{X}, Y) \sim \mathcal{NIG}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \alpha, \beta).$$

*Then, the marginal distribution of* $\mathbf{X}$ *is the multivariate Student-t distribution* $\mathbf{X} \sim \mathcal{T}_{2\alpha}(\boldsymbol{\mu}, \beta\boldsymbol{\Sigma}/\alpha)$.

*Proof.* Recall that the normal-inverse-gamma density has the form:

$$f(\mathbf{x}, y) = (2\pi)^{-\frac{n}{2}}|\boldsymbol{\Sigma}|^{-\frac{1}{2}}\frac{\beta^\alpha}{\Gamma(\alpha)}y^{-(\alpha+1+n/2)}\mathrm{e}^{-\frac{1}{y}\left[\beta+\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]}.$$

The marginal distribution of $\mathbf{X}$ can be obtained by integrating out $Y$:

$$\begin{aligned}
f(\mathbf{x}) &= \int_0^\infty f(\mathbf{x}, y)\mathrm{d}y \\
&= (2\pi)^{-\frac{n}{2}}|\boldsymbol{\Sigma}|^{-\frac{1}{2}}\frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^\infty y^{-(\alpha+1+\frac{n}{2})}\mathrm{e}^{-\frac{1}{y}\left[\beta+\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]}\mathrm{d}y \\
&= (2\pi)^{-\frac{n}{2}}|\boldsymbol{\Sigma}|^{-\frac{1}{2}}\frac{\beta^\alpha}{\Gamma(\alpha)}\left[\beta+\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]^{-(\alpha+\frac{n}{2})}\Gamma\left(\alpha+\frac{n}{2}\right) \\
&= (2\pi)^{-\frac{n}{2}}|\boldsymbol{\Sigma}|^{-\frac{1}{2}}\frac{\Gamma\left(\alpha+\frac{n}{2}\right)}{\Gamma(\alpha)\beta^{\frac{n}{2}}}\left[1+\frac{1}{2\beta}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]^{-(\alpha+\frac{n}{2})} \\
&= \pi^{-\frac{n}{2}}\left|\frac{\beta}{\alpha}\boldsymbol{\Sigma}\right|^{-\frac{1}{2}}\frac{\Gamma\left(\alpha+\frac{n}{2}\right)}{\Gamma(\alpha)(2\alpha)^{\frac{n}{2}}}\left[1+\frac{1}{2\alpha}(\mathbf{x}-\boldsymbol{\mu})'\left(\frac{\beta}{\alpha}\boldsymbol{\Sigma}\right)^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]^{-(\alpha+\frac{n}{2})},
\end{aligned}$$

which is the density of the $\mathcal{T}_{2\alpha}(\boldsymbol{\mu}, \beta\boldsymbol{\Sigma}/\alpha)$ distribution. □

Using the above lemma, the posterior predictive density is a multivariate $t$ distribution:

$$(\boldsymbol{\varepsilon}^* \,|\, \mathbf{E}_k, \boldsymbol{\Sigma}_{1:T}) \sim \mathcal{T}_{2\widehat{\nu}_k}\left(\widehat{\boldsymbol{\mu}}_k, \frac{\widehat{S}_k}{\widehat{\nu}_k}(\Sigma^* + \widehat{\mathbf{V}}_{\boldsymbol{\mu}_k})\right).$$

We denote this posterior predictive density for cluster $k$ by

$$f_k(\cdot) \equiv p(\cdot \mid \mathbf{E}_k, \mathbf{\Sigma}_{1:T}),$$

which is the expression used in Section 3 for sampling the cluster assignment $z_t$.

# C Additional Forecasting Results

This appendix reports additional forecasting results, including CRPS and RMSE ratios for the competing models relative to DPM-OISV over the full sample and across the three subsamples: pre-COVID, COVID, and post-COVID.

Table 7: Ratios of CRPS for the competing models relative to DPM-OISV over the full sample at forecast horizons $h = 1, 6, 12$. Values below 1 favor DPM-OISV.

| Variable | SVo | | | SVO | | | SV-t | | | DPM-CSV | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 6 | 12 | 1 | 6 | 12 | 1 | 6 | 12 | 1 | 6 | 12 |
| Real Income | 1.15 | 0.89 | 0.86 | 1.12 | 0.90 | 0.91 | 1.13 | 1.04 | 1.31 | 1.09 | 1.04 | 1.29 |
| Real Income ex Transf | $0.78^{*}$ | $0.64^{*}$ | $0.41^{**}$ | $0.77^{*}$ | $0.65^{*}$ | $0.44^{*}$ | 0.97 | 1.01 | 1.00 | 1.03 | 1.04 | 0.98 |
| Real Consumption | 0.99 | $0.80^{***}$ | $0.56^{***}$ | 0.96 | $0.75^{**}$ | $0.54^{**}$ | 1.01 | $0.95^{**}$ | $0.94^{***}$ | 1.00 | $0.92^{***}$ | $0.91^{***}$ |
| IP | 0.88 | $0.86^{*}$ | $0.58^{***}$ | 0.90 | $0.83^{*}$ | $0.57^{**}$ | 1.00 | 0.99 | 0.97 | 1.01 | 0.98 | 0.95 |
| Capacity Util | 0.88 | $0.81^{*}$ | $0.63^{**}$ | 0.91 | $0.81^{*}$ | $0.62^{**}$ | 0.98 | 0.95 | 0.98 | 0.98 | 0.90 | 0.89 |
| Civilian Employment | $0.88^{*}$ | $0.84^{**}$ | 0.79 | $0.90^{*}$ | $0.81^{*}$ | $0.77^{*}$ | 0.99 | 0.99 | 1.20 | 1.00 | 1.00 | 1.18 |
| Unemployment Rate | 0.93 | 1.19 | 1.07 | 0.96 | 1.24 | 1.08 | 1.03 | 1.28 | 1.25 | 1.13 | 1.24 | 1.15 |
| Mean Unemp Duration | 0.89 | 1.15 | 1.12 | 0.89 | 1.14 | 1.12 | 0.95 | 1.12 | 1.26 | 0.98 | 1.18 | 1.28 |
| Initial Claims | 1.23 | $0.87^{***}$ | 1.77 | 1.23 | $0.84^{**}$ | 1.75 | 1.35 | 1.00 | 2.67 | 1.35 | 0.99 | 2.67 |
| Nonfarm Payrolls | 1.28 | 0.94 | 0.94 | 1.32 | 0.92 | 0.93 | 1.41 | 1.09 | 1.41 | 1.43 | 1.07 | 1.36 |
| Hours | 0.84 | 0.91 | $0.66^{*}$ | 0.86 | 0.90 | $0.66^{*}$ | 0.95 | 1.06 | 1.08 | 0.94 | 0.97 | 0.95 |
| Hours: Manufacturing | $0.85^{*}$ | $0.84^{*}$ | $0.62^{**}$ | $0.86^{*}$ | $0.83^{*}$ | $0.61^{*}$ | 0.90 | 0.94 | 1.00 | 0.90 | 0.88 | 0.90 |
| Housing Starts | $0.83^{*}$ | 0.87 | 0.71 | $0.82^{*}$ | 0.87 | 0.71 | 0.90 | 0.93 | 0.88 | 0.91 | 0.97 | 0.90 |
| Building Permits | $0.84^{*}$ | 0.82 | $0.67^{*}$ | $0.84^{*}$ | 0.83 | $0.68^{*}$ | 0.96 | 0.95 | 0.90 | 0.96 | 0.97 | 0.89 |
| Real Mfg & Trade Sales | 0.96 | $0.76^{***}$ | $0.52^{***}$ | 0.94 | $0.73^{**}$ | $0.50^{**}$ | 1.06 | 0.93 | $0.95^{**}$ | 1.07 | 0.95 | $0.91^{***}$ |
| Durable Goods Orders | $0.81^{*}$ | $0.77^{**}$ | $0.60^{***}$ | $0.80^{*}$ | $0.75^{**}$ | $0.59^{**}$ | 0.92 | $0.92^{*}$ | 0.96 | 0.96 | $0.94^{**}$ | $0.97^{*}$ |
| Business Inventories | $0.76^{***}$ | 0.82 | $0.55^{***}$ | $0.75^{***}$ | 0.79 | $0.55^{***}$ | $0.89^{**}$ | 1.03 | 1.09 | 0.95 | 1.04 | 1.09 |
| M1 Money Supply | 1.72 | 1.25 | 1.20 | 1.72 | 1.20 | 1.16 | 1.75 | 1.43 | 1.65 | 1.91 | 1.45 | 1.65 |
| M2 Money Supply | $0.76^{**}$ | $0.68^{**}$ | $0.49^{***}$ | $0.76^{**}$ | $0.64^{**}$ | $0.48^{**}$ | 0.88 | $0.87^{*}$ | 0.90 | 0.99 | 0.93 | $0.90^{*}$ |
| Business Loans | $0.84^{*}$ | 0.91 | $0.63^{**}$ | $0.84^{*}$ | 0.88 | $0.62^{*}$ | $0.89^{*}$ | 1.06 | 1.06 | 0.94 | 1.04 | 1.04 |
| Consumer Credit (NR) | $0.74^{***}$ | $0.65^{***}$ | $0.36^{***}$ | $0.76^{***}$ | $0.68^{**}$ | $0.40^{***}$ | $0.82^{***}$ | $0.84^{*}$ | $0.82^{*}$ | $0.86^{***}$ | $0.87^{**}$ | $0.82^{***}$ |
| Fed Funds Rate | $0.74^{*}$ | $0.70^{*}$ | $0.48^{***}$ | $0.76^{**}$ | $0.68^{*}$ | $0.48^{***}$ | $0.62^{***}$ | $0.62^{***}$ | $0.69^{*}$ | $0.70^{***}$ | $0.64^{***}$ | $0.71^{**}$ |
| 3-Mo T-Bill | $0.50^{***}$ | $0.62^{**}$ | $0.55^{**}$ | $0.51^{***}$ | $0.61^{**}$ | $0.55^{**}$ | $0.49^{***}$ | $0.61^{***}$ | $0.74^{*}$ | $0.50^{***}$ | $0.59^{***}$ | $0.72^{*}$ |
| 1-Year Treasury | $0.56^{***}$ | $0.63^{*}$ | $0.54^{**}$ | $0.57^{***}$ | $0.64^{*}$ | $0.55^{**}$ | $0.52^{***}$ | $0.64^{**}$ | $0.74^{*}$ | $0.54^{***}$ | $0.64^{***}$ | $0.72^{*}$ |
| 5-Year Treasury | $0.81^{**}$ | 0.87 | $0.69^{*}$ | $0.83^{*}$ | 0.88 | $0.70^{*}$ | $0.81^{*}$ | 0.90 | 0.96 | $0.84^{***}$ | $0.87^{*}$ | 0.89 |
| 10-Year Treasury | $0.84^{*}$ | 0.94 | $0.73^{*}$ | $0.86^{*}$ | 0.94 | 0.73 | 0.84 | 0.95 | 0.99 | $0.91^{*}$ | 0.94 | 0.95 |
| Baa Corporate | $0.86^{**}$ | 0.80 | $0.61^{**}$ | $0.83^{**}$ | 0.79 | $0.62^{*}$ | 0.93 | 0.89 | 0.90 | 0.96 | 0.93 | 0.93 |
| Baa Spread | 0.90 | $0.84^{*}$ | $0.52^{***}$ | 0.89 | $0.81^{*}$ | $0.52^{**}$ | $0.87^{*}$ | $0.81^{*}$ | 0.85 | 0.93 | $0.81^{**}$ | $0.84^{**}$ |
| AAA Spread | $0.82^{**}$ | $0.75^{*}$ | $0.47^{***}$ | $0.78^{***}$ | $0.74^{*}$ | $0.47^{***}$ | $0.85^{**}$ | $0.76^{*}$ | 0.82 | $0.87^{**}$ | $0.74^{*}$ | $0.79^{*}$ |
| 5Y-FF Spread | 0.87 | $0.79^{*}$ | $0.45^{***}$ | $0.85^{*}$ | $0.77^{*}$ | $0.46^{***}$ | $0.87^{*}$ | $0.76^{*}$ | 0.81 | $0.89^{**}$ | $0.76^{*}$ | $0.78^{*}$ |
| 10Y-FF Spread | $0.86^{*}$ | $0.76^{*}$ | $0.45^{***}$ | $0.85^{*}$ | $0.73^{*}$ | $0.45^{***}$ | $0.88^{*}$ | $0.76^{*}$ | 0.81 | 0.92 | $0.76^{*}$ | $0.80^{*}$ |
| PPI: Fin Goods | 0.91 | $0.86^{**}$ | $0.62^{***}$ | 0.91 | $0.84^{**}$ | $0.63^{***}$ | 0.94 | 0.98 | 0.99 | 0.97 | 1.01 | 1.00 |
| PCE Prices | $0.84^{*}$ | $0.78^{***}$ | $0.49^{***}$ | $0.83^{**}$ | $0.76^{***}$ | $0.50^{***}$ | 0.90 | 0.96 | 0.94 | 0.95 | 0.99 | 0.94 |
| CPI: All Items | $0.89^{*}$ | $0.78^{***}$ | $0.48^{***}$ | $0.87^{*}$ | $0.76^{**}$ | $0.49^{***}$ | 0.93 | 0.95 | 0.94 | 0.96 | 0.97 | 0.94 |
| Oil Price | $0.87^{*}$ | $0.82^{**}$ | $0.57^{***}$ | $0.86^{*}$ | $0.80^{*}$ | $0.57^{**}$ | 0.91 | $0.94^{*}$ | $0.97^{*}$ | 0.94 | $0.96^{*}$ | $0.97^{**}$ |
| PPI: Metals & Metal Prod | $0.85^{**}$ | $0.82^{**}$ | $0.55^{***}$ | $0.84^{*}$ | $0.80^{**}$ | $0.55^{***}$ | $0.89^{**}$ | $0.91^{*}$ | 0.94 | $0.92^{**}$ | $0.92^{*}$ | $0.93^{*}$ |
| Hourly Earnings | $0.61^{***}$ | $0.61^{***}$ | $0.35^{***}$ | $0.60^{***}$ | $0.59^{***}$ | $0.35^{***}$ | $0.70^{**}$ | $0.82^{***}$ | $0.82^{***}$ | $0.70^{***}$ | $0.84^{***}$ | $0.81^{***}$ |
| S&P 500 | $0.70^{*}$ | $0.76^{**}$ | $0.49^{***}$ | $0.71^{*}$ | $0.73^{*}$ | $0.48^{**}$ | $0.77^{*}$ | 0.93 | $0.95^{*}$ | $0.83^{*}$ | 0.96 | $0.94^{***}$ |
| S&P Div Yield | $0.58^{***}$ | $0.61^{***}$ | $0.45^{***}$ | $0.58^{***}$ | $0.59^{***}$ | $0.44^{***}$ | $0.67^{*}$ | $0.73^{*}$ | $0.72^{*}$ | $0.69^{***}$ | $0.75^{**}$ | $0.76^{*}$ |
| VIX | $0.76^{***}$ | $0.80^{*}$ | $0.55^{***}$ | $0.75^{***}$ | $0.76^{*}$ | $0.54^{**}$ | $0.85^{***}$ | 0.93 | 0.91 | $0.89^{**}$ | 0.99 | 0.96 |
| USD/GBP FX | $0.70^{***}$ | $0.69^{**}$ | $0.43^{***}$ | $0.69^{**}$ | $0.69^{**}$ | $0.44^{***}$ | $0.79^{**}$ | $0.87^{*}$ | $0.91^{***}$ | $0.83^{**}$ | $0.90^{***}$ | $0.90^{***}$ |
| JPY/USD FX | $0.67^{***}$ | $0.71^{**}$ | $0.46^{***}$ | $0.68^{***}$ | $0.69^{**}$ | $0.46^{***}$ | $0.75^{***}$ | $0.89^{*}$ | 0.93 | $0.80^{***}$ | $0.91^{**}$ | $0.92^{**}$ |

Each entry is the ratio of the DPM-OISV forecast score to that of the indicated benchmark model. Values below 1 favor DPM-OISV, while values above 1 favor the benchmark. Asterisks denote Diebold-Mariano-West tests based on Newey-West HAC standard errors: $^{***}$ $p < 0.001$, $^{**}$ $0.001 \leqslant p < 0.01$, and $^{*}$ $0.01 \leqslant p < 0.05$.

Table 8: Ratios of RMSE for the competing models relative to DPM-OISV over the full sample at forecast horizons $h = 1, 6, 12$. Values below 1 favor DPM-OISV.

| Variable | SVo | | | SVO | | | SV-t | | | DPM-CSV | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 6 | 12 | 1 | 6 | 12 | 1 | 6 | 12 | 1 | 6 | 12 |
| Real Income | 1.10 | 0.98 | 2.88 | 1.11 | 1.01 | 2.86 | 1.12 | 1.03 | 2.97 | 1.07 | 1.04 | 2.96 |
| Real Income ex Transf | 0.95 | 0.85 | 0.58 | 0.89 | 0.92 | 0.87 | 0.94 | 1.01 | 1.00 | 0.98 | 1.03 | 1.00 |
| Real Consumption | 1.03 | 0.98 | 0.99 | 1.04 | 0.97 | 0.99 | 0.99 | 1.00 | 1.00 | 1.00 | $0.99^{**}$ | 1.00 |
| IP | 0.88 | 0.93 | 0.98 | 0.89 | 0.95 | 0.99 | 0.99 | 1.00 | 1.03 | 1.05 | 1.00 | 1.03 |
| Capacity Util | 0.86 | 0.92 | 0.91 | 0.91 | 0.96 | 0.92 | 0.93 | 0.98 | 1.08 | 0.99 | 1.00 | 1.07 |
| Civilian Employment | 0.95 | 0.99 | 3.27 | 0.96 | 0.99 | 3.28 | 1.01 | 1.00 | 3.31 | 1.02 | 1.01 | 3.30 |
| Unemployment Rate | 0.97 | 1.38 | 2.53 | 1.00 | 1.40 | 2.53 | 1.03 | 1.44 | 2.66 | 1.14 | 1.61 | 2.89 |
| Mean Unemp Duration | 0.79 | 1.05 | 1.18 | 0.78 | 1.05 | 1.26 | 0.96 | 1.11 | 1.31 | 1.04 | 1.21 | 1.42 |
| Initial Claims | 1.31 | 0.99 | 6.25 | 1.31 | 0.99 | 6.26 | 1.38 | 1.00 | 6.28 | 1.41 | 1.00 | 6.29 |
| Nonfarm Payrolls | 1.38 | 1.02 | 7.29 | 1.38 | 1.02 | 7.30 | 1.42 | 1.04 | 7.32 | 1.45 | 1.04 | 7.32 |
| Hours | 0.75 | 0.98 | 0.94 | 0.75 | 1.01 | 0.97 | 0.94 | 1.05 | 1.10 | 0.93 | 0.97 | 0.99 |
| Hours: Manufacturing | 0.99 | 0.88 | 0.91 | 1.00 | 0.91 | 0.89 | 1.01 | 0.96 | 1.04 | 0.99 | 0.91 | 0.98 |
| Housing Starts | 0.75 | 0.82 | 0.84 | 0.75 | 0.85 | 0.83 | 0.88 | 0.84 | 0.80 | 0.89 | 0.96 | 0.89 |
| Building Permits | 0.82 | 0.69 | 0.79 | 0.81 | 0.70 | 0.77 | 0.89 | 0.75 | 0.77 | 1.00 | 0.94 | 0.87 |
| Real Mfg & Trade Sales | 1.03 | 0.86 | 0.94 | 1.05 | 0.87 | 0.93 | 1.13 | 0.94 | 1.01 | 1.15 | 1.01 | 0.99 |
| Durable Goods Orders | 0.67 | 0.89 | 0.92 | 0.64 | 0.88 | 0.94 | 0.88 | 0.96 | 1.01 | 0.96 | 0.97 | 1.03 |
| Business Inventories | $0.64^{**}$ | 0.77 | 1.00 | $0.64^{**}$ | 0.74 | 0.97 | $0.78^{*}$ | 0.99 | 1.10 | 0.93 | 1.02 | 1.13 |
| M1 Money Supply | 1.58 | 1.18 | 8.01 | 1.58 | 1.18 | 7.96 | 1.56 | 1.19 | 8.06 | 1.59 | 1.19 | 8.06 |
| M2 Money Supply | 0.67 | 0.75 | 0.76 | 0.67 | 0.72 | 0.75 | 0.74 | 0.88 | $0.94^{*}$ | 0.89 | 0.96 | 0.96 |
| Business Loans | 0.76 | 0.99 | 0.98 | 0.77 | 0.98 | 0.96 | 0.83 | 1.02 | 1.04 | $0.93^{*}$ | 1.00 | 1.03 |
| Consumer Credit (NR) | $0.86^{**}$ | 0.84 | 0.57 | $0.84^{*}$ | 0.81 | 0.73 | $0.90^{*}$ | 0.92 | 0.88 | 0.93 | 0.98 | 0.95 |
| Fed Funds Rate | $0.56^{*}$ | $0.59^{*}$ | 0.52 | $0.64^{*}$ | 0.57 | $0.56^{*}$ | $0.53^{*}$ | $0.54^{*}$ | $0.63^{*}$ | $0.71^{**}$ | $0.60^{*}$ | $0.73^{*}$ |
| 3-Mo T-Bill | $0.40^{*}$ | 0.54 | 0.63 | $0.46^{*}$ | 0.55 | 0.68 | $0.45^{*}$ | $0.58^{**}$ | $0.73^{*}$ | $0.54^{***}$ | $0.60^{**}$ | $0.77^{*}$ |
| 1-Year Treasury | $0.44^{**}$ | 0.49 | 0.60 | $0.47^{**}$ | 0.52 | 0.65 | 0.39 | $0.59^{*}$ | $0.72^{*}$ | $0.49^{*}$ | $0.65^{***}$ | $0.76^{*}$ |
| 5-Year Treasury | $0.67^{*}$ | 0.76 | 0.82 | $0.73^{*}$ | 0.81 | 0.85 | 0.60 | 0.83 | 0.94 | 0.74 | $0.87^{*}$ | 0.92 |
| 10-Year Treasury | 0.75 | 0.93 | 0.89 | 0.83 | 0.99 | 0.92 | 0.64 | 0.89 | 0.95 | 0.83 | 0.93 | 0.97 |
| Baa Corporate | $0.83^{*}$ | 0.69 | 0.60 | $0.78^{*}$ | 0.69 | 0.63 | 0.93 | 0.81 | 0.79 | 0.97 | 0.92 | 0.93 |
| Baa Spread | 0.86 | $0.86^{*}$ | 0.70 | 0.84 | 0.82 | 0.58 | 0.78 | 0.71 | 0.83 | 0.91 | 0.74 | 0.88 |
| AAA Spread | 0.82 | $0.75^{*}$ | 0.62 | $0.74^{*}$ | $0.73^{*}$ | 0.63 | $0.82^{*}$ | 0.68 | 0.77 | $0.87^{*}$ | 0.67 | 0.79 |
| 5Y-FF Spread | 0.83 | $0.80^{*}$ | $0.61^{*}$ | $0.82^{*}$ | 0.73 | 0.65 | 0.80 | 0.64 | 0.77 | $0.90^{*}$ | 0.69 | 0.79 |
| 10Y-FF Spread | 0.81 | $0.77^{*}$ | 0.61 | $0.83^{*}$ | 0.68 | $0.59^{*}$ | 0.84 | 0.66 | 0.78 | 0.96 | 0.68 | 0.81 |
| PPI: Fin Goods | 0.82 | $0.93^{*}$ | 0.93 | 0.82 | $0.93^{**}$ | 0.96 | 0.87 | 0.98 | 0.99 | 0.94 | 1.00 | 1.02 |
| PCE Prices | $0.77^{*}$ | 0.91 | 0.87 | $0.79^{*}$ | $0.91^{*}$ | 0.85 | 0.82 | 0.97 | 0.97 | 0.94 | 1.01 | 1.02 |
| CPI: All Items | 0.88 | 0.90 | $0.85^{*}$ | 0.87 | 0.90 | 0.84 | 0.84 | 0.96 | 0.96 | 0.95 | 0.99 | 1.00 |
| Oil Price | 0.83 | $0.92^{*}$ | 0.91 | 0.78 | $0.94^{**}$ | 0.96 | 0.86 | $0.96^{*}$ | 0.99 | 0.94 | 0.97 | 0.99 |
| PPI: Metals & Metal Prod | $0.81^{*}$ | 0.92 | 0.96 | 0.76 | $0.92^{*}$ | 0.91 | $0.90^{*}$ | $0.91^{*}$ | 0.95 | $0.92^{*}$ | 0.93 | 0.95 |
| Hourly Earnings | $0.44^{*}$ | 0.72 | 0.58 | 0.42 | 0.75 | 0.63 | 0.51 | $0.85^{*}$ | 0.80 | 0.55 | 0.92 | $0.90^{*}$ |
| S&P 500 | 0.51 | 0.87 | $0.83^{*}$ | 0.49 | 0.86 | $0.90^{*}$ | 0.55 | 0.93 | 0.97 | 0.73 | 0.98 | 0.97 |
| S&P Div Yield | 0.46 | 0.62 | 0.66 | 0.45 | 0.59 | 0.61 | 0.50 | 0.67 | 0.67 | 0.63 | 0.79 | 0.79 |
| VIX | $0.77^{*}$ | 0.86 | $0.79^{*}$ | $0.78^{*}$ | $0.88^{*}$ | $0.83^{*}$ | $0.86^{**}$ | 0.93 | 0.93 | $0.89^{*}$ | 0.99 | 0.97 |
| USD/GBP FX | $0.54^{*}$ | 0.69 | 0.82 | 0.51 | 0.70 | $0.84^{**}$ | 0.65 | 0.85 | $0.95^{**}$ | $0.76^{*}$ | $0.93^{**}$ | 0.97 |
| JPY/USD FX | 0.53 | 0.78 | 0.82 | 0.54 | 0.78 | $0.88^{**}$ | 0.65 | 0.90 | 0.95 | $0.79^{***}$ | 0.94 | 0.96 |

Each entry is the ratio of the DPM-OISV RMSE to that of the indicated benchmark model. Values below 1 favor DPM-OISV, while values above 1 favor the benchmark. Asterisks denote Diebold-Mariano-West tests based on Newey-West HAC standard errors: $^{***}$ $p < 0.001$, $^{**}$ $0.001 \leqslant p < 0.01$, and $^{*}$ $0.01 \leqslant p < 0.05$.

Table 9: Ratios of CRPS for the competing models relative to DPM-OISV over the pre-COVID sample at forecast horizons $h = 1, 6, 12$. Values below 1 favor DPM-OISV.

| Variable | SVo | | | SVO | | | SV-$t$ | | | DPM-CSV | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 6 | 12 | 1 | 6 | 12 | 1 | 6 | 12 | 1 | 6 | 12 |
| Real Income | 1.00 | $0.92^{*}$ | $0.75^{***}$ | 1.01 | 0.95 | $0.81^{**}$ | 1.01 | 1.03 | 1.00 | 1.02 | 1.04 | 1.00 |
| Real Income ex Transf | 0.98 | $0.89^{*}$ | $0.61^{***}$ | 1.01 | 0.94 | $0.70^{***}$ | 1.04 | 1.06 | 1.02 | 1.05 | 1.06 | 0.99 |
| Real Consumption | $0.90^{**}$ | $0.87^{***}$ | $0.66^{***}$ | $0.90^{**}$ | $0.88^{***}$ | $0.70^{***}$ | 0.95 | $0.96^{**}$ | $0.95^{**}$ | $0.89^{**}$ | $0.92^{***}$ | $0.91^{***}$ |
| IP | 1.02 | 0.97 | $0.69^{***}$ | 1.03 | 0.96 | $0.72^{***}$ | 1.12 | 1.02 | 1.00 | 1.10 | 0.99 | 0.96 |
| Capacity Util | 0.94 | 0.91 | 0.75 | 0.93 | 0.91 | 0.77 | 1.07 | 1.04 | 1.02 | 1.03 | 0.95 | 0.91 |
| Civilian Employment | 0.93 | $0.92^{***}$ | $0.76^{***}$ | 0.93 | $0.92^{***}$ | $0.78^{***}$ | 0.98 | 0.99 | 0.98 | 0.94 | 0.97 | $0.95^{*}$ |
| Unemployment Rate | 0.99 | 0.98 | 0.91 | 1.00 | 0.98 | 0.92 | 1.05 | 1.04 | 1.01 | 1.02 | 0.95 | 0.89 |
| Mean Unemp Duration | 1.01 | 1.22 | 1.26 | 1.01 | 1.21 | 1.28 | 0.97 | 1.11 | 1.24 | $0.95^{*}$ | 1.12 | 1.20 |
| Initial Claims | 0.97 | $0.92^{**}$ | $0.73^{***}$ | 0.96 | $0.93^{**}$ | $0.76^{***}$ | 1.00 | $0.98^{*}$ | $0.98^{*}$ | 0.97 | $0.97^{**}$ | $0.96^{*}$ |
| Nonfarm Payrolls | 0.97 | 0.93 | $0.76^{***}$ | 0.98 | $0.92^{*}$ | $0.78^{**}$ | 1.07 | 0.99 | 0.97 | 1.01 | $0.93^{*}$ | $0.91^{*}$ |
| Hours | 0.94 | 0.99 | $0.79^{**}$ | 0.96 | 1.01 | $0.83^{*}$ | 0.99 | 1.06 | 1.07 | $0.93^{*}$ | 0.94 | 0.94 |
| Hours: Manufacturing | 0.92 | 0.93 | $0.75^{***}$ | 0.94 | 0.94 | $0.78^{**}$ | 0.94 | 0.97 | 0.99 | $0.89^{*}$ | 0.88 | 0.89 |
| Housing Starts | 1.02 | 1.09 | 0.93 | 1.01 | 1.10 | 0.96 | 1.01 | 1.05 | 1.00 | 1.00 | 1.04 | 0.98 |
| Building Permits | 1.05 | 1.06 | 0.88 | 1.05 | 1.07 | 0.91 | 1.13 | 1.14 | 1.04 | 1.04 | 1.08 | 0.99 |
| Real Mfg & Trade Sales | 0.99 | $0.88^{**}$ | $0.62^{***}$ | 0.99 | $0.88^{**}$ | $0.65^{***}$ | 1.06 | 0.99 | 0.96 | 1.03 | 0.96 | $0.91^{***}$ |
| Durable Goods Orders | 1.00 | $0.90^{***}$ | $0.72^{***}$ | 0.99 | $0.90^{***}$ | $0.75^{***}$ | 1.03 | 0.98 | 1.00 | 1.03 | 0.98 | 0.99 |
| Business Inventories | 0.91 | 0.93 | $0.63^{***}$ | $0.91^{*}$ | 0.91 | $0.66^{***}$ | 1.00 | 1.05 | 1.03 | 0.98 | 1.04 | 1.02 |
| M1 Money Supply | 1.02 | $0.94^{***}$ | $0.82^{***}$ | 1.02 | $0.94^{***}$ | $0.83^{***}$ | 0.96 | $0.97^{*}$ | $0.95^{*}$ | 1.01 | 0.97 | 0.95 |
| M2 Money Supply | $0.94^{*}$ | $0.81^{***}$ | $0.60^{***}$ | $0.93^{*}$ | $0.80^{***}$ | $0.62^{***}$ | 1.01 | 0.93 | 0.94 | 1.04 | 0.92 | 0.94 |
| Business Loans | 0.96 | 0.99 | $0.77^{**}$ | 0.95 | 0.98 | $0.80^{*}$ | 0.99 | 1.10 | 1.11 | 0.97 | 1.11 | 1.10 |
| Consumer Credit (NR) | $0.87^{***}$ | $0.77^{***}$ | $0.42^{***}$ | $0.90^{*}$ | $0.83^{***}$ | $0.49^{***}$ | $0.93^{*}$ | 0.93 | 0.92 | $0.92^{*}$ | 0.91 | $0.87^{**}$ |
| Fed Funds Rate | 0.99 | $0.83^{*}$ | $0.56^{***}$ | 0.93 | $0.80^{*}$ | $0.57^{***}$ | $0.75^{***}$ | $0.66^{**}$ | $0.72^{*}$ | $0.72^{***}$ | $0.64^{***}$ | $0.70^{**}$ |
| 3-Mo T-Bill | $0.70^{***}$ | $0.71^{**}$ | $0.61^{**}$ | $0.70^{***}$ | $0.70^{**}$ | $0.63^{**}$ | $0.61^{***}$ | $0.62^{**}$ | $0.74^{*}$ | $0.56^{***}$ | $0.57^{***}$ | $0.69^{*}$ |
| 1-Year Treasury | $0.69^{***}$ | $0.72^{***}$ | $0.61^{***}$ | $0.71^{***}$ | $0.72^{***}$ | $0.63^{**}$ | $0.63^{***}$ | $0.65^{***}$ | $0.73^{*}$ | $0.55^{***}$ | $0.59^{***}$ | $0.67^{**}$ |
| 5-Year Treasury | 0.94 | 0.98 | $0.80^{*}$ | 0.93 | 1.01 | $0.84^{*}$ | $0.93^{*}$ | 0.94 | 0.98 | $0.86^{***}$ | $0.83^{**}$ | $0.84^{*}$ |
| 10-Year Treasury | 0.95 | 1.01 | $0.83^{**}$ | 0.93 | 1.05 | $0.86^{*}$ | 0.97 | 0.99 | 1.04 | $0.92^{*}$ | 0.91 | 0.91 |
| Baa Corporate | 0.95 | $0.89^{**}$ | $0.71^{***}$ | $0.93^{*}$ | $0.91^{*}$ | $0.74^{***}$ | 0.97 | 0.95 | 1.00 | $0.92^{*}$ | $0.89^{*}$ | 0.90 |
| Baa Spread | 1.02 | 0.99 | $0.67^{***}$ | 0.98 | 0.96 | $0.68^{***}$ | 0.95 | 0.88 | 0.91 | 0.92 | 0.88 | 0.90 |
| AAA Spread | 0.94 | 0.93 | $0.62^{***}$ | $0.89^{*}$ | 0.90 | $0.63^{***}$ | 0.93 | 0.84 | 0.87 | $0.85^{**}$ | 0.80 | 0.84 |
| 5Y-FF Spread | 0.97 | $0.88^{*}$ | $0.56^{***}$ | $0.89^{*}$ | 0.88 | $0.58^{***}$ | 0.95 | 0.81 | 0.85 | $0.88^{**}$ | 0.78 | 0.82 |
| 10Y-FF Spread | 0.96 | 0.90 | $0.58^{***}$ | $0.90^{*}$ | 0.88 | $0.59^{***}$ | 0.96 | 0.83 | 0.87 | $0.88^{*}$ | 0.80 | 0.86 |
| PPI: Fin Goods | 1.05 | $0.91^{*}$ | $0.68^{***}$ | 1.04 | $0.92^{**}$ | $0.71^{***}$ | 1.01 | 0.97 | 0.94 | 1.01 | 0.97 | 0.96 |
| PCE Prices | 0.97 | $0.84^{**}$ | $0.54^{***}$ | 0.96 | $0.83^{***}$ | $0.57^{***}$ | 1.04 | 0.96 | 0.88 | 1.01 | 0.95 | 0.89 |
| CPI: All Items | 1.01 | $0.85^{**}$ | $0.54^{***}$ | 0.99 | $0.85^{***}$ | $0.57^{***}$ | 1.06 | 0.96 | $0.88^{*}$ | 1.01 | 0.94 | $0.89^{*}$ |
| Oil Price | 0.95 | $0.92^{***}$ | $0.69^{***}$ | 0.95 | $0.93^{***}$ | $0.74^{***}$ | 0.97 | 0.97 | 0.98 | 0.95 | 0.97 | 0.98 |
| PPI: Metals & Metal Prod | 0.99 | $0.92^{**}$ | $0.64^{***}$ | 0.99 | $0.91^{***}$ | $0.66^{***}$ | 0.97 | 0.93 | 0.94 | 0.96 | $0.92^{*}$ | $0.92^{*}$ |
| Hourly Earnings | $0.79^{***}$ | $0.71^{***}$ | $0.40^{***}$ | $0.78^{***}$ | $0.71^{***}$ | $0.42^{***}$ | $0.84^{***}$ | $0.87^{***}$ | $0.83^{***}$ | $0.81^{***}$ | $0.85^{***}$ | $0.78^{***}$ |
| S&P 500 | $0.90^{*}$ | $0.88^{***}$ | $0.58^{***}$ | $0.92^{*}$ | $0.88^{***}$ | $0.60^{***}$ | 0.94 | 0.99 | $0.96^{*}$ | $0.89^{*}$ | 0.97 | $0.93^{***}$ |
| S&P Div Yield | $0.78^{***}$ | $0.73^{***}$ | $0.54^{***}$ | $0.79^{***}$ | $0.72^{***}$ | $0.56^{***}$ | $0.88^{**}$ | 0.84 | 0.79 | $0.78^{***}$ | 0.80 | 0.80 |
| VIX | $0.87^{***}$ | $0.91^{*}$ | $0.66^{***}$ | $0.88^{***}$ | $0.91^{*}$ | $0.68^{***}$ | $0.91^{**}$ | 0.96 | 0.90 | $0.88^{**}$ | 1.00 | 0.95 |
| USD/GBP FX | $0.91^{**}$ | $0.84^{***}$ | $0.52^{***}$ | $0.90^{**}$ | $0.85^{***}$ | $0.56^{***}$ | 0.94 | $0.95^{*}$ | $0.93^{**}$ | $0.92^{*}$ | $0.93^{**}$ | $0.91^{***}$ |
| JPY/USD FX | $0.84^{***}$ | $0.82^{***}$ | $0.52^{***}$ | $0.83^{***}$ | $0.82^{***}$ | $0.54^{***}$ | $0.87^{***}$ | 0.96 | 0.94 | $0.85^{***}$ | $0.94^{*}$ | $0.91^{**}$ |

Each entry is the ratio of the DPM-OISV forecast score to that of the indicated benchmark model.

Values below 1 favor DPM-OISV, while values above 1 favor the benchmark. Asterisks denote Diebold-Mariano-West tests based on Newey-West HAC standard errors: $^{***}$ $p < 0.001$, $^{**}$ $0.001 \leqslant p < 0.01$, and $^{*}$ $0.01 \leqslant p < 0.05$.

Table 10: Ratios of CRPS for the competing models relative to DPM-OISV over the COVID sample at forecast horizons $h = 1, 6, 12$. Values below 1 favor DPM-OISV.

| Variable | SVo | | | SVO | | | SV-$t$ | | | DPM-CSV | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 6 | 12 | 1 | 6 | 12 | 1 | 6 | 12 | 1 | 6 | 12 |
| Real Income | 1.28 | 0.87 | 1.08 | 1.22 | 0.85 | 1.12 | 1.23 | 1.05 | 2.03 | 1.13 | 1.03 | 1.96 |
| Real Income ex Transf | $0.44^{*}$ | $0.19^{**}$ | $0.10^{***}$ | $0.42^{**}$ | $0.18^{***}$ | $0.10^{***}$ | 0.81 | $0.81^{*}$ | $0.80^{*}$ | 0.96 | 1.08 | 0.95 |
| Real Consumption | 1.14 | $0.56^{**}$ | $0.26^{***}$ | 1.04 | $0.40^{*}$ | $0.20^{**}$ | 1.13 | 0.95 | $0.85^{***}$ | 1.21 | 1.01 | 0.97 |
| IP | 0.71 | $0.46^{***}$ | $0.22^{***}$ | 0.74 | $0.39^{**}$ | $0.18^{***}$ | 0.85 | 0.83 | 0.75 | 0.92 | 0.89 | 0.88 |
| Capacity Util | 0.80 | $0.61^{**}$ | $0.37^{***}$ | 0.87 | $0.58^{**}$ | $0.32^{***}$ | 0.86 | $0.70^{*}$ | 0.95 | 0.92 | 0.82 | 1.11 |
| Civilian Employment | $0.83^{*}$ | $0.51^{*}$ | 0.94 | $0.87^{*}$ | $0.41^{**}$ | 0.76 | 1.01 | 1.00 | 3.88 | 1.07 | 1.34 | 4.55 |
| Unemployment Rate | 0.88 | 1.74 | 1.66 | 0.94 | 1.97 | 1.67 | 1.02 | 1.94 | 2.18 | 1.27 | 2.18 | 2.44 |
| Mean Unemp Duration | 0.66 | 1.07 | 0.87 | 0.67 | 1.06 | 0.84 | 0.93 | 1.21 | 1.46 | 1.08 | 1.53 | 2.07 |
| Initial Claims | 1.83 | $0.65^{***}$ | 5.23 | 1.82 | $0.53^{*}$ | 4.38 | 2.23 | 1.12 | 13.77 | 2.43 | 1.13 | 16.02 |
| Nonfarm Payrolls | 1.42 | 1.04 | 1.70 | 1.49 | 0.95 | 1.47 | 1.59 | 1.99 | 5.82 | 1.67 | 2.71 | 8.04 |
| Hours | 0.55 | $0.44^{***}$ | $0.19^{***}$ | $0.57^{*}$ | $0.39^{***}$ | $0.17^{***}$ | 0.76 | 0.80 | $0.72^{***}$ | 0.89 | 0.96 | 0.84 |
| Hours: Manufacturing | 0.69 | $0.42^{***}$ | $0.20^{***}$ | 0.67 | $0.39^{***}$ | $0.18^{**}$ | 0.79 | $0.67^{**}$ | $0.74^{***}$ | 0.89 | 0.71 | 0.77 |
| Housing Starts | $0.37^{**}$ | $0.25^{**}$ | $0.18^{***}$ | $0.36^{***}$ | $0.26^{**}$ | $0.16^{***}$ | $0.54^{*}$ | 0.38 | 0.34 | $0.62^{*}$ | 0.65 | $0.52^{***}$ |
| Building Permits | $0.37^{***}$ | $0.30^{*}$ | $0.22^{***}$ | $0.36^{***}$ | $0.31^{*}$ | $0.21^{***}$ | $0.49^{**}$ | 0.39 | 0.38 | 0.69 | 0.61 | $0.53^{**}$ |
| Real Mfg & Trade Sales | 0.97 | $0.43^{***}$ | $0.21^{***}$ | 0.92 | $0.35^{**}$ | $0.17^{**}$ | 1.15 | 0.70 | $0.84^{***}$ | 1.26 | 0.94 | $0.92^{***}$ |
| Durable Goods Orders | $0.36^{*}$ | $0.21^{***}$ | $0.13^{***}$ | $0.36^{*}$ | $0.17^{**}$ | $0.11^{**}$ | $0.56^{*}$ | $0.40^{**}$ | $0.53^{***}$ | 0.76 | $0.58^{***}$ | $0.70^{***}$ |
| Business Inventories | $0.45^{***}$ | 0.63 | $0.36^{***}$ | $0.44^{***}$ | 0.58 | $0.31^{**}$ | $0.61^{***}$ | 1.05 | 1.56 | 0.89 | 1.17 | 1.77 |
| M1 Money Supply | 2.16 | 3.08 | 2.49 | 2.14 | 2.42 | 1.97 | 2.30 | 10.13 | 12.86 | 2.59 | 12.85 | 14.70 |
| M2 Money Supply | $0.50^{**}$ | $0.30^{**}$ | $0.22^{***}$ | $0.49^{**}$ | $0.25^{*}$ | $0.19^{*}$ | 0.65 | 0.57 | 0.72 | 1.00 | 1.04 | $0.82^{***}$ |
| Business Loans | $0.63^{*}$ | $0.66^{**}$ | $0.20^{***}$ | $0.64^{*}$ | $0.54^{*}$ | $0.17^{**}$ | $0.72^{*}$ | 0.92 | 0.83 | 0.86 | 0.81 | 0.73 |
| Consumer Credit (NR) | $0.36^{***}$ | $0.28^{***}$ | $0.14^{***}$ | $0.34^{***}$ | $0.27^{**}$ | $0.14^{***}$ | $0.45^{***}$ | $0.49^{**}$ | $0.45^{***}$ | $0.58^{**}$ | $0.67^{***}$ | $0.59^{***}$ |
| Fed Funds Rate | $0.25^{***}$ | $0.27^{**}$ | $0.23^{**}$ | $0.31^{***}$ | $0.27^{**}$ | $0.22^{*}$ | $0.26^{***}$ | $0.31^{***}$ | 0.48 | $0.51^{***}$ | $0.43^{***}$ | 0.64 |
| 3-Mo T-Bill | $0.07^{***}$ | $0.17^{**}$ | $0.27^{*}$ | $0.08^{***}$ | $0.17^{**}$ | $0.27^{*}$ | $0.09^{***}$ | $0.23^{***}$ | 0.56 | $0.13^{***}$ | $0.30^{***}$ | 0.70 |
| 1-Year Treasury | $0.15^{***}$ | $0.24^{*}$ | $0.30^{*}$ | $0.15^{***}$ | $0.26^{*}$ | $0.29^{*}$ | $0.15^{**}$ | $0.34^{*}$ | 0.60 | $0.24^{**}$ | 0.50 | 0.82 |
| 5-Year Treasury | $0.38^{***}$ | $0.43^{*}$ | $0.39^{*}$ | $0.42^{**}$ | 0.42 | $0.37^{*}$ | $0.37^{**}$ | 0.55 | 0.77 | $0.58^{*}$ | 0.86 | 1.09 |
| 10-Year Treasury | $0.41^{***}$ | $0.54^{*}$ | $0.42^{**}$ | $0.45^{***}$ | $0.50^{*}$ | $0.38^{*}$ | $0.37^{**}$ | 0.59 | 0.72 | $0.61^{*}$ | 0.91 | 1.08 |
| Baa Corporate | $0.41^{***}$ | 0.43 | $0.35^{*}$ | $0.38^{***}$ | 0.41 | $0.33^{*}$ | $0.58^{**}$ | 0.56 | 0.57 | 0.90 | 1.01 | 1.02 |
| Baa Spread | $0.37^{***}$ | $0.36^{***}$ | $0.16^{***}$ | $0.38^{***}$ | $0.36^{***}$ | $0.16^{***}$ | $0.38^{***}$ | $0.43^{***}$ | $0.56^{***}$ | $0.68^{**}$ | $0.50^{***}$ | $0.61^{**}$ |
| AAA Spread | $0.33^{***}$ | $0.27^{***}$ | $0.13^{***}$ | $0.30^{***}$ | $0.26^{***}$ | $0.13^{***}$ | $0.38^{***}$ | $0.35^{***}$ | $0.47^{***}$ | $0.67^{**}$ | $0.41^{***}$ | $0.50^{***}$ |
| 5Y-FF Spread | $0.36^{**}$ | $0.40^{*}$ | $0.20^{***}$ | $0.37^{***}$ | $0.37^{*}$ | $0.18^{***}$ | $0.36^{**}$ | $0.43^{*}$ | $0.63^{***}$ | $0.58^{**}$ | $0.53^{***}$ | $0.61^{**}$ |
| 10Y-FF Spread | $0.37^{***}$ | $0.31^{**}$ | $0.14^{***}$ | $0.40^{***}$ | $0.28^{**}$ | $0.14^{***}$ | $0.41^{**}$ | $0.38^{***}$ | $0.50^{***}$ | 0.78 | $0.46^{***}$ | $0.53^{***}$ |
| PPI: Fin Goods | $0.47^{**}$ | $0.60^{**}$ | $0.43^{***}$ | $0.48^{**}$ | $0.54^{*}$ | $0.40^{*}$ | $0.59^{*}$ | 0.98 | 1.28 | 0.77 | 1.27 | 1.31 |
| PCE Prices | $0.37^{***}$ | $0.60^{*}$ | $0.38^{***}$ | $0.38^{***}$ | 0.54 | $0.34^{**}$ | $0.43^{**}$ | 0.99 | 1.26 | $0.65^{*}$ | 1.33 | 1.35 |
| CPI: All Items | $0.45^{**}$ | $0.56^{*}$ | $0.37^{***}$ | $0.44^{**}$ | 0.51 | $0.33^{**}$ | $0.46^{*}$ | 0.96 | 1.24 | 0.70 | 1.24 | 1.30 |
| Oil Price | $0.59^{*}$ | $0.40^{***}$ | $0.21^{***}$ | $0.56^{*}$ | $0.34^{**}$ | $0.18^{**}$ | 0.69 | $0.73^{*}$ | $0.90^{*}$ | 0.91 | 0.88 | 0.95 |
| PPI: Metals & Metal Prod | $0.35^{**}$ | $0.42^{**}$ | $0.28^{***}$ | $0.32^{**}$ | $0.36^{*}$ | $0.25^{**}$ | $0.45^{**}$ | 0.68 | 1.00 | 0.65 | 0.87 | 1.05 |
| Hourly Earnings | $0.23^{***}$ | $0.27^{***}$ | $0.17^{***}$ | $0.22^{***}$ | $0.24^{**}$ | $0.15^{**}$ | $0.32^{*}$ | $0.58^{*}$ | 0.72 | 0.38 | 0.87 | 0.98 |
| S&P 500 | $0.20^{*}$ | $0.28^{**}$ | $0.20^{***}$ | $0.20^{*}$ | $0.24^{**}$ | $0.17^{**}$ | 0.24 | $0.49^{*}$ | 0.84 | 0.42 | 0.74 | 0.92 |
| S&P Div Yield | $0.14^{**}$ | $0.25^{***}$ | $0.19^{***}$ | $0.14^{**}$ | $0.24^{**}$ | $0.17^{***}$ | $0.17^{*}$ | $0.33^{***}$ | $0.37^{**}$ | $0.28^{*}$ | $0.51^{***}$ | $0.54^{***}$ |
| VIX | $0.29^{***}$ | $0.28^{***}$ | $0.16^{***}$ | $0.26^{***}$ | $0.24^{**}$ | $0.13^{**}$ | $0.42^{***}$ | $0.49^{***}$ | $0.59^{***}$ | $0.69^{*}$ | $0.67^{**}$ | $0.70^{**}$ |
| USD/GBP FX | $0.17^{***}$ | $0.22^{**}$ | $0.17^{***}$ | $0.17^{***}$ | $0.21^{**}$ | $0.15^{**}$ | $0.24^{**}$ | $0.43^{*}$ | $0.76^{*}$ | $0.34^{**}$ | $0.64^{**}$ | 0.89 |
| JPY/USD FX | $0.11^{***}$ | $0.26^{**}$ | $0.22^{***}$ | $0.12^{***}$ | $0.23^{**}$ | $0.19^{**}$ | $0.17^{**}$ | $0.44^{*}$ | 0.81 | $0.28^{***}$ | 0.65 | 0.91 |

Each entry is the ratio of the DPM-OISV forecast score to that of the indicated benchmark model. Values below 1 favor DPM-OISV, while values above 1 favor the benchmark. Asterisks denote Diebold-Mariano-West tests based on Newey-West HAC standard errors: $^{***}$ $p < 0.001$, $^{**}$ $0.001 \leqslant p < 0.01$, and $^{*}$ $0.01 \leqslant p < 0.05$.

Table 11: Ratios of CRPS for the competing models relative to DPM-OISV over the post-COVID sample at forecast horizons $h = 1, 6, 12$. Values below 1 favor DPM-OISV.

| Variable | SVo | | | SVO | | | SV-$t$ | | | DPM-CSV | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 6 | 12 | 1 | 6 | 12 | 1 | 6 | 12 | 1 | 6 | 12 |
| Real Income | 1.09 | $0.86^{*}$ | $0.53^{***}$ | 1.05 | 0.90 | $0.56^{**}$ | 1.19 | 1.14 | 1.23 | 1.27 | 1.07 | 1.10 |
| Real Income ex Transf | 0.92 | $0.66^{***}$ | $0.35^{***}$ | $0.87^{*}$ | $0.69^{***}$ | $0.37^{***}$ | 0.97 | $0.91^{*}$ | 0.97 | 1.01 | $0.89^{*}$ | 0.89 |
| Real Consumption | 0.85 | $0.76^{***}$ | $0.57^{***}$ | 0.85 | $0.77^{***}$ | $0.59^{***}$ | 0.89 | $0.88^{***}$ | 0.97 | 0.84 | $0.82^{***}$ | $0.88^{***}$ |
| IP | $0.85^{*}$ | $0.89^{*}$ | $0.60^{***}$ | $0.85^{*}$ | $0.88^{*}$ | $0.63^{***}$ | 0.94 | 0.99 | 1.01 | 0.90 | 0.94 | 0.93 |
| Capacity Util | 0.91 | $0.61^{*}$ | $0.39^{***}$ | 0.93 | $0.62^{*}$ | $0.41^{***}$ | 0.96 | 0.80 | $0.66^{*}$ | 0.94 | $0.67^{*}$ | $0.51^{***}$ |
| Civilian Employment | 0.91 | $0.86^{*}$ | $0.70^{**}$ | 0.92 | $0.88^{**}$ | $0.72^{**}$ | 0.96 | 0.97 | 1.04 | 0.95 | 0.95 | 0.98 |
| Unemployment Rate | 0.97 | 0.93 | $0.62^{***}$ | 0.96 | 0.93 | $0.64^{***}$ | 0.98 | 0.99 | 1.03 | 0.91 | 0.95 | $0.85^{***}$ |
| Mean Unemp Duration | 0.96 | 0.96 | $0.88^{**}$ | 0.95 | 0.96 | $0.90^{*}$ | 0.92 | 1.07 | 1.09 | 0.95 | 1.05 | 1.04 |
| Initial Claims | $0.87^{*}$ | $0.89^{*}$ | 1.73 | $0.87^{*}$ | $0.89^{*}$ | 1.77 | 0.94 | 1.02 | 2.65 | 0.91 | 0.99 | 2.54 |
| Nonfarm Payrolls | 0.97 | $0.72^{***}$ | $0.43^{***}$ | 0.97 | $0.75^{***}$ | $0.46^{***}$ | $0.85^{*}$ | $0.79^{*}$ | $0.79^{***}$ | $0.82^{***}$ | $0.74^{***}$ | $0.72^{***}$ |
| Hours | 1.09 | 1.35 | 0.97 | 1.11 | 1.39 | 1.02 | 1.10 | 1.41 | 1.61 | 1.01 | 1.11 | 1.05 |
| Hours: Manufacturing | 0.90 | 1.05 | 0.74 | 0.90 | 1.06 | 0.76 | 0.94 | 1.13 | 1.44 | 0.96 | 1.11 | 1.16 |
| Housing Starts | $0.83^{**}$ | $0.71^{**}$ | $0.43^{***}$ | $0.82^{**}$ | $0.69^{**}$ | $0.43^{***}$ | $0.86^{***}$ | $0.83^{*}$ | $0.69^{*}$ | $0.84^{***}$ | $0.75^{*}$ | $0.56^{*}$ |
| Building Permits | $0.81^{**}$ | $0.66^{***}$ | $0.41^{***}$ | $0.81^{**}$ | $0.64^{***}$ | $0.41^{***}$ | 0.91 | 0.86 | 0.71 | 0.88 | $0.73^{*}$ | $0.55^{*}$ |
| Real Mfg & Trade Sales | 0.82 | $0.80^{***}$ | $0.55^{***}$ | 0.81 | $0.80^{***}$ | $0.56^{***}$ | 0.90 | $0.90^{**}$ | 0.95 | 0.88 | $0.87^{***}$ | $0.86^{***}$ |
| Durable Goods Orders | $0.78^{***}$ | $0.85^{***}$ | $0.74^{***}$ | $0.79^{***}$ | $0.85^{***}$ | $0.76^{***}$ | $0.86^{***}$ | 0.96 | $0.97^{*}$ | $0.83^{***}$ | $0.91^{*}$ | $0.93^{*}$ |
| Business Inventories | 0.80 | $0.66^{***}$ | $0.47^{***}$ | 0.77 | $0.65^{***}$ | $0.48^{***}$ | 0.92 | 0.92 | 1.05 | 0.90 | $0.87^{*}$ | 0.95 |
| M1 Money Supply | 1.18 | 1.04 | 2.49 | 1.22 | 1.06 | 2.61 | 1.07 | 1.07 | 3.88 | 1.05 | 1.07 | 3.63 |
| M2 Money Supply | $0.82^{**}$ | 0.82 | $0.60^{***}$ | $0.81^{**}$ | $0.79^{*}$ | $0.59^{***}$ | $0.88^{**}$ | 0.93 | 0.90 | $0.82^{**}$ | 0.88 | 0.83 |
| Business Loans | 0.91 | $0.88^{*}$ | $0.66^{***}$ | 0.93 | $0.90^{*}$ | $0.68^{***}$ | 0.95 | 0.95 | 0.96 | 0.95 | 0.99 | 0.95 |
| Consumer Credit (NR) | $0.85^{*}$ | $0.78^{**}$ | $0.56^{***}$ | 0.92 | $0.84^{**}$ | $0.59^{***}$ | $0.91^{**}$ | $0.89^{*}$ | $0.85^{*}$ | $0.90^{*}$ | $0.86^{**}$ | $0.77^{*}$ |
| Fed Funds Rate | 1.09 | 1.05 | $0.63^{**}$ | 1.12 | 1.02 | $0.63^{**}$ | 0.88 | 1.03 | 0.92 | 0.86 | 0.93 | 0.83 |
| 3-Mo T-Bill | 0.86 | 1.12 | 0.82 | 0.88 | 1.12 | 0.83 | $0.77^{***}$ | 1.14 | 1.06 | $0.72^{***}$ | 1.03 | 0.95 |
| 1-Year Treasury | 0.99 | 1.13 | 0.77 | 1.01 | 1.11 | 0.76 | 0.94 | 1.19 | 1.10 | 0.89 | 1.07 | 0.96 |
| 5-Year Treasury | 1.09 | 1.26 | $0.81^{*}$ | 1.13 | 1.20 | $0.80^{**}$ | 1.14 | 1.28 | 1.19 | 1.08 | 1.09 | 0.97 |
| 10-Year Treasury | 1.09 | 1.27 | 0.86 | 1.12 | 1.20 | 0.85 | 1.15 | 1.30 | 1.20 | 1.14 | 1.13 | 1.01 |
| Baa Corporate | 1.07 | 1.11 | $0.78^{**}$ | 1.09 | 1.07 | $0.79^{**}$ | 1.10 | 1.15 | 1.15 | 1.11 | 1.06 | 0.98 |
| Baa Spread | 1.14 | $0.74^{*}$ | $0.44^{***}$ | 1.19 | 0.75 | $0.46^{***}$ | 1.14 | 0.82 | $0.73^{***}$ | 1.14 | $0.77^{*}$ | $0.68^{***}$ |
| AAA Spread | 1.04 | $0.73^{*}$ | $0.43^{***}$ | 1.07 | $0.75^{*}$ | $0.45^{***}$ | 1.08 | 0.84 | $0.84^{*}$ | 1.12 | $0.78^{*}$ | $0.72^{***}$ |
| 5Y-FF Spread | 1.18 | 0.98 | $0.39^{***}$ | 1.26 | 0.99 | $0.41^{***}$ | 1.18 | 0.98 | $0.74^{***}$ | 1.15 | 0.92 | $0.68^{***}$ |
| 10Y-FF Spread | 1.17 | 0.84 | $0.38^{***}$ | 1.24 | 0.84 | $0.40^{***}$ | 1.16 | 0.91 | $0.76^{*}$ | 1.18 | 0.87 | $0.70^{***}$ |
| PPI: Fin Goods | 0.99 | $0.92^{*}$ | $0.67^{***}$ | 0.99 | 0.93 | $0.69^{***}$ | 1.01 | 1.01 | 1.00 | 0.97 | 1.01 | 0.99 |
| PCE Prices | 0.99 | $0.71^{***}$ | $0.44^{***}$ | 0.98 | $0.72^{***}$ | $0.45^{***}$ | 0.97 | 0.94 | $0.90^{*}$ | 1.00 | $0.91^{*}$ | $0.82^{***}$ |
| CPI: All Items | 0.93 | $0.70^{***}$ | $0.40^{***}$ | 0.91 | $0.71^{***}$ | $0.41^{***}$ | 1.00 | $0.92^{***}$ | $0.94^{**}$ | 0.96 | $0.87^{***}$ | $0.83^{***}$ |
| Oil Price | 0.99 | $0.83^{***}$ | $0.56^{***}$ | 0.96 | $0.86^{***}$ | $0.58^{***}$ | 0.94 | 0.98 | 0.97 | 0.94 | 0.97 | $0.91^{***}$ |
| PPI: Metals & Metal Prod | $0.90^{**}$ | $0.86^{**}$ | $0.59^{***}$ | $0.90^{**}$ | $0.87^{*}$ | $0.60^{***}$ | 0.94 | 0.98 | $0.91^{*}$ | $0.89^{*}$ | 0.95 | $0.87^{***}$ |
| Hourly Earnings | $0.74^{**}$ | $0.75^{***}$ | $0.50^{***}$ | $0.73^{**}$ | $0.75^{***}$ | $0.51^{***}$ | $0.77^{***}$ | $0.83^{**}$ | $0.86^{***}$ | $0.75^{**}$ | $0.82^{***}$ | $0.80^{***}$ |
| S&P 500 | 0.93 | 0.94 | $0.63^{***}$ | 0.93 | 0.95 | $0.64^{***}$ | 0.98 | 1.05 | 1.02 | 1.00 | 1.03 | 0.99 |
| S&P Div Yield | $0.73^{***}$ | $0.65^{***}$ | $0.46^{***}$ | $0.74^{***}$ | $0.62^{***}$ | $0.46^{***}$ | $0.85^{*}$ | 0.94 | 1.09 | $0.79^{**}$ | $0.73^{***}$ | $0.76^{**}$ |
| VIX | 1.00 | 0.94 | 0.73 | 1.00 | 0.91 | 0.74 | 1.02 | 1.23 | 1.57 | 1.05 | 1.31 | 1.49 |
| USD/GBP FX | $0.88^{*}$ | 0.83 | $0.48^{***}$ | $0.89^{*}$ | 0.84 | $0.50^{***}$ | 0.97 | 0.95 | $0.92^{*}$ | 0.93 | $0.90^{*}$ | $0.85^{***}$ |
| JPY/USD FX | $0.90^{*}$ | $0.89^{*}$ | $0.63^{***}$ | $0.91^{*}$ | $0.89^{*}$ | $0.65^{***}$ | 0.96 | 0.97 | 1.00 | 0.93 | 0.94 | $0.95^{**}$ |

Each entry is the ratio of the DPM-OISV forecast score to that of the indicated benchmark model. Values below 1 favor DPM-OISV, while values above 1 favor the benchmark. Asterisks denote Diebold-Mariano-West tests based on Newey-West HAC standard errors: $^{***}$ $p < 0.001$, $^{**}$ $0.001 \leqslant p < 0.01$, and $^{*}$ $0.01 \leqslant p < 0.05$.